

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
DEPARTAMENTO DE ESTATÍSTICA  
CURSO DE ESTATÍSTICA

Liliane da Silva Oliveira

**Um Estudo Empírico do Desempenho de  
Combinações de Previsões**

Juiz de Fora

2011

Leiliane da Silva Oliveira

*Um Estudo Empírico do Desempenho de Combinações de  
Previsões*

Monografia apresentada ao Curso de Estatística  
da UFJF, como requisito para a obtenção do grau  
de Bacharel em Estatística.

**Orientador: Henrique Steinherz Hippert**

**Doutor em Engenharia Elétrica-PUC/Rio**

Juiz de Fora

2011

Oliveira, Leiliane

Um Estudo Empírico do Desempenho de Combinações de Previsões

/ Leiliane Oliveira - 2011

63 .p

1.Previsão.2.Combinação de Previsão. I.Título.

CDU N/A

Leiliane da Silva Oliveira

*Um Estudo Empírico do Desempenho de Combinações de  
Previsões*

Monografia apresentada ao Curso de Estatística  
da UFJF, como requisito para a obtenção do grau  
de Bacharel em Estatística.

Aprovado em 28 de Novembro de 2011.

**BANCA EXAMINADORA**

---

Henrique Steinherz Hippert

Doutor em Engenharia Elétrica-PUC/Rio

---

Marcel de Toledo Vieira

Ph.D. em Estatística - University of Southampton

---

Ronaldo Rocha Bastos

Ph.D. em Urban and Regional Planning - Liverpool University

*Aos meus queridos pais.*

*“Somos o que somos, por sermos assim:  
JOVENS!”*

*Silamara Ferreira*

*“Todos os modelos são errados, mas alguns são úteis.”*

*George Box*

*“Não se conhece completamente uma ciência,  
a menos que se saiba sua história”*

*Auguste Comte<sup>1</sup>*

## Agradecimentos

À Deus, pela proteção e pela oportunidade de ter realizado mais um objetivo da minha vida.

Ao meu pai, por todo o amor, companheirismo, compreensão, carinho e dedicação, pelo esforço contínuo e trabalho árduo para a concretização deste sonho.

À minha mãe, pelo amor incondicional, por sempre estar ao meu lado e torcer por mim em todos os momentos, pelos conselhos e compreensão, e pela paciência nos momentos mais difíceis.

Ao Daniel, meu porto seguro, pelos incontáveis momentos de aprendizagem e alegria, pelo companheirismo, compreensão, incentivo, amor e carinho, essenciais para a minha completa felicidade.

Ao meu irmão, por me ajudar nos momentos em que os dias pareciam pequenos.

À minha querida amiga, Silamara, que esteve ao meu lado nos momentos alegres e nos mais difíceis, pelos conselhos e compreensão que teve com a minha ausência, e que em nada abalou nossa amizade.

À Anna Cláudia, por ter me apresentado a estatística, que eu amo e com o que eu quero trabalhar para o resto de minha vida

À Raquel Borges e Sarah, companheiras fiéis da faculdade, pelo companheirismo, carinho, incentivo, loucuras e fé mútua nos momentos desesperadores do ICE.

Aos queridos colegas de faculdade, especialmente a Raquel Morani, Vanessa Morani, Carlos Eduardo e Vanessa Félix, pelos incontáveis momentos de alegria e loucuras no ICE.

Aos meus antigos professores, Solange, Irma, Dinoléia, Meiri e Mauro, por

terem despertado em mim o gosto pelo conhecimento.

Ao professor Clécio, pelas aulas desesperadoras e, ao mesmo tempo engraçadas, que me foram muito úteis, pelos conhecimentos repassados na Iniciação Científica, e pela completa dedicação ao curso.

Meu agradecimento especial ao meu orientador, professor Henrique Hippert, por ser um exemplo de profissional e pessoa, pelo envolvimento e supervisão deste trabalho, e principalmente pela confiança depositada em mim.

A todos os amigos e todas as pessoas que por mim torceram, e que de uma forma ou de outra contribuíram para a concretização deste trabalho.



## Resumo

A combinação de previsões é um procedimento que tenta melhorar a precisão das previsões, que aproveita a disponibilidade de vários métodos de previsão individual. Este trabalho tem como objetivo aplicar os conhecimentos aprendidos durante o curso e fazer um estudo sobre o desempenho de diferentes combinações de previsões. Realizamos uma comparação entre métodos de previsão individuais e métodos de combinação, escolhendo para isso apenas uma medida de desempenho, o erro quadrático médio (MSE). As previsões individuais foram feitas pelo método de Amortecimento Exponencial de Holt, por modelos ARIMA e por modelos estruturais de tendência local. Realizamos seis tipos de combinação: média simples, mediana, média dos extremos, regressão, combinação ótima e combinação ótima com independência. Para verificar o melhor desempenho das combinações de previsões, ajustamos diversos métodos em dez séries reais, estacionárias e não-sazonais.

**Palavras-chave:** Amortecimento Exponencial, Combinação de Previsões, Métodos de Previsão, Modelos ARIMA, Modelos Estruturais, Previsão, Séries Temporais.

# Abstract

The combination of forecasts is a procedure that attempts to improve their precision and which takes advantage of the availability of many individual forecast methods. This paper intends to apply the acquired theoretical knowledge throughout the course and thus make a study about the performance of different forecast combinations. We compared individual forecast methods and combination methods, using only one performance measure, the Mean Square Error (MSE). The individual forecasts were made through the Holt Exponential Smoothing, as well as through ARIMA models and structural models of local tendency. We did six types of combination: simple average, median, extreme average, regression, optimal combination and optimal combination with independency. To check the best performance of forecast combinations, we fitted several methods in ten real series, stationary and non-seasonal.

**Keywords:** Exponential Smoothing, Forecasts combination, Forecast methods, ARIMA models, Structural methods, Forecast, Time Series.

# Sumário

<b>Lista de Figuras</b>	<b>9</b>
<b>Lista de Tabelas</b>	<b>11</b>
<b>1 Introdução</b>	<b>12</b>
<b>2 Fundamentação Teórica</b>	<b>15</b>
2.1 Técnicas Descritivas . . . . .	16
2.2 Métodos e Modelos de Previsão . . . . .	19
2.2.1 Naïve, Média Global e Média Móvel . . . . .	20
2.2.2 Amortecimento exponencial . . . . .	21
2.2.3 ARIMA . . . . .	24
2.2.4 Modelo Estrutural . . . . .	29
2.3 Combinação de Previsões . . . . .	30
2.4 Medidas de Erros de Previsão . . . . .	32
<b>3 Metodologia</b>	<b>34</b>
<b>4 Resultados</b>	<b>38</b>
4.1 Análise Descritiva . . . . .	38
4.2 Modelos de Previsões Individuais . . . . .	44
4.2.1 Amortecimento Exponencial de Holt . . . . .	44

4.2.2	ARIMA . . . . .	45
4.2.3	Modelo Estrutural . . . . .	47
4.3	Combinação de Previsões . . . . .	47
4.4	Resultados na Amostra de Teste . . . . .	49
<b>5</b>	<b>Discussão e Conclusão</b>	<b>53</b>
<b>A</b>	<b>Anexos</b>	<b>55</b>
A.1	Função no R . . . . .	55
A.2	Modelos ARIMA . . . . .	58
	<b>Referências Bibliográficas</b>	<b>60</b>

## Lista de Figuras

2.1	Séries simuladas com tendência constante, linear e quadrática . . . . .	17
2.2	Decomposição Clássica da Série Airline . . . . .	18
2.3	Correlograma . . . . .	18
2.4	Efeito do valor dos pesos . . . . .	22
4.1	Gráfico de linha, FAC e FACP da série Cow . . . . .	38
4.2	Gráfico de linha, FAC e FACP da série Buffsnow . . . . .	39
4.3	Gráfico de linha, FAC e FACP da série Calfem . . . . .	39
4.4	Gráfico de linha, FAC e FACP da série DowJones . . . . .	40
4.5	Gráfico de linha, FAC e FACP da série DowJones Diferenciada . . . . .	40
4.6	Gráfico de linha, FAC e FACP da série Huron . . . . .	40
4.7	Gráfico de linha, FAC e FACP da série Globtp . . . . .	41
4.8	Gráfico de linha, FAC e FACP da série Globtp Diferenciada . . . . .	41
4.9	Gráfico de linha, FAC e FACP da série Sheep . . . . .	42
4.10	Gráfico de linha, FAC e FACP da série Sheep Diferenciada . . . . .	42
4.11	Gráfico de linha, FAC e FACP da série Summer . . . . .	42
4.12	Gráfico de linha, FAC e FACP da série Ausgundeaths . . . . .	43
4.13	Gráfico de linha, FAC e FACP da série Ausgundeaths Diferenciada . . . . .	43
4.14	Gráfico de linha, FAC e FACP da série Carga . . . . .	43
4.15	Gráfico de linha, FAC e FACP da série Carga Diferenciada . . . . .	44

4.16	Previsão por AEH da série Sheep . . . . .	45
4.17	Previsão por AEH da série Globtp . . . . .	45
4.18	FAC dos erros de previsão por ARIMA da Série Cow . . . . .	46
4.19	Previsão por ARIMA da série Ausgundeaths . . . . .	46
4.20	Previsão por ARIMA da série Summer . . . . .	46
4.21	Previsão por Modelos Estruturais da série Huron. . . . .	48
4.22	Previsão por Modelos Estruturais da série DowJones. . . . .	48
4.23	Parâmetros da Combinação Ótima e Ótima com Independência da Série Buffsnow . . . . .	50
4.24	Parâmetros de Combinação por Regressão da Série Buffsnow . . . . .	50
4.25	MSE: Cow. . . . .	52
4.26	MSE: Buffsnow. . . . .	52
4.27	MSE: Calfem. . . . .	52
4.28	MSE: DowJones. . . . .	52
4.29	MSE: Globtp. . . . .	52
4.30	MSE: Huron. . . . .	52
4.31	MSE: Sheep. . . . .	52
4.32	MSE: Summer. . . . .	52
4.33	MSE: Ausgundeaths. . . . .	52
4.34	MSE: Carga. . . . .	52

## Lista de Tabelas

2.1	Padrões teóricos da FAC e FACP para modelos ARMA (Fonte: [Hippert, 2010]) . . . . .	28
2.2	Medidas de Erros de Previsão . . . . .	32
4.1	Parâmetros otimizados do modelo AEH . . . . .	45
4.2	Modelos ARIMA . . . . .	46
4.3	Variâncias dos Modelos Estruturais . . . . .	47
4.4	Parâmetros da Combinação por Regressão. . . . .	49
4.5	Pesos da Combinação Ótima. . . . .	49
4.6	Pesos da Combinação Ótima com Independência. . . . .	49
4.7	MSE na série de teste. . . . .	51

# 1 Introdução

Uma *série temporal* é qualquer conjunto de observações ordenadas no tempo. Os índices diários da Bolsa de Valores de São Paulo e os índices pluviométricos mensais de uma cidade são exemplos de séries temporais.

A *análise* e a *previsão* das séries temporais englobam um conjunto de técnicas importantes em várias áreas do conhecimento. As técnicas de *análise* permitem investigar o mecanismo gerador de uma série, descrever o seu comportamento e procurar periodicidades relevantes nos dados. As técnicas de *previsão* procuram construir modelos matemáticos a partir dos quais seja possível prever valores futuros da série; previsões nos permitem preparar para o que nos espera no futuro, fazer planos, e tomar decisões.

Existem várias técnicas de se fazer previsão de séries temporais, dependendo do tipo de série em estudo. O comportamento das séries encontradas no mundo real pode ser muito variável. Algumas séries podem seguir um nível constante, ou seja, variar em torno de uma média central. Outras podem exibir um nível crescente ou decrescente, caracterizando uma série temporal com *tendência*. Algumas séries exibem comportamento de repetição a períodos regulares de tempo, a chamada *sazonalidade*; por exemplo, é natural esperar que as vendas mensais de brinquedos tenham um pico no mês de dezembro e talvez um pico secundário em outubro (outros exemplos de séries com sazonalidade seriam a temperatura mensal e o consumo de energia elétrica). Para cada tipo de série, têm sido desenvolvidas técnicas de previsão adequadas.

Dentre os métodos de previsão existentes, de ampla aplicação, os mais difundidos são provavelmente os métodos de *amortecimento exponencial* e os modelos ARIMA. Os primeiros, de base empírica, são de aplicação mais simples, pois não exigem grandes conhecimentos em matemática. Os segundos, também conhecidos como metodologia de [Box e Jenkins, 1994], são mais complicados, e exigem do usuário um conhecimento mais profundo da Estatística. Ambos são bastante eficientes, o que explica sua grande



---

utilização. Os *modelos estruturais*, propostos como alternativa aos ARIMA, são menos utilizados na prática, embora apresentem algumas vantagens [Morettin e Tolo, 2006].

Diante da diversidade de métodos ou modelos que podem ser usados para fazer previsões de uma mesma série, é preciso definir critérios para a escolha. Um destes critérios, que será usado neste trabalho, é a avaliação empírica da acurácia, através de medidas de erro, como o Erro Quadrático Médio (MSE). Outros critérios (como por exemplo o custo) podem também vir a ser usados, em aplicações reais.

Cada modelo de previsão retira informações distintas dos dados; dessa forma, surge a idéia de que, combinando essas previsões individuais, poderíamos talvez obter previsões mais precisas. A *combinação de previsões* é portanto um procedimento que visa melhorar a precisão das previsões, aproveitando a disponibilidade de vários métodos de previsão individual. De acordo com [Menezes et al., 2000], muitos estudos empíricos e teóricos têm mostrado que o desempenho de combinações de previsões é melhor que o de modelos de previsão individual.

O objetivo deste trabalho é verificar empiricamente o desempenho de diferentes combinações de previsões, comparado ao das técnicas de previsão individual. A intenção de se estudar o desempenho das combinações de previsões é a de procurar estabelecer bases teóricas ou empíricas que justifiquem o desempenho das combinações, e permitam, por exemplo, determinar quais combinações são melhores para a previsão de cada tipo de série.

Fizemos para isto um estudo empírico do desempenho de diversas combinações de previsões, aplicadas sobre um conjunto de séries reais não-sazonais, possuindo um número de observações em torno de 100. Como forma de verificação da adequação das combinações, usamos apenas uma medida de erro de previsão, o Erro Quadrático Médio (MSE), que é a medida que tem sido usada na maioria dos estudos feitos sobre combinação de previsão [Menezes et al., 2000].

Iniciamos este texto com uma revisão bibliográfica das técnicas de previsão individual e de combinação de previsões, no Capítulo 2. Em seguida, fizemos previsões aplicando diversos métodos de previsão individual, e calculamos diversas combinações

destas previsões. Para a análise do desempenho, calculamos os erros de previsão e retiramos algumas medidas destes. A metodologia de análise dos dados e dos modelos de previsão é apresentada no Capítulo 3. Os resultados obtidos, assim como alguns comentários são apresentados no Capítulo 4. Por fim, apresentamos nossas conclusões e algumas sugestões de possíveis estudos futuros no Capítulo 5.

