



Pós-Graduação em Ciência da Computação

**Extensão de técnicas clássicas para análise  
de séries temporais do tipo intervalo**

**Por**

***André Luis Santiago Maia***

**Tese de doutorado**



Universidade Federal de Pernambuco  
posgraduacao@cin.ufpe.br  
www.cin.ufpe.br/~posgraduacao

RECIFE, MARÇO/2010



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE INFORMÁTICA  
PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

ANDRÉ LUIS SANTIAGO MAIA

EXTENSÃO DE TÉCNICAS CLÁSSICAS PARA ANÁLISE  
DE SÉRIES TEMPORAIS DO TIPO INTERVALO

*ESTE TRABALHO FOI APRESENTADO À PÓS-GRADUAÇÃO EM  
CIÊNCIA DA COMPUTAÇÃO DO CENTRO DE INFORMÁTICA DA  
UNIVERSIDADE FEDERAL DE PERNAMBUCO COMO REQUISITO  
PARCIAL PARA OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIA  
DA COMPUTAÇÃO.*

*ORIENTADOR: PROF. DR. FRANCISCO DE A. T. DE CARVALHO*

RECIFE, MARÇO/2010

**Maia, André Luis Santiago**  
**Extensão de técnicas clássicas para análise de séries temporais do tipo intervalo / André Luis Santiago Maia. - Recife: O Autor, 2010.**  
**xiv, 239 folhas : il., fig., tab.**

**Tese (doutorado) – Universidade Federal de Pernambuco. Cln. Ciência da computação, 2010.**

**Inclui bibliografia e apêndice.**

**1. Inteligência artificial. 2. Análise de séries temporais. 3. Inteligência computacional. I. Título.**

**006.3**

**CDD (22. ed.)**

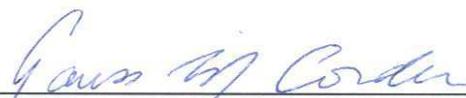
**MEI2010 – 035**

Tese de Doutorado apresentada por **André Luis Santiago Maia**, a Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, sob o título “**Extensão de Técnicas Clássicas para Análise de Séries Temporais do Tipo Intervalo**”, elaborada sob a orientação do **Prof. Francisco de Assis Tenório de Carvalho** e aprovada pela Banca Examinadora formada pelos professores:



---

Profa. Teresa Bernarda Ludermir  
Centro de Informática / UFPE



---

Prof. Gauss Moutinho Cordeiro  
Departamento de Estatística e Informática / UFRPE



---

Prof. Gerson Zaverucha  
Programa de Engenharia de Sistemas / UFRJ



---

Prof. Ricardo José Gabrielli Barreto Campello  
Instituto de Ciências Matemáticas e de Computação / USP



---

Prof. Eufrásio de Andrade Lima Neto  
Departamento de Estatística / UFPB

Visto e permitida a impressão.  
Recife, 08 de março de 2010.



---

**Prof. NELSON SOUTO ROSA**

Coordenador da Pós-Graduação em Ciência da Computação do  
Centro de Informática da Universidade Federal de Pernambuco.

*Dedico este trabalho à minha esposa  
& aos meus pais.*

# Agradecimentos

Expresso toda minha sincera gratidão a todos que, direta ou indiretamente, contribuíram e tornaram possível o desenvolvimento deste trabalho. Sei que cometerei injustiça citando apenas alguns nomes, mas citarei os que foram realmente imprescindíveis.

- A Deus, por permitir que eu cumprisse este objetivo.
- À minha esposa, Gecy, principal pessoa nesta caminhada, pelo companheirismo e colaboração e pelos momentos de alegria nas horas vagas indispensáveis para a conclusão do trabalho. Só ela sabe com exatidão o quanto foi difícil a trajetória percorrida até este momento.
- Ao Prof. Francisco de A. T. de Carvalho, meu orientador, pela inserção nesta área de pesquisa promissora, pela compreensão e pela confiança necessárias para a conclusão e para o êxito deste trabalho. A orientação do Prof. Francisco é