## 2 Fundamentação Teórica

Nesta seção apresentaremos alguns conceitos básicos sobre séries temporais; as referências serão [Morettin e Toloi, 2006] e [Ehlers, 2011].

Uma série temporal é qualquer conjunto de observações feitas sequencialmente ao longo do tempo. Uma característica importante deste tipo de dados é o fato de serem auto-correlacionados; como a maioria dos métodos estatísticos foi desenvolvida para a análise de dados independentes, o estudo de séries temporais requer técnicas específicas, que levam em consideração a ordem temporal das observações. Uma série temporal pode ser dita *discreta*, quando geralmente as observações são equiespaçadas no tempo; ou *contínua*, quando são feitas continuamente no tempo. Uma série temporal contínua pode ser discretizada, ou seja, seus valores são registrados a certos intervalos de tempo.

O estudo de séries temporais tem como objetivo investigar o processo gerador da série temporal; descrever as propriedades da série, verificando a existência de tendência e sazonalidade; fazer previsões de valores futuros, a curto ou longo prazo; e explicar a variação de uma série através da variação de outra série.

Existe um número muito grande de modelos distintos para analisar o comportamento de uma série temporal. A escolha de um modelo adequado depende de vários fatores, como o conhecimento *a priori* da natureza dos dados e o objetivo da análise. Dessa forma, é interessante que se faça primeiramente uma análise descritiva dos dados antes de se ajustar qualquer modelo (ver seção 2.1).

Uma das formas de utilização de um modelo ajustado à série temporal é para fazer previsões de valores futuros, permitindo o fornecimento de informações para tomadas de decisão, visando determinados objetivos. A escolha do melhor modelo é contudo difícil, e existem várias formas de se avaliar a adequação do modelo aos dados, baseadas nos erros de previsão. Uma opção é combinar previsões feitas por modelos distintos; acredita-se que, retirando informações de vários modelos de previsão e

combinando-os, podem-se compensar as fraquezas de cada um e assim obter uma melhor previsão, ou seja, menores erros de previsão [Lemke e Gabrys, 2010].

## 2.1 Técnicas Descritivas

A primeira ferramenta para a análise de séries temporais é o gráfico dos dados ao longo do tempo, muito útil para a identificação de pontos discrepantes e dos componentes de uma série temporal: *tendência*, *ciclo* e *sazonalidade* [Morettin e Tolo, 2006].

Não existe uma definição comum para *tendência* entre diferentes autores, mas pode-se dizer que esta reflete a mudança de longo prazo do nível médio da série [Ehlers, 2011]. Nos casos mais comuns, trabalha-se com tendência constante (eq. 2.2), linear (eq. 2.3) ou quadrática (eq. 2.4), como mostrado na Figura 2.1.

$$Z_t = \mu_t + \varepsilon_t, \quad \text{onde:} \quad (2.1)$$

$$\mu_t = a_t \quad \text{ou} \quad (2.2)$$

$$\mu_t = a_t + b_t t \quad \text{ou} \quad (2.3)$$

$$\mu_t = a_t + b_t t + c_t t^2 \quad (2.4)$$

Existem várias formas de se estimar a tendência de uma série, as mais utilizadas [Morettin e Tolo, 2006] consistem em:

- Ajustar uma função do tempo, como um polinômio, uma exponencial ou outra função suave;
- Suavizar, ou filtrar, os valores da série ao redor de um ponto, para estimar a tendência naquele ponto;
- Suavizar os valores da série através de sucessivos ajustes de retas de mínimos quadrados ponderados (“lowess”).

Muitas séries temporais apresentam também um comportamento de repetição espaçado por um determinado período de tempo, a chamada *sazonalidade*. Defini-la é

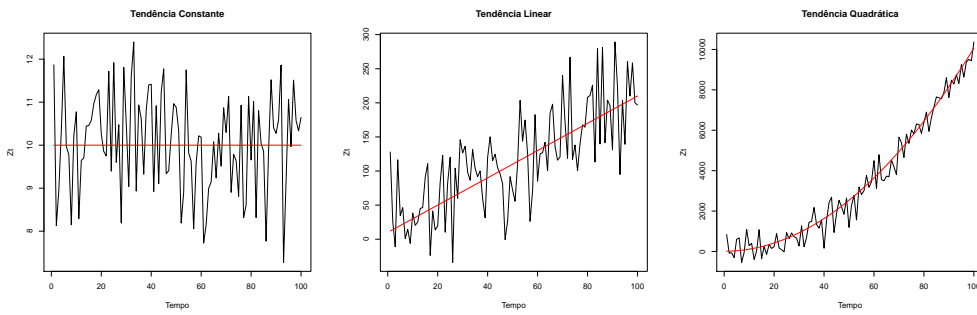


Figura 2.1: Séries simuladas com tendência constante, linear e quadrática

algo bem difícil, tanto conceitualmente como estatisticamente [Morettin e Tolo, 2006], mas podemos dizer que uma série sazonal é aquela em que a média varia ao longo do tempo de acordo com um padrão repetitivo, o qual pode ser previsto [Hippert, 2010]. A *sazonalidade* pode ser aditiva, flutuações sazonais mais ou menos constantes desconsiderando a média global da série; ou multiplicativa, onde as flutuações sazonais têm amplitudes distintas dependendo da média global da série.

Para séries sazonais é interessante fazer a *decomposição clássica* da série, pois dessa forma podem-se identificar muitas de suas propriedades. O modelo para a decomposição tem a forma 2.5; como exemplo, é apresentada a decomposição do logaritmo da série Airline <sup>1</sup>(Figura 2.2).

$$Z_t = \mu_t + S_t + \varepsilon_t \quad (2.5)$$

Na Figura 2.2, o primeiro gráfico mostra a série original; o segundo, o componente sazonal  $S_t$ ; o terceiro o componente de tendência  $\mu_t$ ; e por fim, o resíduo  $\varepsilon_t$  que é uma variável aleatória independente e identicamente distribuída,  $E(\varepsilon_t) = 0$ ;  $V(\varepsilon_t) = \sigma^2$ , com *Distribuição Normal*.

Outras técnicas importantes para a análise de séries temporais são os coeficientes de *autocorrelação* que medem a correlação entre as observações da série temporal em diferentes instantes de tempo. Ou seja, ao invés de termos duas variáveis distintas, X e Y, temos os valores observados de uma mesma variável Z, a série temporal, observados

<sup>1</sup>Disponível em [Hyndman, 2011]

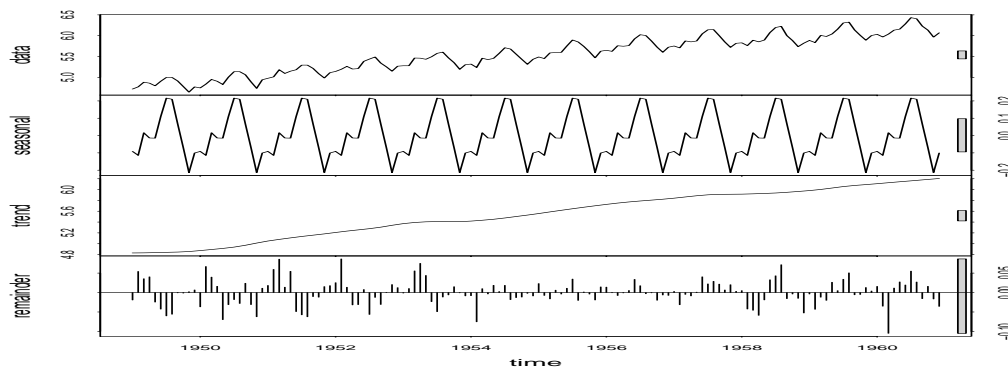


Figura 2.2: Decomposição Clássica da Série Airline

em instantes distintos no tempo  $t$ ,  $Z_t$  e  $Z_{t-k}$ . A fórmula mais simples da autocorrelação amostral [Morettin e Tolo, 2006] é dada por:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^N (z_t - \bar{Z})(z_{t-k} - \bar{Z})}{\sum_{t=1}^N (z_t - \bar{Z})^2} \quad (2.6)$$

O *correlograma* (Figura 2.3), um gráfico de barras mostrando os primeiros  $k$  coeficientes de autocorrelação, também é útil para identificar características da série temporal. Podem associar-se certos padrões do correlograma com certas características de uma série temporal, como aleatoriedade, correlação de curto prazo, correlação negativa, estacionariedade e variação sazonal [Ehlers, 2011]. Esse gráfico também é muito útil para a identificação do tipo de modelo ARIMA, que será apresentado adiante (seção 2.2.3).

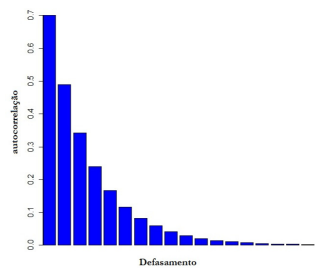


Figura 2.3: Correlograma

## 2.2 Métodos e Modelos de Previsão

Os modelos de previsão são baseados na idéia de que as observações passadas contêm informações sobre o padrão da série. A função desses modelos é distinguir qualquer padrão que esteja presente nas observações, e então usar esse padrão para prever valores futuros da série [Morettin e Toloi, 2006]. Suponha uma série:

$$Z(t) = [z_1, z_2, z_3, \dots, z_t]$$

A previsão do valor do instante  $t + k$  feita no instante  $t$  será denotada por  $\hat{Z}_{t+k}$ , e é definida como a esperança condicional de  $Z_{t+k}$  dado todos os valores passados, isto é:

$$\hat{Z}_{t+k} = E(Z_{t+k} | Z_t, Z_{t-1}, \dots, Z_1) \quad (2.7)$$

Neste trabalho vamos adotar  $k = 1$ , ou seja, todas as previsões serão feitas um passo à frente.

Uma prática comum para verificar o desempenho do modelo de previsão é fazer previsões para todas as observações da série temporal e em seguida compará-las com as observações. Denotaremos a diferença entre as previsões e os valores observados,  $Z_{t+k} - \hat{Z}_{t+k}$ , por  $e_{t+k}$ , o erro de previsão  $k$  passos à frente.

Existem vários tipos de métodos de previsão: *métodos empíricos*, que não exigem o ajuste de modelos de distribuição estatística e *métodos de base estatística*, que são baseados na aplicação de modelos estatísticos cujas propriedades são conhecidas [Hippert, 2010]. Como vimos anteriormente, uma série temporal possui vários componentes; devido a isso, existem modelos que são capazes de “enxergar” esses componentes isoladamente ou conjuntamente. A seguir são apresentados alguns modelos de previsão, com um enfoque maior nos modelos que serão usados neste trabalho.

### 2.2.1 Naïve, Média Global e Média Móvel

Séries não-sazonais e de média  $a$  constante, que podem ser modeladas pela (eq. 2.2), têm os parâmetros do modelo de previsão estimados de forma bem simples, por métodos com forte apelo intuitivo e que não usam técnicas estatísticas (como mínimos quadrados).

O método de previsão *Naïve*, conhecido também por modelo de previsão *ingênuo*, nada mais é do que: “A previsão de amanhã é a observação de hoje”, isto é:

$$\hat{Z}_{t+1} = \hat{a}_t = Z_t \quad (2.8)$$

O método de previsão *Média Global* é baseado na média aritmética de todos os valores passados até o instante  $t$ , isto é:

$$\hat{Z}_{t+1} = \hat{a}_t = \frac{Z_t + Z_{t-1} + \cdots + Z_1}{t} \quad (2.9)$$

O método de previsão *Média Móvel*, consiste em calcular a média aritmética das  $N$  observações mais recente, isto é:

$$\hat{Z}_{t+1} = \hat{a}_t = M_t = \frac{Z_t + Z_{t-1} + \cdots + Z_{t-N+1}}{N} \quad (2.10)$$

Assim,  $M_t$ , é uma estimativa do nível da série que não leva em conta as observações mais antigas, o que é sensato devido ao fato do nível variar suavemente ao longo do tempo. A dificuldade deste método é encontrar o tamanho da “janela”  $N$ , o qual pode ser escolhido como o inteiro positivo que minimiza a soma dos quadrados dos erros de previsão [Barros, 2011]. Na prática esse modelo não é muito utilizado, pois o *Modelo de Amortecimento Exponencial Simples*, que será apresentado a seguir (eq. 2.16), possui a mesma lógica e é mais simples de ser implementado.

Quando a série apresenta uma tendência linear (eq. 2.3), precisamos calcular as “médias móveis duplas” (eq. 2.11), pois temos mais um parâmetro para estimar, o parâmetro  $b$ , que representa a declividade.

$$M_t^{(2)} = \frac{M_t + M_{t-1} + \cdots + M_{t-N+1}}{N} \quad (2.11)$$



Dessa forma, as expressões dos estimadores dos parâmetros  $a$  e  $b$ , e da previsão são:

$$\hat{a}_t = 2M_t - M_t^{(2)} \quad (2.12)$$

$$\hat{b}_t = \frac{2}{N-1}[M_t - M_t^{(2)}] \quad (2.13)$$

$$\hat{Z}_{t+1} = \hat{a}_t + \hat{b}_t \quad (2.14)$$

Seguindo o mesmo raciocínio, quando a série apresenta tendência quadrática (eq. 2.4), precisamos das “médias móveis triplas” (eq. 2.15). Para melhor entendimento deste modelo, consulte [Hippert, 2010].

$$M_t^{(3)} = \frac{M_t^{(2)} + M_{t-1}^{(2)} + \dots + M_{t-N+1}^{(2)}}{N} \quad (2.15)$$

### 2.2.2 Amortecimento exponencial

Existem vários métodos de amortecimento exponencial, aplicáveis a diferentes séries temporais. Eles se baseiam na idéia de média ponderada de todas as observações passadas, usando porém pesos decrescentes: quanto mais velha uma observação, menos peso ela tem.

No método de *Amortecimento Exponencial Simples (AES)*, próprio para modelo constante (eq. 2.2), existe apenas uma constante  $\alpha$ , e os pesos são dados por potências sucessivas desta constante. Isto pode ser expresso da forma:

$$\hat{Z}_{t+1} = \alpha Z_t + \alpha(1-\alpha)Z_{t-1} + \alpha(1-\alpha)^2 Z_{t-2} + \alpha(1-\alpha)^3 Z_{t-3} + \dots + \alpha(1-\alpha)^t Z_1 \quad (2.16)$$

Nesta equação, a previsão é dada pela soma de parcelas da forma  $\phi_k Z_{t-k}$ , onde os coeficientes  $\phi_k$  são funções de potências sucessivas de  $(1-\alpha)$ , ou seja,  $\phi_k = \alpha(1-\alpha)^k$ .

Como  $\alpha$  pode assumir qualquer valor entre 0 e 1, é interessante conhecer seu efeito. Se o valor de  $\alpha$  é pequeno, o decaimento dos pesos  $\phi_k$  é lento, ou seja, a observação mais recente não recebe muito mais peso que as observações mais antigas; se  $\alpha$  é grande, a observação mais recente recebe muito peso, e as antigas pouco [Barros, 2011]; Quando

$\alpha = 1$  recaímos no modelo de previsão de *Naïve*, como podemos observar pela Figura 2.4.

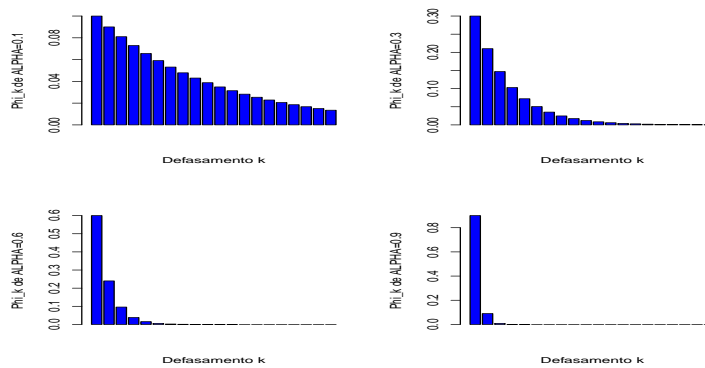


Figura 2.4: Efeito do valor dos pesos

Devido a esses diversos valores que a constante  $\alpha$  pode assumir, é preciso otimizá-la de forma que os erros de previsão sejam mínimos (veja seção 2.4).

Existe uma forma recursiva para calcular a média de forma mais simplificada que a eq. 2.16. A cada instante, a média é dada por uma combinação linear entre a média anterior e a observação mais recente, ou seja:

$$M_t = \alpha Z_t + (1 - \alpha)M_{t-1} \quad \text{onde } M_1 = Z_1 \quad (2.17)$$

Como  $\hat{Z}_{t+1} = M_t$ , podemos reescrever a eq. 2.17 em termos das previsões:

$$\hat{Z}_{t+1} = \alpha Z_t + (1 - \alpha)\hat{Z}_t \quad (2.18)$$

O modelo de *amortecimento exponencial simples* pode, ainda, ser interpretado como uma correção da previsão, a cada passo, pela parcela do último erro observado. Em outras palavras, se o último erro de previsão foi positivo (ou seja, o valor observado foi maior do que o previsto), a próxima previsão será aumentada de uma parcela deste erro; se o erro foi negativo (ou seja, o valor observado foi menor que o previsto), a próxima previsão será reduzida de uma parcela deste erro, [Hippert, 2010]. Isso pode ser expresso pela fórmula a seguir:

$$\hat{Z}_{t+1} = \hat{Z}_t + \alpha(Z_t - \hat{Z}_t) = \hat{Z}_t + \alpha\varepsilon_t \quad (2.19)$$

Para modelos com tendência linear (2.3), existe o *Amortecimento Exponencial de Brown (AEB)* com um parâmetro, que se baseia em médias móveis simples (2.10) e duplas (2.11) calculadas recursivamente pelas equações (2.17) e (2.20).

$$M_t^{(2)} = \alpha M_t + (1 - \alpha)M_{t-1}^{(2)} \quad (2.20)$$

Onde os valores iniciais são:  $M_1 = M_1^{(2)} = Z_1$

Outra possibilidade é *Amortecimento Exponencial de Holt (AEH)*, que será visto com mais detalhes na próxima subseção.

Quando a tendência é quadrática (eq. 2.4), deve-se usar o amortecimento exponencial de Brown triplo com um parâmetro. Neste caso, consulte [Hippert, 2010].

Para séries temporais que apresentam o componente de sazonalidade, o modelo de previsão tem que explicar além deste componente, o nível da série, que pode ser constante ou apresentar uma tendência. A sazonalidade pode ser modelada<sup>2</sup> por modelos aditivos (eq. 2.21) e modelos multiplicativos (eq. 2.22).

$$Z_t = \mu_t + \rho^{m(t)} + \varepsilon_t \quad (2.21)$$

$$Z_t = \mu_t \rho^{m(t)} + \varepsilon_t \quad (2.22)$$

onde  $\mu_t = a$  ou  $\mu_t = a + bt$

### Amortecimento Exponencial de Holt (AEH)

Quando uma série temporal apresenta tendência linear e usa-se AES para prever valores futuros, induz-se um erro sistemático nessas previsões, pois esse tipo de amortecimento segue a tendência com certo atraso [Morettin e Tolo, 2006]. O AEH, que é bastante semelhante ao AES, procura reduzir esse erro. A diferença, agora, consiste nas constantes de amortecimento, que são duas:  $\alpha$  e  $\beta$ , responsáveis pela atualização do nível e da tendência da série, respectivamente.

<sup>2</sup>Para mais informações consulte método de Holt-Winters em [Makridakis et al., 1998]

O AEH é implementado pelas seguintes equações:

$$\hat{a}_t = \alpha Z_t + (1 - \alpha)(\hat{a}_{t-1} + \hat{b}_{t-1}) \quad (2.23)$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1 - \beta)\hat{b}_{t-1} \quad (2.24)$$

$$\hat{Z}_{t+1} = \hat{a}_t + \hat{b}_t \quad (2.25)$$

Esse modelo necessita de valores iniciais, os quais são dados em função das primeiras observações da série. Neste trabalho, adotamos os seguintes valores iniciais [Hippert, 2010]:

$$\begin{aligned} a_1 &= \frac{Z_4 - Z_1}{3} \\ b_1 &= \frac{Z_1 + Z_2 + Z_3 + Z_4}{4} \end{aligned} \quad (2.26)$$

A desvantagem deste modelo é o fato de existirem duas constantes de amortecimento que devem ser otimizadas.

### 2.2.3 ARIMA

Os modelos ARIMA, desenvolvidos por [Box e Jenkins, 1994], têm uma abordagem estatística, isto é, os modelos utilizados para séries temporais são processos controlados por leis probabilísticas. Nesta abordagem, série temporal é definida como uma realização de um processo estocástico (definição 2.2.1, abaixo).

A metodologia de [Box e Jenkins, 1994] consiste em ajustar modelos *auto-regressivos* (eq. 2.30), de *médias móveis* (eq. 2.31), ou uma combinação de ambos, a séries que podem ser estacionárias (modelo ARMA) ou não (modelo ARIMA), sazonais ou não (modelo SARIMA). A construção do modelo é feita de forma iterativa. Primeiramente, faz-se a especificação de uma classe geral de modelos (ARMA, ARIMA ou SARIMA); em segundo lugar, faz-se a identificação do modelo com base na análise da *Função de Autocorrelação* (FAC, eq. 2.6) e da *Função de Autocorrelação Parcial*, vista a seguir (FACP). Em terceiro lugar, vem a estimação dos parâmetros do modelo identificado. Finalmente, faz-se o diagnóstico do modelo ajustado através da análise de resíduos. Se o

modelo encontrado não é adequado, repetem-se os passos anteriores até que se encontre o modelo que forneça os menores erros de previsão [Morettin e Toloi, 2006].

Nos modelos de previsão de amortecimento exponencial adota-se a suposição de que as observações são independentes, mas na prática a maioria das séries temporais apresentam dependência entre seus valores observados, e conseqüentemente os erros não serão independentes. Os modelos ARIMA reproduzem essa correlação existente na série e tentam reproduzir a forma de como os dados mais recentes se relacionam com os mais antigos.

Como os modelos ARIMA tem base estatística, precisamos de alguma definições [Morettin e Toloi, 2006]:

**Definição 2.2.1. Processo Estocástico:** *Seja  $T$  um conjunto de índices e  $A \in \Omega$ . Um processo estocástico indexado por  $T$  com espaço de estados  $A$ , é uma família de variáveis aleatórias  $Z = (Z_t : t \in T)$ , definidas num espaço amostral  $\Omega$  e tomando valores no conjunto  $A$ .*

$t \in T$  pode ser considerado como o instante do tempo em que se observa a variável  $Z_t$ , que é chamada de *espaço de estado* no processo neste instante. Se  $T$  for um conjunto enumerável, então dizemos que o processo estocástico é a *tempo discreto*; caso contrário o processo estocástico é a *tempo contínuo*. Se  $A \subseteq \mathbb{Z}$  então o processo tem espaço de estados discreto; se  $A \subseteq \mathbb{R}$  então o processo tem espaço de estados contínuo.

Sejam  $t_1 < t_2 < \dots < t_n$  elementos quaisquer de  $T$ , temos:

$$F(z_1, \dots, z_n; t_1, \dots, t_n) = P\{Z_{t_1} \leq z_1, \dots, Z_{t_n} \leq z_n\} \quad (2.27)$$

O processo estocástico  $Z = (Z_t : t \in T)$  estará especificado se conhecermos as distribuições finito-dimensionais, dada pela equação 2.27, para todo  $n \geq 1$ .

A inferência sobre a distribuição geradora de uma série temporal requer algumas restrições ao processo estocástico, principalmente a de *estacionariedade*.

**Definição 2.2.2. Processo Estocástico Estritamente Estacionário:** *Um processo*

estocástico  $Z = Z_t : t \in T$ ) é dito estritamente estacionário se as distribuições finito-dimensionais (eq. 2.27) são invariantes no tempo.

Isto implica que a média e a variância do processo são constantes sob translações no tempo, isto é:

$$E_t = \mu, \text{ e} \quad (2.28)$$

$$V_t = \sigma^2, \text{ para todo } t \in T. \quad (2.29)$$

**Definição 2.2.3. Processo Estocástico Estacionário no Sentido Amplo:** Um processo estocástico  $Z = (Z(t) : t \in T)$  é estacionário no sentido amplo ou estacionário de segunda ordem se e somente se:

- $E\{Z_t\} = E_t = \mu$ , constante, para todo  $t \in T$ ;
- $E\{Z_t^2\} = C < \infty$ , para todo  $t \in T$ ;
- $\gamma(t_1, t_2) = \text{Cov}\{Z_{t_1}, Z_{t_2}\}$  é uma função de  $|t_1 - t_2|$ ,

A partir de agora estaremos interessados somente nesta classe de processos, que denominaremos apenas por *processo estacionário*.

Os modelos de [Box e Jenkins, 1994] são casos particulares de um modelo de *filtro linear*<sup>3</sup>. Qualquer série  $Z_t$  que apresente dependência entre seus valores pode ser representada por uma combinação linear de uma série de ruído branco:

$$\varepsilon_t \rightarrow [FT] \rightarrow Z_t$$

Onde  $FT$  é a função de transferência e  $\varepsilon_t \sim N(0, \sigma^2)$ .

Um filtro linear pode ser apresentado na sua forma invertida, ou seja, o ruído branco  $\varepsilon_t$  pode ser representado com função da série temporal  $Z_t$ :

$$\varepsilon_t \leftarrow [FT] \leftarrow Z_t$$

---

<sup>3</sup>Para mais detalhes consulte [Morettin e Tolo, 2006]

Dessa forma, se um processo estacionário pode ser representado por um filtro linear invertido dizemos que o processo é inversível; porém existem algumas condições para a invertibilidade de filtros.

Um processo aleatório é chamado de *processo autoregressivo* de ordem  $p$ , ou  $AR(p)$  se

$$Z_t = \xi + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + \varepsilon_t \quad (2.30)$$

Ou seja, o valor mais recente da série é dado pela combinação linear de valores passados da série, mais um erro aleatório. Pode-se mostrar que, [Box e Jenkins, 1994], um processo  $AR$  sempre é inversível e suas condições de estacionariedade são bem conhecidas, da teoria dos sistemas lineares. Não é possível fazer a identificação de uma provável ordem para esse modelo através de sua FAC, pois esta apresenta decaimento infinito; dessa forma, precisa-se de uma ferramenta que meça as correlações que são definidas diretamente por esse modelo. Trata-se da função de *autocorrelação parcial*<sup>4</sup>. A ordem do modelo  $AR$  é dada pelos  $p$  primeiros valores de defasamento dessa função.

Um processo aleatório é chamado de processo de *médias móveis* de ordem  $q$ , ou  $MA(q)$  se:

$$Z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (2.31)$$

Ou seja, o valor mais recente da série é dado pela combinação linear de erros aleatórios passados. Pode-se mostrar [Box e Jenkins, 1994], que o processo  $MA$  sempre é estacionário e suas condições de inversibilidade são bem conhecidas, da teoria dos sistemas lineares. A identificação de uma provável ordem para modelos  $MA$  é feita pela análise de sua FAC, pois essa só tem os primeiro  $q$  valores de defasamento, já sua FACP apresenta um decaimento infinito.

Ao combinarmos modelos  $AR$  e  $MA$  é possível obter uma representação mais adequada da série temporal com um número menor de parâmetros [Ehlers, 2011]. Um processo aleatório é chamado processo *autoregressivo-médias móveis* de ordem  $p$  e  $q$ , ou

---

<sup>4</sup>Para mais detalhes verifique [Morettin e Tolo, 2006]

$ARMA(p, q)$  se:

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (2.32)$$

As condições de estacionariedade e inversibilidade do modelo  $ARMA$  são as mesmas dos modelos  $AR$  e  $MA$ . A identificação da ordem deste modelo através da FAC e FACP se torna mais complicada pois sua FACP se comporta como a de um processo  $MA$  puro e sua FAC como a de um processo  $AR$  puro. Pode-se observar que se  $p = 0$  recaímos no modelo  $MA$  e se  $q = 0$  recaímos no modelo  $AR$ .

Vimos até o momento modelos que são aplicados a series estacionárias, ou seja, aquelas em que se supõe que a média e a variância são constantes em qualquer instante  $t$ . Quando essa suposição é quebrada (o que é comum, pois a maioria das séries reais são não-estacionárias), para aplicarmos o modelo  $ARMA$  é necessário, primeiramente, tornar a série estacionária. Para isso, é indicado que se diferencie a série; consulte [Morettin e Toloi, 2006]. Esse modelo que leva em consideração a diferenciação da série não-estacionária é chamado  $ARIMA(p, d, q)$ , onde  $d$  é a ordem de diferenciação.

Na Tabela 2.1 apresentamos um resumo de como identificar a ordem dos modelos ARIMA.

Tabela 2.1: Padrões teóricos da FAC e FACP para modelos ARMA

(Fonte: [Hippert, 2010])

	FAC	FACP
AR(p)	Barras decaem exponencialmente ou em forma de senóide amortecida	Há $p$ barras, de $k = 1$ a $k = p$
MA(q)	Há $q$ barras, de $k = 1$ a $k = q$	Barras decaem exponencialmente ou em forma de senóide amortecida
ARMA(p,q)	Sequencia infinita de barras, dominada por exponenciais ou senóides amortecidas, para $k > p - q$	Sequencia infinita de barras, dominada por exponenciais ou senóides amortecidas, para $k > p - q$

A adequação dos modelos ARIMA é confirmada através da verificação das suposições do modelo; ou seja, verifica-se, por teste de hipóteses, se os erros de previsão são descorrelacionados e normais, com média zero e variância constante.



### 2.2.4 Modelo Estrutural

Nos modelos de previsão vistos anteriormente as componentes da série, *média*, *tendência e sazonalidade* deveriam se manter constantes no tempo, pois havia a suposição de estacionariedade. Nos *Modelos Estruturais* esta suposição não é necessária, pois eles conseguem modelar séries que possuem componentes fixos ou aleatórios ao longo do tempo [Morettin e Toloi, 2006].

Os *Modelos Estruturais* têm a vantagem de serem facilmente interpretáveis. Nesses modelos a previsão é feita por uma atualização sequencial de estimativas iniciais, usando o *filtro de Kalman*<sup>5</sup>. Dessa forma, pode-se fazer previsões extrapolando a partir dos últimos valores observados [Hippert, 2011].

Todo modelo linear de séries temporais pode ser representado por equações de *espaço de estados* como nas equações 2.33 e 2.34. Na equação 2.33, chamada de *equação das observações*, o vetor de observações  $\mathbf{Z}_t$  é escrito como uma função linear de um vetor de variáveis aleatórias  $\mathbf{X}_t$ . Como este deve conter informações sobre toda a série, mas possui variáveis que nem sempre são observáveis, deve ser atualizado a cada instante por um processo markoviano, dado pela equação 2.34, chamada de *equação de estados*.

$$\mathbf{Z}_t = \mathbf{F}\mathbf{X}_t + \boldsymbol{\nu}_t \quad (2.33)$$

$$\mathbf{X}_t = \mathbf{G}\mathbf{X}_{t-1} + \mathbf{v}_t \quad (2.34)$$

Onde:  $\mathbf{X}_t$ : vetor de estado do sistema ( $dX1$ ),  $\mathbf{F}$ : vetor multiplicador ( $1Xd$ ),  $\mathbf{G}$ : matriz multiplicadora ( $dXd$ ),  $\mathbf{v}_t$ : vetor de erros ( $dX1$ ) e  $\boldsymbol{\nu}_t \sim N(0, \sigma_\nu)$

A atualização das estimativas do vetor  $\mathbf{X}_t$  é feita pelo *filtro de Kalman* a cada instante. Este filtro é um algoritmo de estimação recursiva, ou seja, a estimação a cada instante do vetor  $\mathbf{X}_t$  é dada por toda a informação da série obtida até aquele instante. Ao final do algoritmo é possível fazer previsões.

A seguir são apresentados os principais *modelos estruturais* em séries temporais.

<sup>5</sup>Para mais detalhes consulte [Morettin e Toloi, 2006]

- **Modelo de Nível Local:**

$$Z_t = \mu_t + \varepsilon_t \quad (2.35)$$

$$\mu_t = \mu_{t-1} + \eta_t \quad (2.36)$$

Em que  $\varepsilon_t \sim N(0, \sigma_\varepsilon)$ ,  $\eta_t \sim N(0, \sigma_\eta)$  independentes e não-correlacionadas entre si.

- **Modelo de Tendência Local:**

$$Z_t = \mu_t + \varepsilon_t \quad (2.37)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.38)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.39)$$

Em que  $\varepsilon_t \sim N(0, \sigma_\varepsilon)$ ,  $\eta_t \sim N(0, \sigma_\eta)$  e  $\xi_t \sim N(0, \sigma_\xi)$  com  $\xi_t$  e  $\eta_t$  mutuamente não-correlacionados, e não correlacionados com  $\varepsilon_t$ .

- **Modelo de Tendência Local e Componente Sazonal:**

$$Z_t = \mu_t + S_t + \varepsilon_t \quad (2.40)$$

Onde  $S_t$  é a componente sazonal e  $\mu_t$  é dada pelas equações 2.38-2.39

## 2.3 Combinação de Previsões

A idéia de combinar previsões de diferentes origens é relativamente antiga, tendo surgido antes mesmo dos modelos ARIMA. A primeira referência citada com frequência é [Bates e Granger, 1969]. Estes autores estudaram formas de combinação linear de duas previsões, e sugeriram a forma de combinação ótima (ver capítulo 3). [Newbold e Granger, 1974] estenderam depois a idéia para mais de duas previsões. Uma ampla revisão sobre o tema foi publicada por [Clemen, 1989]. Desde então, a literatura sobre o tema tem crescido bastante. Uma importante revisão posterior, que citaremos diversas vezes ao longo deste trabalho, foi a de [Menezes et al., 2000].

O uso de combinações de previsões, contudo, não é em geral mencionado nos livros-texto sobre previsão. Por exemplo, o mais completo livro sobre séries temporais e previsão em língua portuguesa [Morettin e Tolo, 2006], não menciona o assunto. Mesmo [Makridakis et al., 1998], que é provavelmente o mais amplo tratamento dos métodos de previsão (desde os estatísticos até os qualitativos), não dedica mais que duas páginas (embora discuta os resultados obtidos em um estudo com 1001 séries, que mostrou que diferentes formas de combinações desempenharam melhor do que os métodos individuais). [Montgomery et al., 1990] estuda empiricamente a utilização de diversas formas de combinação na previsão de 10 séries reais (p.191), e compara vários métodos para a escolha dos pesos, inclusive a combinação ótima.

Atualmente, quando se trata de estudos sobre métodos de previsão, a combinação de previsões tem sido muito estudada por diversos pesquisadores [Rocha, 2008]. A combinação de previsões tem sido geralmente estudada como uma forma de compensar as deficiências de um método, pela combinação de suas previsões com as de outro método. É desejável que se tenha uma diversidade entre as previsões individuais, o que garante que um método compense a fraqueza de outro. A forma padrão de se olhar para a diversidade de um conjunto de métodos de previsão é examinar os coeficientes de correlação entre as previsões [Lemke e Gabrys, 2010]; a diversidade por si só, contudo, não é a chave para uma boa combinação; a precisão das previsões individuais também é muito importante para se obter um melhor resultado. Para [Zou e Yang, 2004] a vantagem é que, quando há muita incerteza na escolha do melhor modelo, como é o caso em muitas aplicações, a combinação pode reduzir a instabilidade da previsão e, portanto, melhorar a acurácia da previsão. De acordo com [Menezes et al., 2000], pesquisas teóricas e de simulação tem mostrado que o desempenho relativo da combinação de previsão depende de variância do erro das previsões, da correlação entre os erros de previsão e da amostra de dados históricos utilizados para a estimativa.

## 2.4 Medidas de Erros de Previsão

Usar testes de hipóteses estatísticos para selecionar o melhor modelo apresenta várias dificuldades, como a questão de testes múltiplos e o fato de que o modelo preferencial de um teste não necessariamente tem o melhor desempenho na prática. Alternativamente, a seleção de modelos pode ser feita por critérios como o Critério de Informação de Akaike (AIC) e o Critério de Informação Bayesiano (BIC) [Morettin e Tolo, 2006]. A desvantagem seria a instabilidade destes critérios, pois séries temporais com poucas observações, onde os diferentes modelos de previsão são semelhantes, os valores dos critérios são bem próximos e fica difícil escolher o melhor modelo [Zou e Yang, 2004]

Denotamos anteriormente (veja página 19)  $e_{t+k}$  como o erro de previsão  $k$  passos à frente. Esses erros são úteis para a otimização dos parâmetros do modelo e para a verificação de sua adequação, através de algumas medidas [Makridakis et al., 1998], como as apresentadas na Tabela 2.2.

Tabela 2.2: Medidas de Erros de Previsão

<b>Erro Médio:</b> (Mean Error,ME)=	$\frac{1}{N} \sum_{i=1}^N e_t$
<b>Erro Médio Absoluto:</b> (Mean Absolute Error,MAE)=	$\frac{1}{N} \sum_{i=1}^N  e_t $
<b>Erro Percentual Médio:</b> (Mean Percent Error,MPE) =	$\frac{100}{N} \sum_{i=1}^N \left( \frac{e_t}{Z_t} \right)$
<b>Erro Absoluto Percentual Médio:</b> (Mean Absolute Percent Error,MAPE) =	$\frac{100}{N} \sum_{i=1}^N \left  \left( \frac{e_t}{Z_t} \right) \right $
<b>Erro Quadrático Médio:</b> (Mean square Error,MSE) =	$\frac{1}{N} \sum_{i=1}^N e_t^2$

Essas medidas não devem ser calculadas na série de erros completa. Como alguns modelos necessitam de que sejam atribuídos valores iniciais, as primeiras previsões provavelmente não serão muito adequadas, pois sofrem influência destes valores iniciais. A parte inicial das séries de erros será portanto desprezada, quando do cálculo das medidas [Morettin e Tolo, 2006].

Para a verificação da adequação de um modelo é comum comparar vários modelos ajustados através de seus erros. Entretanto, o modelo pode se ajustar bem aos dados usados na sua estimação, e ainda assim não produzir boas previsões. Dessa

---

forma, é indicada a separação dos últimos valores da série temporal para a comparação de modelos; esses últimos valores não são usados para a estimação. A parte central da série de erros, portanto, é usada para a otimização dos parâmetros do modelo: ajusta-se o mesmo modelo com diferentes valores para os parâmetros e escolhem-se aqueles valores que levaram à menor medida de erro.

### 3 Metodologia

A proposta deste trabalho foi verificar se a combinação de previsões é capaz de melhorar a precisão das previsões feitas por diferentes métodos. Fizemos previsões individuais por três modelos, em seguida vários tipos de combinações com essas previsões, e por fim, uma comparação do Erro Quadrático Médio (MSE) das previsões de cada modelo e cada combinação. Não serão feitas comparações pelo Erro Absoluto Percentual Médio (MAPE) devido ao fato de se ter séries contendo observações próximas de zero.

As previsões individuais foram feitas pelo método de amortecimento exponencial de Holt, por modelos ARIMA e por modelos estruturais de tendência local. Optamos por esses modelos devido à sua consolidação na literatura e à sua aplicação prática.

O *método de amortecimento* de Holt, apesar de ser um método empírico, que não se preocupa com a distribuição de probabilidade geradora da série, tem as vantagens de produzir rapidamente estimativas razoavelmente precisas, e de ser um método de simples entendimento.

Já os *modelos ARIMA*, que se baseiam na análise das propriedades probabilísticas da série, dependem de um bom conhecimento teórico de Estatística. Eles permitem que se façam análises estatísticas nas previsões, como testes de hipóteses e intervalos de confiança, e se destacam pela precisão de suas previsões.

Os *modelos estruturais* são uma alternativa para o estudo das séries temporais não-estacionárias. Eles são capazes de modelar essas séries sem a necessidade de diferenciá-las como nos modelos ARIMA.

Os modelos de combinação de previsões utilizados foram:

- ***Média Simples***: Segundo [Menezes et al., 2000] essa combinação é muito utilizada devido à sua robustez. São necessárias, no mínimo, duas previsões individu-

ais.

$$Z_t = P_1 + P_2 + P_3 \quad (3.1)$$

Onde  $P_i$  é uma previsão individual.

- **Mediana:** Esta será estudada devido ao fato de sua robustez, uma vez que a mediana não é tão influenciada por valores extremos quanto a média. No entanto, para usar essa combinação são necessárias no mínimo três previsões individuais.
- **Média dos Extremos:** Nesta combinação, calcula-se a média entre a maior e a menor previsão individual.

$$\frac{\min[\mathbf{P}_i] + \max[\mathbf{P}_i]}{2} \quad (3.2)$$

Onde  $\mathbf{P}$  é a matriz de previsões individuais por vários métodos.

- **Regressão:** Esta combinação usa as previsões individuais como variáveis regressoras para explicar a variabilidade da série temporal  $Z_t$ .

$$Z_t = \zeta + \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3 + \xi_t \quad (3.3)$$

Onde  $\zeta$  é uma constante,  $P_i$  é uma previsão individual e  $\xi_t$  é um erro aleatório.

- **Ótima:** Estudada por Bates [Menezes et al., 2000], consiste em uma média ponderada pela covariância dos erros de previsão, em que os pesos são dados por:

$$\omega = \frac{\mathbf{S}^{-1}\mathbf{1}}{\mathbf{1}^T \mathbf{S}^{-1}\mathbf{1}} \quad (3.4)$$

Onde  $\mathbf{S}$  é a matriz de covariância dos erros de previsão e  $\mathbf{1}$  é uma vetor unitário.

- **Ótima com independência:** Consiste em uma média ponderada, como a do item anterior, mas considerando a independência entre os erros de previsão. Dessa forma, a matriz  $\mathbf{S}$  na fórmula (3.4) é uma matriz diagonal, contendo apenas as variâncias dos erros.

A ferramenta computacional utilizada para a análise exploratória dos dados, obtenção das previsões individuais e combinada, e cálculo do MSE, foi o *software* livre R (versão 2.13). Foi necessário o desenvolvimento de algumas funções no R para facilitar a otimização dos parâmetros dos modelos estimados. Tais funções encontram-se no anexo A.1.

Para analisar a precisão das combinações de previsões, usamos dez séries de dados reais. Escolhemos séries com cerca de 100 observações, que não possuem sazonalidade nem quebra estrutural (mudança brusca do nível da série). As séries de 1 a 9 estão disponíveis em [Hyndman, 2011]. A série 10 foi fornecida pelo professor Henrique Hippert.

1. **Série Cow:** Temperatura matutina diária de uma vaca.
2. **Série Buffsnow:** Precipitação anual de neve, na cidade de Buffalo (NY), nos Estados Unidos da América (1910-1972).
3. **Série Calfem:** Total de nascimentos diários do sexo feminino, no estado da Califórnia nos Estados Unidos da América, dados do primeiro trimestre de 1959.
4. **Série DowJones:** Índice DowJones da bolsa de valores de Nova York, dados de 28 agosto à 18 dezembro 1972.
5. **Série Globtp:** Mudanças na temperatura global, em graus Celsius, dados anuais (1880-1985).
6. **Série Huron:** Nível do Lago Huron<sup>1</sup> em pés (reduzido de 570 pés), dados de 1875-1972.
7. **Série Sheep:** População de carneiros (em milhões) da Inglaterra e País de Gales, dados de 1867-1939.
8. **Série Summer:** Temperatura média no verão, em graus Celsius, na cidade de Munique-Riem, dados anuais (1781-1988).

---

<sup>1</sup>O Lago Huron é um dos cinco Grandes Lagos, localizado entre os estado de Michigam (EUA) e a província canadense de Ontário



9. ***Série Ausgundeaths*** : Taxa de suicídio sem arma de fogo na Austrália, dados anuais (1915-2004).
10. ***Série Carga***: Carga elétrica semanal de uma concessionária da cidade do Rio de Janeiro, em 1996-1997.

Para a análise dos erros de previsão dividimos as séries de erros em duas partes. A partição contendo 80% da série de erros, que chamaremos de *série de estimação*, foi usada para a estimação dos parâmetros dos modelos, sendo que os primeiros 20% não foram usados para o cálculo das medidas de erros de previsão, de forma a evitar a influência dos valores iniciais atribuídos em cada modelo. A segunda partição, contendo os últimos 20% da série de erros, foi utilizada para a comparação entre os modelos.

## 4 Resultados

No total, ajustamos 9 modelos para cada série; portanto, temos 90 modelos. Para uma melhor visualização dos resultados, dividimos este capítulo em quatro seções: Análise Descritiva (seção 4.1), Modelos de Previsões Individuais (seção 4.2), Combinações de Previsões (seção 4.3) e Resultados na Amostra de Teste (seção 4.4).

### 4.1 Análise Descritiva

Primeiramente é necessária a análise descritiva de todas as séries que serão analisadas; isso facilitará a escolha dos modelos a serem ajustados em cada uma.

- **Cow:** Pelo gráfico de linha e pelas FAC e FACP da Figura 4.1 pode-se pensar que

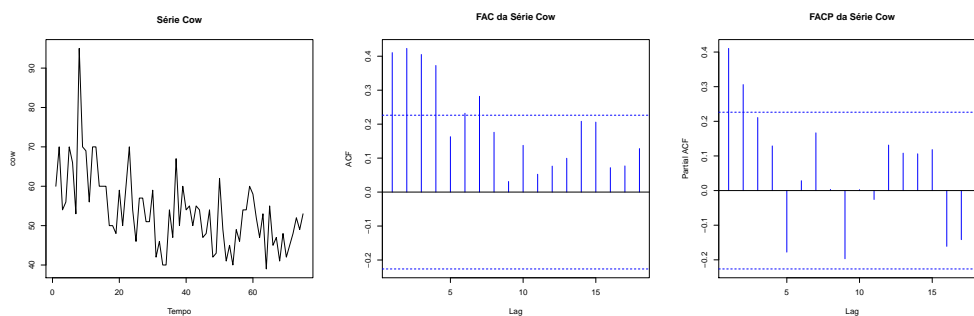


Figura 4.1: Gráfico de linha, FAC e FACP da série Cow

a série é não-estacionária. Mas como se trata da temperatura de uma vaca não faz sentido ser não-estacionária, do contrário a vaca morreria.

- **Busffsnow:** Considerando a média em vários trechos diferentes da série e suas FAC e FACP (Figura 4.2), percebemos que a série parece ser estacionária.
- **Calfem:** Pelo gráficos da Figura 4.3 percebemos que a série é estacionária, porém

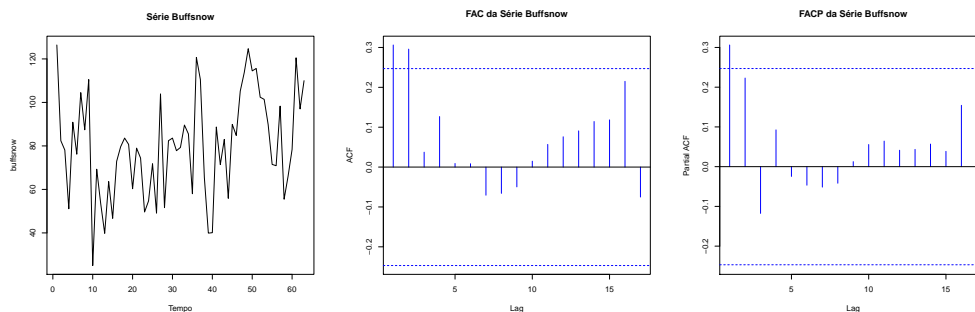


Figura 4.2: Gráfico de linha, FAC e FACP da série Buffsnow

parece apresentar uma alta variância. Percebemos, ainda, que a série pode ser modelada por uma constante mais um ruído, uma vez que nenhum coeficiente de auto-correlação e auto-correlação parcial foram significativos.

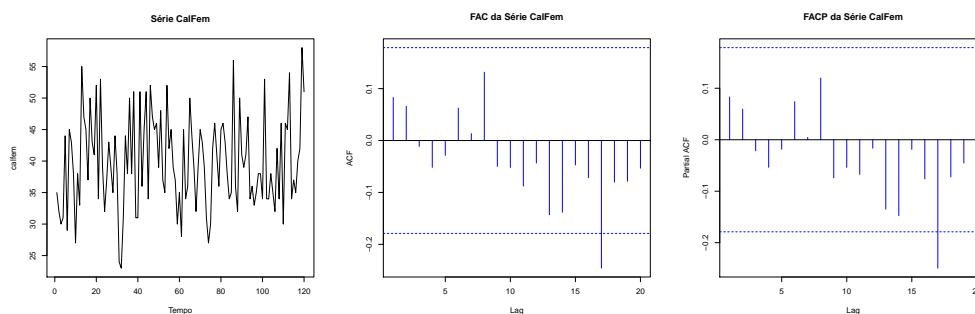


Figura 4.3: Gráfico de linha, FAC e FACP da série CalFem

- **DowJones:** Pelos gráficos da Figura 4.4 percebemos que a série é não-estacionária, portanto é necessária uma diferenciação. Feita uma diferenciação, a série se tornou estacionária. Observe a Figura 4.5.
- **Huron:** Pelas médias em diferentes trechos da série e pelos gráficos da Figura 4.6 percebemos que a série parece ser estacionária. Pelas FAC e FACP, um modelo AR(2) parece ser um modelo de ajuste a essa série; confira subseção 4.2.2.
- **Globtp:** Analisando os gráficos da Figura 4.7, percebemos que a série parece ser não-estacionária. Diferenciando uma vez, a série se tornou estacionária, veja Figura 4.8.

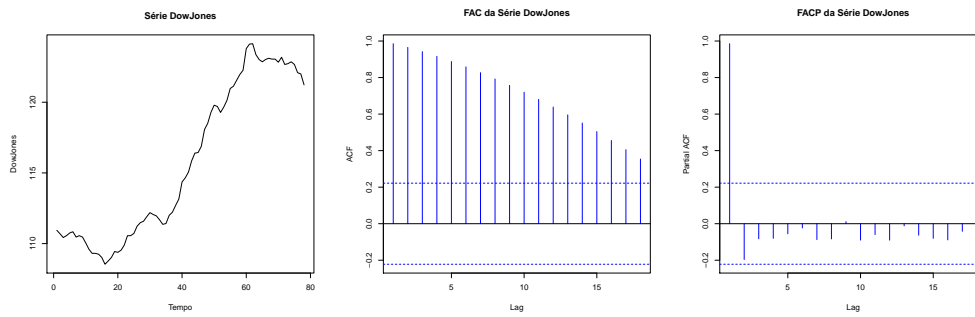


Figura 4.4: Gráfico de linha, FAC e FACP da série DowJones

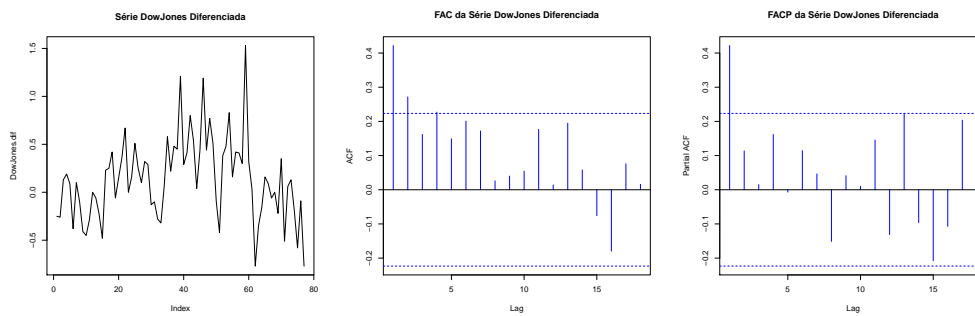


Figura 4.5: Gráfico de linha, FAC e FACP da série DowJones Diferenciada

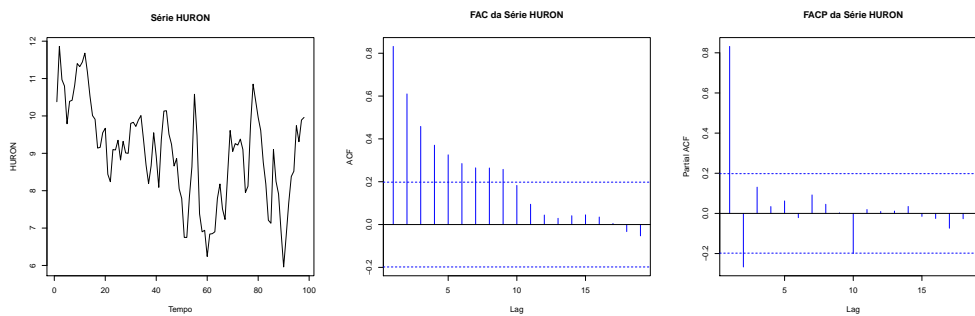


Figura 4.6: Gráfico de linha, FAC e FACP da série Huron

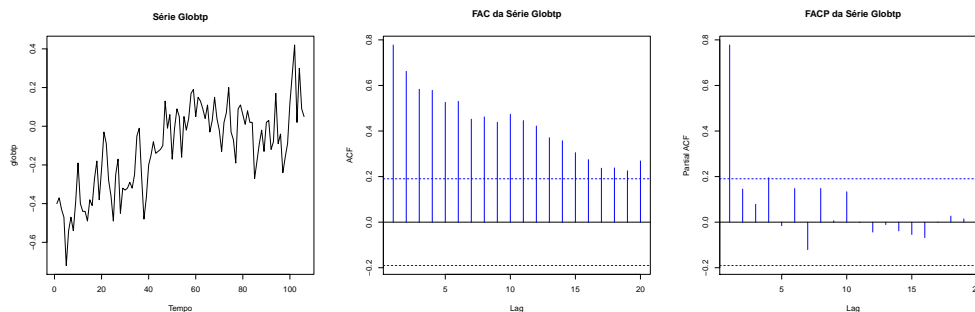


Figura 4.7: Gráfico de linha, FAC e FACP da série Globtp

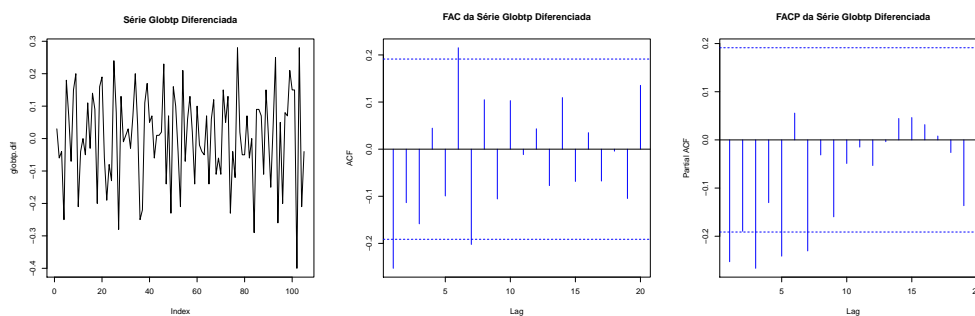


Figura 4.8: Gráfico de linha, FAC e FACP da série Globtp Diferenciada

- **Sheep:** Pelas médias de diferentes trechos da série e pelas FAC e FACP da Figura 4.9, percebemos que a série parece ser não-estacionária. Dada a série diferenciada percebemos, pelos gráficos da Figura 4.10, que a série parece ser estacionária e pode-se dizer que segue um modelo ARMA.
- **Summer:** Pelas médias de diferentes trechos da série, pelo gráfico de linha e pelas FAC e FACP da Figura 4.11 percebemos que a série é estacionária e que um bom modelo seria um ARMA.
- **Ausgundeaths:** Pelos gráficos da Figura 4.12 percebemos que a série parece ser não-estacionária, pois o decaimento da correlação dos erros se dá de forma lenta. Dada a série diferenciada, percebemos, pelos gráficos da Figura 4.13, que a série passa a ser estacionária. Não fica muito claro pelas FAC e FACP qual o modelo ARIMA seria mais adequado.

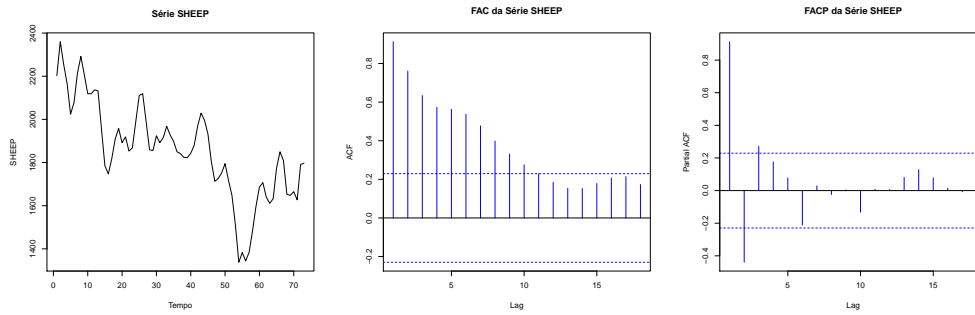


Figura 4.9: Gráfico de linha, FAC e FACP da série Sheep

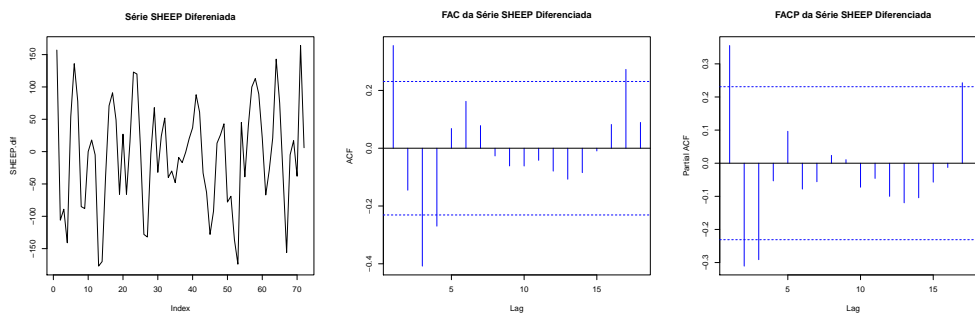


Figura 4.10: Gráfico de linha, FAC e FACP da série Sheep Diferenciada

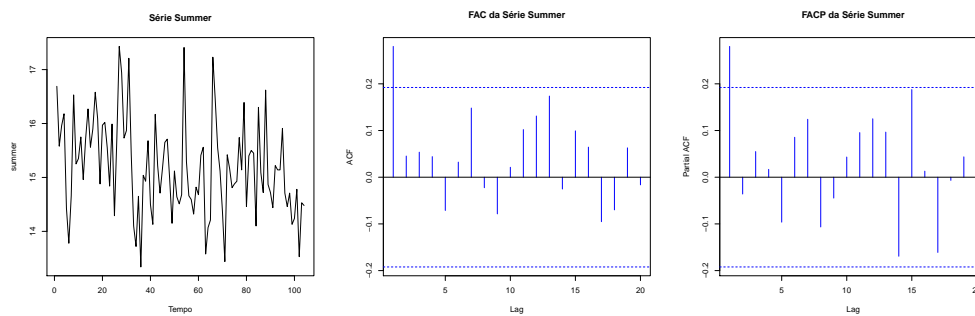


Figura 4.11: Gráfico de linha, FAC e FACP da série Summer

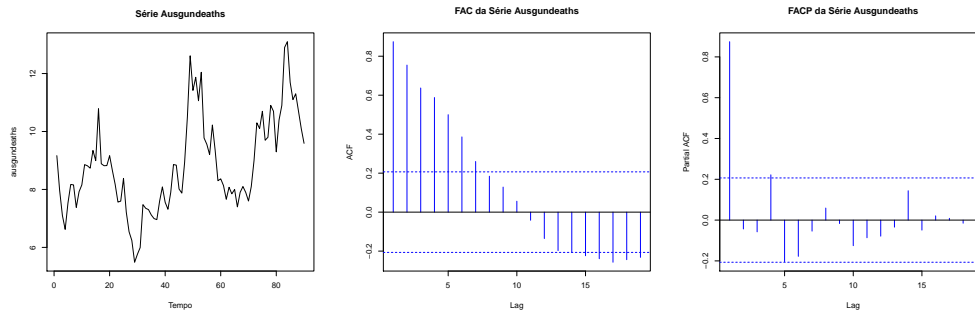


Figura 4.12: Gráfico de linha, FAC e FACP da série Ausgundeaths

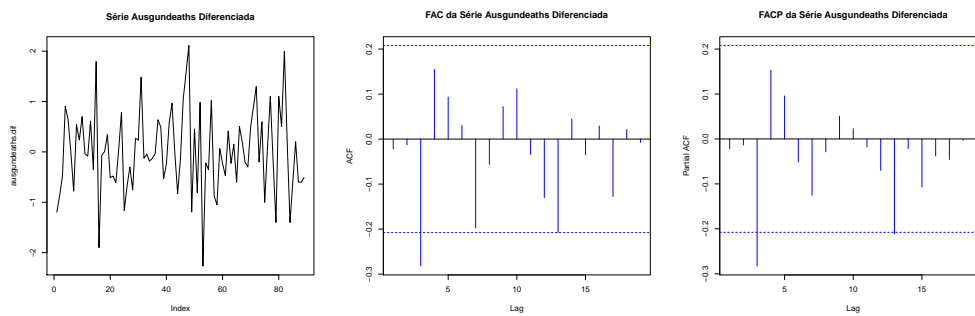


Figura 4.13: Gráfico de linha, FAC e FACP da série Ausgundeaths Diferenciada

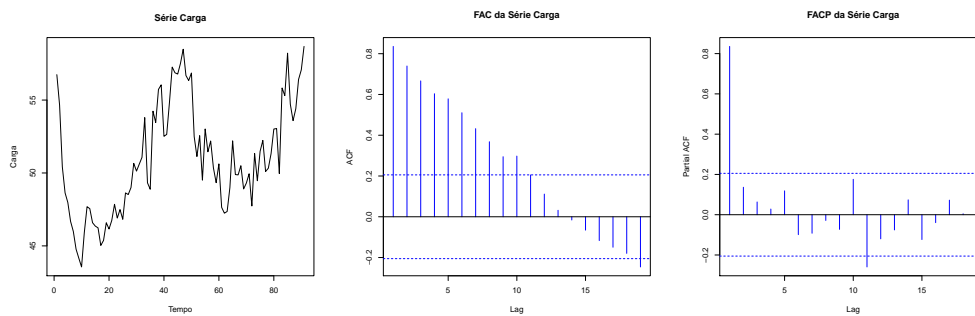


Figura 4.14: Gráfico de linha, FAC e FACP da série Carga

- **Carga:** Pelos gráficos da Figura 4.14 percebemos que a série parece ser não-estacionária. Os gráficos da Figura 4.15 são da série diferenciada. Um possível modelo seria  $ARIMA(0,1,0)$

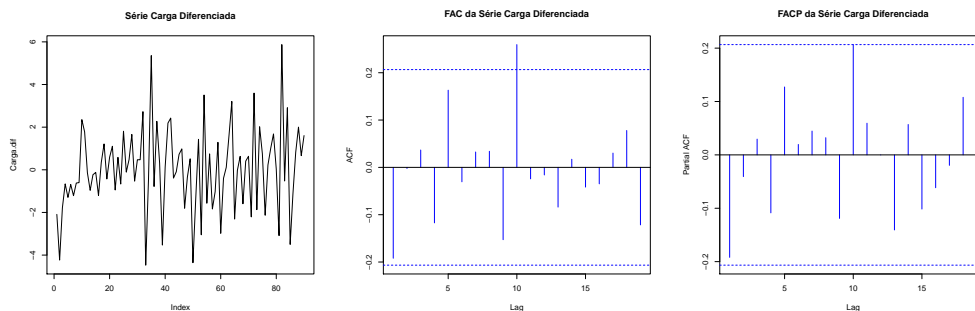


Figura 4.15: Gráfico de linha, FAC e FACP da série Carga Diferenciada

## 4.2 Modelos de Previsões Individuais

Analisamos vários modelos de previsão para cada série, e otimizamos os parâmetros observando o MSE. A seguir são apresentados os modelos de previsão com o menor MSE no trecho da série de estimação.

### 4.2.1 Amortecimento Exponencial de Holt

Pelas fórmulas 2.23 e 2.24 podemos fazer previsões pelo método AEH, e com os parâmetros apresentados na Tabela 4.1 obtemos os menores erros de previsão. Percebemos, ainda, que algumas séries possuem a constante  $\alpha$  próxima de 1; conseqüentemente suas previsões são semelhantes às previsões pelo método de Naïve (onde as observações mais antigas não são importantes) e não são muito precisas. Dentre as 10 séries analisadas a série Sheep (veja Figura 4.16), que teve as duas constantes próximas de 1, foi a que apresentou as piores previsões e a série Globtp, que obteve valores baixos para as constantes, apresentou boas previsões. Podemos verificar pela Figura 4.17 que as previsões acompanham a linha de tendência das observações da série Globtp.



Tabela 4.1: Parâmetros otimizados do modelo AEH

Modelo\Série	Cow	Buff.	Calf.	DowJ.	Globtp	Huron	Sheep	Sum.	Ausg.	Carga
MSE	52,674	510,568	56,813	0,1944	0,015	0,606	6309	0,9502	0,621	3,632
$\alpha$	0,05	0,35	0,15	0,95	0,25	0,95	0,95	0,2	0,95	0,55
$\beta$	0,15	0,2	0,05	0,2	0,05	0,05	0,95	0,05	0,1	0,15

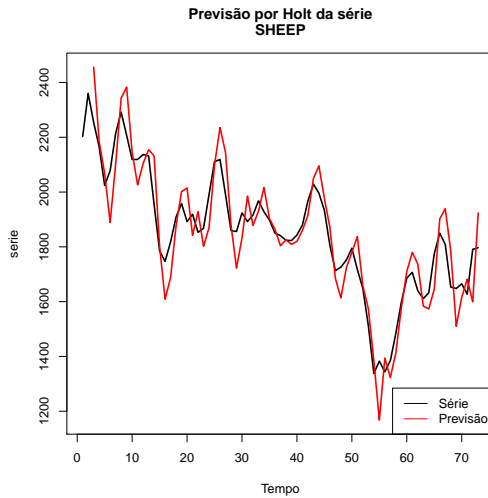


Figura 4.16: Previsão por AEH da série Sheep

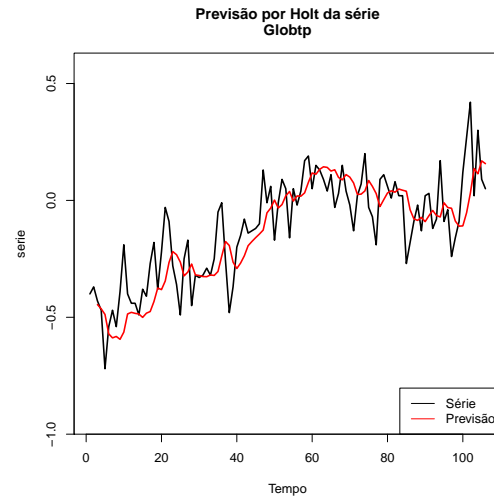


Figura 4.17: Previsão por AEH da série Globtp

### 4.2.2 ARIMA

Ajustamos diversos modelos ARIMA para cada série e escolhemos o melhor modelo através do MSE, mas sempre verificando os pressupostos desses modelos: média nula, normalidade e decorrelação dos erros. Para tal fizemos os testes  $t$ , normalidade de Lilliefors e Durbin Watson<sup>1</sup>, representados na Tabela 4.2 por  $t$ ,  $L$  e  $DW$ , respectivamente, respectivamente, ao nível de significância de 1%. Como forma de verificação da decorrelação dos erros de previsão também foi feito a FAC dos erros; como exemplo apresentamos a FAC dos erros de previsão da série Cow (veja Figura 4.18).

Pela Tabela 4.2 temos o melhor modelo ARIMA para cada série, seus respectivos MSE e estatísticas. É apresentado também o Critério de Informação de Akaike(AIC) sobre toda a série temporal; não adotamos esta medida para a escolha do melhor mo-

<sup>1</sup>Os erros são decorrelacionados quando DW está próximo de 2.

delo pois queremos otimizar os parâmetros na sub-série de estimação. Por essa tabela, verificamos que os pressupostos dos modelos ARIMA não foram quebrados em nenhuma série.

Tabela 4.2: Modelos ARIMA

Série	Modelo	MSE	AIC	DW	t	L
Cow	ARIMA(1,0,2)	52,602	435,8951	1,7951	0,1749	0,7962
Buffsnow	ARIMA(1,0,1)	430,3222	469,2364	1,8501	0,9929	0,2553
Calfem	ARIMA(0,0,0)	50,5278	656,4667	1,7741	0,9902	0,1314
DowJones	ARIMA(1,1,2)	0,1776	64,0239	1,8354	0,1498	0,71
Globtp	ARIMA(1,1,1)	0,0137	-113,4217	1,8573	0,4773	0,4384
Huron	ARIMA(2,0,0)	0,4703	168,9344	1,8416	0,2821	0,9007
Sheep	ARIMA(2,1,1)	3741,705	661,0407	2,0568	0,2402	0,1995
Summer	ARIMA(1,0,0)	0,781	217,0226	1,9406	0,4476	0,1971
Ausgundeaths	ARIMA(1,1,1)	0,5440	167,315	1,8408	0,9445	0,0039
Carga	ARIMA(0,1,1)	3,5005	293,6501	2,2548	0,6474	0,4044

A seguir apresentamos como exemplo os gráficos das previsões de duas séries. Podemos observar pelo gráfico da Figura 4.19 que as previsões de Ausgundeaths não são muito boas, pois a linha de previsão parece seguir a linha das observações, como no método Naïve. Já as previsões da série Summer parecem melhores (Figura 4.20). Os parâmetros do modelo de cada série são apresentados no anexo A.2.

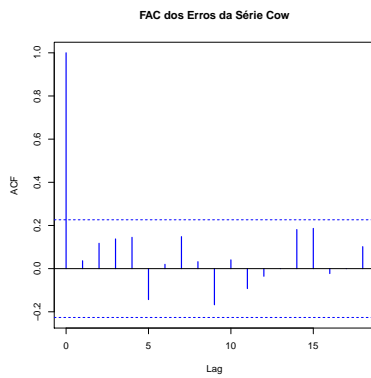


Figura 4.18: FAC dos erros de previsão por ARIMA da Série Cow

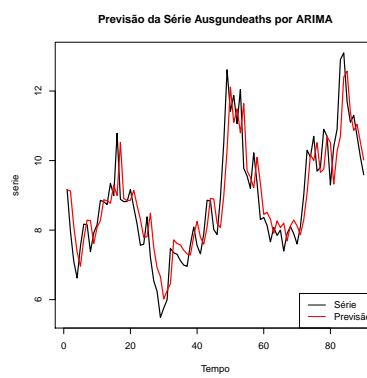


Figura 4.19: Previsão por ARIMA da série Ausgundeaths

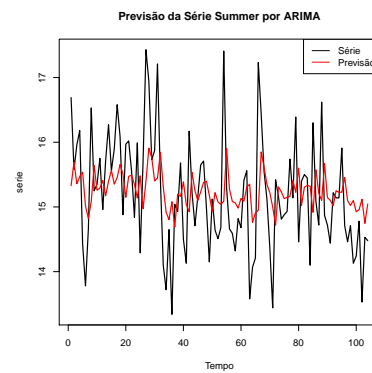


Figura 4.20: Previsão por ARIMA da série Summer

### 4.2.3 Modelo Estrutural

No modelo estrutural de tendência local não temos que otimizar os parâmetros, pois esses são atualizados em todos os instantes  $t$  através do filtro de Kalman (seção 2.2.4). Mas como temos que entrar com valores iniciais para as variâncias dos erros do modelo (veja fórmulas 2.38 e 2.39), a otimização do modelo foi feita em cima desses valores iniciais; ou seja, calculamos o MSE na série de estimação para diversos valores iniciais da variância e adotamos aqueles em que o MSE foi o menor. Usamos a função `Var_Mod_Est`, disponível no anexo A.1 .

A Tabela 4.3 mostra os valores iniciais das variâncias usadas em cada série, e os respectivos MSE na série de estimação. Podemos observar que a maior variabilidade se encontra no erro  $\varepsilon$  na maioria das séries; exceto na série DowJones e Sheep, onde a maior variabilidade está no erro  $\xi$  referente à tendência. A explicação está nas próprias séries, que apresentam uma tendência mais forte.

Tabela 4.3: Variâncias dos Modelos Estruturais

Série	Cow	Buff.	Calf.	DowJ.	Globtp	Huron	Sheep	Sum.	Ausg.	Carga
MSE	65,7497	516,7902	58,3130	0,2330	0,0167	0,9743	6092,0890	0,9028	0,7648	3,7224
$\sigma_\varepsilon$	100,0100	120,6770	100,0234	10,0808	200,00	100,0241	1,2373e-6	100,0078	95,7015	99,7167
$\sigma_\eta$	1,0	1,0	1,0	10,0	10,00	10,0	1,0	1,0	1,0	15
$\sigma_\xi$	1,87	8,2350	0,2020	99,9920	9,9840	9,7560	100,7050	02148	42,0210	12,5810

Como ilustração desse método de previsão apresentamos os gráficos da Figura 4.21 e da Figura 4.22. Podemos observar por estes que a série Huron não obteve previsões muito precisas, e a série DowJones sim.

## 4.3 Combinação de Previsões

Dadas diferentes previsões por métodos distintos, podemos combiná-las e obter uma nova previsão, a qual acreditamos ser mais precisa. Na prática, é comum ter acesso às previsões individuais sem o conhecimento do método usado em sua estimação, e

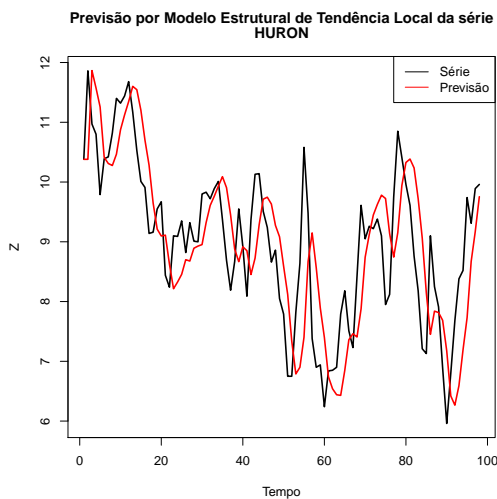


Figura 4.21: Previsão por Modelos Estruturais da série Huron.

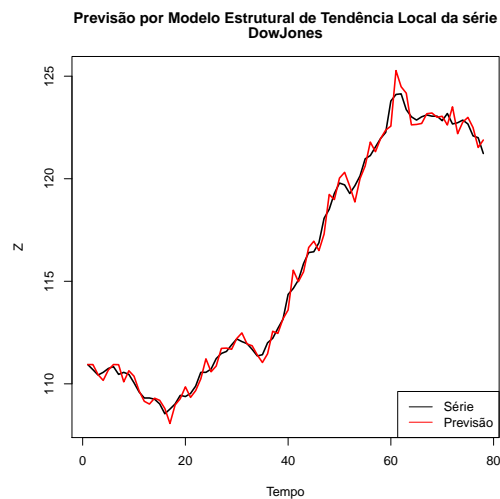


Figura 4.22: Previsão por Modelos Estruturais da série DowJones.

mesmo assim fazer uma combinação. Como, neste caso, não podemos alterar o modelo de previsão individual com o intuito de melhorar a precisão da previsão, é interessante que a cada instante façamos a reestimação do modelo de previsão combinada para aqueles métodos de combinação que exigem a estimação de alguns parâmetros, como no caso da combinação por regressão, ótima e ótima com independência.

As tabelas a seguir mostram os parâmetros estimados dos modelos de combinação por regressão (Tabela 4.4), combinação ótima (Tabela 4.5) e combinação ótima com independência (Tabela 4.6), usando a série de estimação. Após a estimação dos parâmetros de cada modelo foram feitas previsões usando um modelo com esses parâmetros fixos, segundo o modelo 4.1.

$$Z_t = \zeta + \lambda_{AEH} + \lambda_{ARIMA} + \lambda_{Mod.Estrutural} \quad (4.1)$$

A série Calfem não apresenta o parâmetro  $\lambda_{ARIMA}$  pois suas previsões pelo método ARIMA foram uma constante e pelos conceitos da regressão essas previsões não trazem nenhuma informação nova.

Fizemos previsões com os métodos de combinação de regressão, ótima e ótima

Tabela 4.4: Parâmetros da Combinação por Regressão.

Série	Cow	Buff.	Calf.	DowJ.	Globtp	Huron	Sheep	Sum.	Ausg.	Carga
$\zeta$	-18,4005	-11,1995	36,4211	0,5869	-0,0129	-0,6031	63,7465	2,142	1,2648	10,4397
$\lambda_{AEH}$	-0,5241	0,2772	0,0075	0,1660	0,1429	-0,3748	0,6983	-0,9152	0,3898	0,8667
$\lambda_{ARIMA}$	2,0218	1,2884	NA	0,7482	0,8272	1,275	0,9246	1,3642	0,3611	0,2075
$\lambda_{Mod.Estrutural}$	-0,2135	-0,4402	0,0758	0,0809	-0,0518	0,162	-0,6579	0,3993	0,0989	-0,2764

com independência reestimando seus parâmetros a cada instante. Pelos gráficos abaixo (Figura 4.23 e Figura 4.24) podemos verificar como os parâmetros da série Buffsnow variam no decorrer no tempo. As tabelas da seção 4.4 permite comparar as precisões desses métodos de combinação feitos sem a reestimação dos seus parâmetros e com a reestimação, através do MSE.

Tabela 4.5: Pesos da Combinação Ótima.

Modelo\Série	Cow	Buff.	Calf.	DowJ.	Globtp	Huron	Sheep	Sum.	Ausg.	Carga
Peso_Holt	-0,689	0,2402	0,0075	0,0252	-0,105	-0,141	0,7585	-0,889	-0,024	0,7009
Peso_ARIMA	1,7544	1,0995	0,9167	0,9556	1,0234	1,0500	1,0300	1,4793	0,9419	0,6539
Peso_Mod.Estrutural	-0,066	-0,339	0,0758	0,0192	0,0813	0,0905	-0,789	0,4104	0,0805	-0,355

Tabela 4.6: Pesos da Combinação Ótima com Independência.

Modelo\Série	Cow	Buff.	Calf.	DowJ.	Globtp	Huron	Sheep	Sum.	Ausg.	Carga
Peso_Holt	0,3245	0,2679	0,2949	0,3394	0,3091	0,3472	0,2592	0,3048	0,3243	0,3372
Peso_ARIMA	0,3769	0,4063	0,3783	0,3806	0,3843	0,4420	0,4861	0,3847	0,3954	0,3672
Peso_Mod.Estrutural	0,2986	0,3258	0,3268	0,2800	0,3065	0,2108	0,2548	0,3104	0,2803	0,2956

## 4.4 Resultados na Amostra de Teste

O MSE calculado sobre a série de teste nos permite comparar distintos modelos de previsão, e assim verificar qual é mais adequado. Desejamos verificar o desempenho das combinações de previsões, o qual acreditamos que seja melhor do que o das previsões individuais pois, de acordo com [Lemke e Gabrys, 2010], combinando as informações de

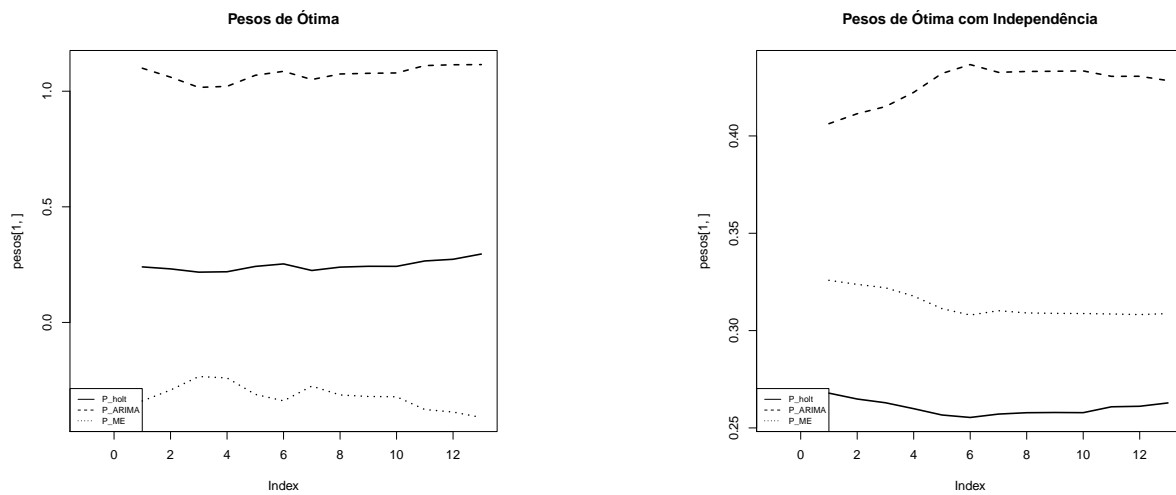


Figura 4.23: Parâmetros da Combinação Ótima e Ótima com Independência da Série Buffsnow

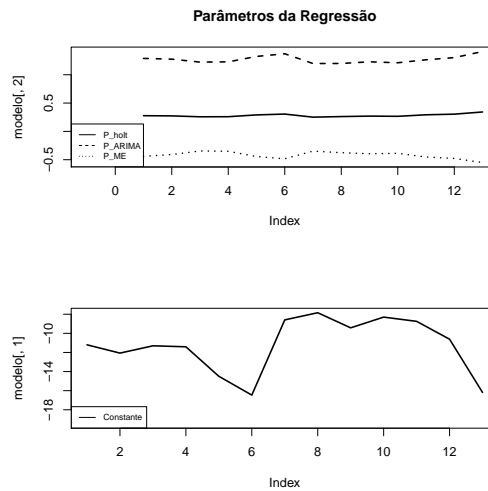


Figura 4.24: Parâmetros de Combinação por Regressão da Série Buffsnow

diversos modelos de previsão, podemos compensar as fraquezas de cada um e obter previsões mais precisas. A tabela 4.7, onde o menor MSE de cada série encontra-se destacado em negrito e os modelos de REG\*, ÓTIMA\* e ÓTIMAI\*, tiveram seus parâmetros reestimados; nos permite observar que na maioria das séries pelos menos uma combinação de previsões foi melhor do que as previsões individuais.

Comparando os três últimos modelos de combinação de previsões entre si

percebemos que quando reestimamos os parâmetros das combinações por Regressão, combinação Ótima e combinação Ótima com Independência ocorre uma pequena melhora na precisão das previsões dessas combinações.

Tabela 4.7: MSE na série de teste.

Modelo\Série	Cow	Buff.	Calf.	DowJ.	Globtp	Huron	Sheep	Sum.	Ausg.	Carga
<b>Holt</b>	<b>23,73</b>	697	59,18	0,1174	0,0258	0,7287	10751	<b>0,4602</b>	0,9164	4,20,25
<b>ARIMA</b>	40,92	346	<b>56,7</b>	0,1283	0,0218	0,5926	6700	0,5937	0,7741	4,6951
<b>Mod.Estrutural</b>	29,52	709	59,35	0,1745	0,0274	1,4140	11082	0,4951	1,2271	<b>4,1004</b>
<b>Med.Simples</b>	28,25	528	57,81	0,1162	0,0238	0,7445	8350	0,4982	0,9001	4,1703
<b>Mediana</b>	28,41	664	58,05	<b>0,1142</b>	0,0258	0,6520	9883	0,5103	0,9519	4,2492
<b>Med.Extremo</b>	28,37	476	57,72	0,1199	0,0230	0,8106	7866	0,4933	<b>0,6614</b>	4,1446
<b>REG</b>	35,58	341,66	68,63	0,1147	<b>0,0218</b>	<b>0,59</b>	<b>5798</b>	0,6732	0,8691	4,4707
<b>REG*</b>	30,41	<b>318</b>	70,94	<b>0,0871</b>	<b>0,0204</b>	<b>0,5502</b>	<b>5029</b>	0,6231	0,7710	4,1352
<b>ÓTIMA</b>	71,72	<b>335</b>	56,82	0,1272	0,0219	0,5940	6422	0,7405	0,7826	4,5960
<b>ÓTIMA*</b>	58,00	330	<b>56,58</b>	0,1208	0,0213	0,5757	6269	0,6918	0,7568	4,4777
<b>ÓTIMA.I</b>	28,66	498	57,7	0,1158	0,0235	0,6741	7573	0,8754	0,8741	4,1887
<b>ÓTIMA.I*</b>	52,43	490	57,68	0,1155	0,0235	0,6698	7597	0,5033	0,8741	4,1795

Os gráficos a seguir facilitam a visualização das diferenças entre as precisões das previsões de cada modelo, de acordo com a Tabela 4.7. Podemos perceber por estes que os erros de previsão de todos os métodos são bem semelhantes, mas que na maioria das séries os erros de previsão por combinação são menores que os erros de previsão individual.

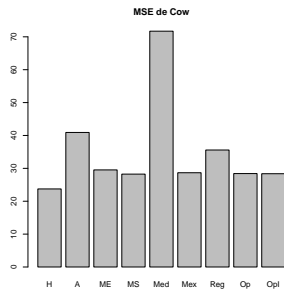


Figura 4.25: MSE:  
Cow.

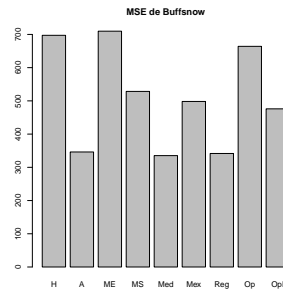


Figura 4.26: MSE:  
Buffsnow.

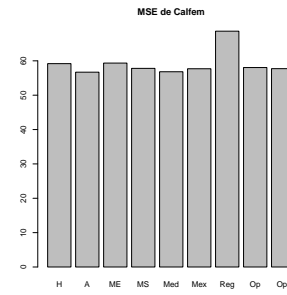


Figura 4.27: MSE:  
Calfem.

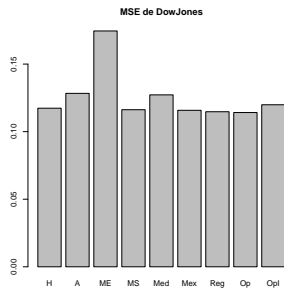


Figura 4.28: MSE:  
DowJones.

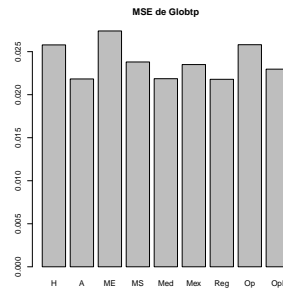


Figura 4.29: MSE:  
Globtp.

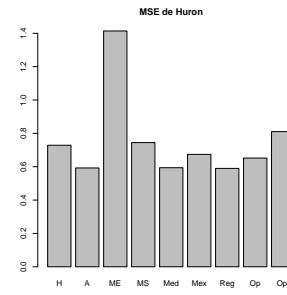


Figura 4.30: MSE:  
Huron.

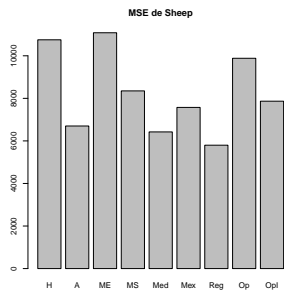


Figura 4.31: MSE:  
Sheep.

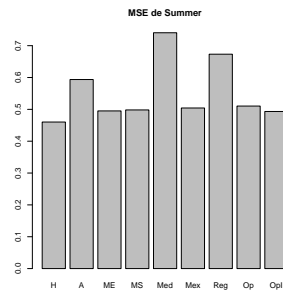


Figura 4.32: MSE:  
Summer.

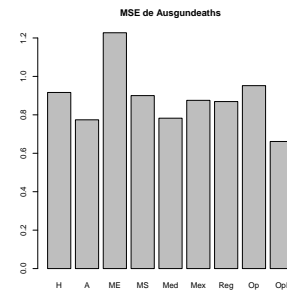


Figura 4.33: MSE:  
Ausgundeaths.

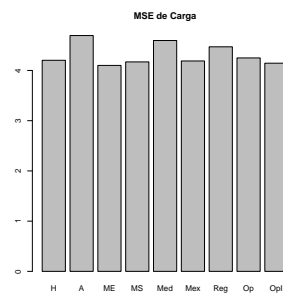


Figura 4.34: MSE:  
Carga.



## 5 Discussão e Conclusão

O objetivo deste trabalho foi verificar o desempenho de diferentes combinações de previsões, através de um estudo empírico, comparando-o com o desempenho de cada método individual de previsão (ARIMA, AEH, ME). Usamos o Erro Quadrático Médio (MSE) como medida de precisão, assim como a maioria dos estudos sobre combinação de previsões.

Segundo [Box e Jenkins, 1994], os modelos ARIMA funcionam bem nas séries que possuem pelo menos 50 observações para a estimação do modelo. Dentre os métodos de previsão individual, podemos perceber pela Tabela 4.7 que o ARIMA obteve um melhor desempenho em seis séries dentre as dez estudadas, apesar das séries não serem tão longas quanto desejável. Já o método AEH obteve melhor desempenho em três séries. O desempenho inferior, em relação ao ARIMA, talvez seja devido à dificuldade de implementação do método, da otimização das constantes de amortecimento e até mesmo da escolha dos valores iniciais. O modelo estrutural apresentou um melhor desempenho em apenas uma série, o que mostra o quanto é complicado trabalhar com esse método iterativo, que utiliza o filtro de Kalman.

Em geral a combinação de previsões apresentou um melhor desempenho do que as previsões individuais, confirmando os estudos feitos por [Lemke e Gabrys, 2010], [Menezes et al., 2000] e [Zou e Yang, 2004]. Considerando as combinações com reestimação dos parâmetros, percebemos que a regressão obteve um melhor desempenho em cinco das dez séries, concordando com a revisão de [Menezes et al., 2000], que cita vários estudos que mostram a eficiência da regressão, e argumentam que a regressão é melhor que a combinação ótima quando as previsões individuais são tendenciosas. Analisando a Tabela 4.7, percebemos que em oito séries a regressão teve realmente um desempenho melhor do que a combinação ótima. Segundo os estudo de [Menezes et al., 2000], o método de combinação ótima é equivalente a uma regressão de mínimos quadrados em

que a constante é suprimida e os pesos são restritos de forma a somarem à unidade. O problema com esta abordagem está no fato de se ter que estimar a matriz  $\mathbf{S}$  de correlação dos erros. [Menezes et al., 2000] afirma ainda que alguns estudos tem mostrado que as combinações que assumem independência entre os erros de previsão individuais apresentam um desempenho consideravelmente melhor do que aqueles que tentam dar conta de correlação; entretanto nosso estudo aponta o contrário. Observando a Tabela 4.7 temos que em 60% das séries a combinação ótima, que considera a correlação entre os erros de previsão, teve um desempenho melhor do que a combinação ótima com independência.

Em suma, as conclusões foram: (a) combinar as diversidades de métodos de previsão individuais garantiu efetivamente uma melhor precisão das previsões; (b) a combinação por regressão mostrou melhor desempenho que as outras combinações. Entretanto, seria interessante fazer as análises levando em consideração outras características das séries de erros, como a estabilidade da variância; a normalidade (ou pelo menos a simetria da distribuição) dos erros; e autocorrelação, pois segundo a revisão de [Menezes et al., 2000] os erros da regressão tendem a ser autocorrelacionados. Estas análises, assim como um estudo sobre o número ideal de previsões individuais a serem combinadas, ou sobre a importância do número de observações das séries como fator para a escolha do método de combinação e um estudo de simulação, são sugestões para pesquisas futuras nesta área.



```

# Thp(t): Theta(t/t-1) (Theta 'previsto')
# Thc(t): Theta(t/t) (Theta 'corrigido')
# A mesma convencao vale para o R: Rp, Rc
# Thps,Thcs,Rps,Rcs,es: vetores contendo as series de
Thp,Thc,Rp,Rc,e
#
##### apriori#####
# Calcula media e variancia apriori
apriori=function(Thc,Rc,G,W) {
  Thp=G%*%Thc
  Rp=G%*%Rc%*%t(G)+W
  mat=cbind(Thp,Rp)
  return(mat)}
#
##### aposteriori#####
# Calcula media e variância aposteriori
aposteriori=function(Thp,Rp,FF,V,e){
  ss= Rp%*% t(FF)%*% solve(V+FF%*%Rp%*%t(FF))
  Thc= Thp + ss%*% e
  Rc = Rp - ss%*% FF%*%Rp
  mat=cbind(Thc,Rc)
  return(mat)}
#
##### fk_forward #####
# INPUT
# serie : valores observados (nx1)
# F,G : matrizes multiplicadoras (1xd, dxd)
# V : variancia do erro (1x1)
# M : variancia do erro w (dxd)
# estado: vetor de estado (dx1)
#
# OUTPUT (dx8)
# Formato : [y Thcs Rcs] onde as colunas sao
# coluna 1: valores observados (x) (dx1)
# colunas 2-3: vetores de estado (Thcs) [mu b] (dx2)
# colunas 4-7: matrizes de covariancia (Rcs) (dx4)
# na forma [s11 s12 s21 s22]
# coluna 8: valores previstos (yp) (dx1)
#
# Usa funções 'apriori', 'aposteriori'.
fk_forward=function(serie,FF,G,v,W,estado) {
  # % Loading e initialization
  n=dim(serie)[1]
  d=dim(FF)[2] # número de colunas
  Thc=rbind(estado)
  Rc=diag(1,2)
  #
  xx=apriori(Thc,Rc,G,W)
  Thp=cbind(xx[,1])
  Rp=xx[,2:3]
  yp=matrix(c(FF%*% Thp),1,1)
  e=serie[1,1]- FF%*% Thp
  es=matrix(c(e),1,1)
  #
  xx=aposteriori(Thp,Rp,FF,v,e)
  Thc=cbind(xx[,1])
  Rc=xx[,2:3]
  Thcs=cbind(t(Thc))
  Rcs=matrix(c(Rc[1,],Rc[2,]),1,4)
  #
  # ----- LOOP -----
  for(i in 2:n){
    xx=apriori(Thc,Rc,G,W)
    Thp=cbind(xx[,1])
    Rp=xx[,2:3]
    yp=cbind(yp,FF%*% Thp)
    e=serie[i,1]-yp[1,i]
    es=cbind(es,e)
  }
  #
  xx=aposteriori(Thp,Rp,FF,v,e)
  Thc=cbind(xx[,1])
  Rc=xx[,2:3]
  Thcs=rbind(Thcs,t(Thc))
  aux=matrix(c(Rc[1,],Rc[2,]),1,4)
  Rcs=rbind(Rcs,aux)}
  saida=cbind(serie,Thcs,Rcs,t(yp))
  return(saida)}
#
##### Var_Mod_Est_Op #####
# Otimiza a variância do ME
# os parâmetros a e b são os limites do trecho da série para
# analisar os erros.
# z é a série a ser estudada, sigs é o vetor da variâncias #
# iniciais do modelo.
# param é o vetor de variancias do modelo.
#
## Var_Mod_Est_Opt=function(z,a,b,sigs)## {
# Função onde param será otimizado
  Mod_Est=function(param) {
    mu=z[1,1]

```

```

bet=0
sig_a=param[1]
sig_b=param[2]
sig_c=param[3]
# definicao das matrizes do modelo.
FF=matrix(c(1,0),1,2)
G=matrix(c(1,0,1,1),2,2)
v=sig_a^ 2
estado=matrix(c(mu,bet),2,1)
W=matrix(c(sig_b^ 2,0,0,sig_c^ 2),2,2)
# amortecimento usando as equacoes de Infield
v=1
lambda=(sig_c/sig_a)^ 2
Qr=matrix(0,2,2)
Qr[2,2]=lambda
x=fk_forward(z,FF,G,v,Qr,estado)
prevs=x[,8]
# cálculo dos erros
Z=z[,1]
erro=erros(Z,c(a,b),prevs)
return(erro[2])}
op=optim(sigs,Mod_Est,method="BFGS") # otimizando
return(op)}
#
#### Mod_Est_opt ####
# Função para previsão onde param já está otimizado
# nome=nome da série tem que ser entre aspas
#
Mod_Est_opt=function(z,param,nome) {
  mu=z[1,1]
  bet=0
  sig_a=param[1]
  sig_b=param[2]
  sig_c=param[3]
  #% definicao das matrizes do modelo.
  FF=matrix(c(1,0),1,2)
  G=matrix(c(1,0,1,1),2,2)
  v=sig_a^ 2
  estado=matrix(c(mu,bet),2,1)
  W=matrix(c(sig_b^ 2,0,0,sig_c^ 2),2,2)
  # amortecimento usando as equacoes de Infield
  v=1
  lambda=(sig_c/sig_a)^2
  Qr=matrix(0,2,2)
  Qr[2,2]=lambda
  x=fk_forward(z,FF,G,v,Qr,estado)
  prevs=x[,8]
  Z=z[,1]
  nome=c("Previsão por Modelo Estrutural de Tendência
Local da série",nome)
  plot(Z,lwd=2,type='l',ylim=c(min(Z,prevs[-c(1,2)]),max(Z,prevs[-
c(1,2)])),
  main=nome, xlab="Tempo")
  lines(prevs,col="red",lwd=2)
  return(prevs)}

Matriz de peso para Optimal
# A variável Ind indaga se é para ser Optimal com inde-
pendência, default é sem independência.
Peso.function=function(serie,serie_holt,serie_arima,serie_ME,Ind=FALSE)
{
  Erro_serie=cbind(serie-serie_holt,serie-serie_arima,serie-serie_ME)
  S=cov(Erro_serie[c(-1,-2),])
  E=cbind(rep(1,3))
  if (Ind==TRUE){
    v=diag(S)
    v=rbind(v,v,v)
    S=v*diag(3)}
  w=(solve(S)%*%E)/as.numeric((t(E)%*%solve(S)%*%E))
  #matrix de pesos.
  rownames(w)=list("Peso_holt","Peso_arima","Peso_ME")
  return(w)}

```

## A.2 Modelos ARIMA

Os parâmetros dos modelos ARIMA estimados pelo R para cada série seguem abaixo. Os valores entre parênteses são os desvios-padrão dos parâmetros.

- **Cow:**

$$\hat{Z}_t = 8,567583(3,1487) + 0,8463(0,1261)Z_{t-1} - 0,6612(0,1691)a_{t-1} + 0,1296(0,1668)a_{t-2}$$

- **Buffsnow:**

$$\hat{Z}_t = 24,2521(6,2927) + 0,6980(0,25127)Z_{t-1} - 0,4129(0,2920)a_{t-1}$$

- **Calfem:**

$$\hat{Z}_t = 39,671(0,7464)$$

- **DowJones:**

$$\hat{Z}_t = 0,0196(0,1205) + 0,8840(0,1932)Z_{t-1} - 0,5117(0,2577)a_{t-1} - 0,1873(0,1595)a_{t-2}$$

- **Globtp:**

$$\hat{Z}_t = 0,0034(0,0027) + 0,4019(0,1382)Z_{t-1} - 0,8881(0,0778)a_{t-1}$$

- **Huron:**

$$\hat{Z}_t = 1,855(0,3658) + 1,0559(0,1098)Z_{t-1} - 0,2575(0,1140)Z_{t-2}$$

- **Sheep:**

$$\hat{Z}_t = -7,0096(6,9220) + 0,9984(0,1837)Z_{t-1} - 0,6167(,1130)Z_{t-2} - 0,5294(0,226)a_{t-1}$$

- **Summer:**

$$\hat{Z}_t = 10,7098(0,1330) + 0,2983(0,1062)Z_{t-1}$$

- **Ausgundeaths:**

$$\hat{Z}_t = 0,0011(0,0233) + 0,8524(0,0683)Z_{t-1} - 1(0,0369)a_{t-1}$$

- **Carga:**

$$\hat{Z}_t = -0,0770(0,1851) - 0,1445(0,1233)a_{t-1}$$

## Referências Bibliográficas

- [Barros, 2011] Barros, M. *Capítulo 5: Introdução às Séries Temporais e aos Modelos ARIMA*. Disponível em: <[http://www.mbarros.com/documentos/upload/Livro\\_Processos\\_Cap\\_5\\_parte.pdf](http://www.mbarros.com/documentos/upload/Livro_Processos_Cap_5_parte.pdf)>. Acessado em: 05 de out. de 2011.
- [Bates e Granger, 1969] Bates, J.; Granger, C. The combination of forecasts. *Operations Research Quarterly*, 20, 451-468.
- [Box e Jenkins, 1994] Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. *Time series analysis-forecasting and control*, 3ed. New Jersey: Prentice Hall.
- [Clemen, 1989] Clemen, R.T. Combining forecasts: A review and annotated bibliography. *Int J of Forecasting*, 5, 559-583.
- [Ehlers, 2011] Ehlers, R.S. *Análise de Séries Temporais*. Curso de Séries Temporais . Agosto de 2009. Notas de Aula. Digitalizado. Disponível em: <<http://www.icmc.usp.br/~ehlers/stemp/stemp.pdf>>. Acessado em: 05 de Abril de 2011.
- [Hyndman, 2011] <<http://robjhyndman.com/TSDL/>>. Acessado em: 20 de julho de 2011.
- [Hippert, 2010] Hippert, H.S. *Séries Temporais I*. Disciplina: Análise e Previsão de Séries Temporais I. Agosto-dezembro de 2010 Notas de Aula. Digitalizado.
- [Hippert, 2011] Hippert, H.S. *Séries Temporais II*. Disciplina: Análise e Previsão de Séries Temporais II. Março-Julho de 2011 Notas de Aula. Digitalizado.



- [Hollauer et al., 2008] Hollauer, G.; Issler, J.V.; Notini, H.H. Prevendo o Crescimento da Produção Industrial Usando um Número Limitado de Combinações de Previsões. *Economia Aplicada*, v.12, p.177-198.
- [Lemke e Gabrys, 2008] Lemke, C. and Gabrys, B. *Forecasting and Forecast Combination in Airline Revenue Management Applications*. In: Nguyen, N.T., Kolaczek, G. and Gabrys, B., eds. Knowledge Processing and Reasoning for Information Society. Warsaw, Poland: EXIT Publishing House, p. 231-247. Disponível em: <<http://eprints.bournemouth.ac.uk/8502/>>. Acessado em: 02 de Ago.2011.
- [Lemke e Gabrys, 2010] Lenke, C.; Gabrys, B. Meta-Learning for Time Series Forecasting and Forecast Combination. *Neurocomputing*, v.73, p.2006-2016.
- [Makridakis et al., 1998] Makridakis, S.; Wheelwright, S.C.; Hyndman, R.J. *Forecasting: Methods and Applications* 3ed. New York: John WILEY & Sons.
- [Menezes et al., 2000] Menezes, L. M. de; Bunn, D.W.; Taylor, J.W. Review of Guidelines for the Use of Cambined Forecasting. *European Journal of Operational Research*, v.120, p.190-204.
- [Montgomery et al., 1990] Montgomery, D. C.; Johnson, L. A.; Gardner, J. S. *Forecasting and time series analysis*. 2ed. NY: McGraw-Hill.
- [Morettin e Toloi, 2006] Morettin, P.A.; Toloi, C.M. *Análise de Séries Temporais*. 2ed. São Paulo: Egard Blucher.
- [Newbold e Granger, 1974] Newbold, P.; Granger, C. W. J. Experience with forecasting univariate times series and the combination of forecasts. *Journal of the Royal Statistical Society*, ser. A, vol. 137, pp 131-146
- [Rocha, 2008] Rocha, V.B. *Uma Abordagem de WAVELETS Aplicada à Combinação de Previsões: Uma Análise Teórica e Experimental*. 155f. Dissertação (Mestrado em Ciências)-Universidade Federal do Paraná, Curitiba.

- 
- [Zou e Yang, 2004] Zou, H.; Yang, Y. Combining Time Series Model for Forecasting.  
*International Journal of Forecasting*, v.20, p.69-84.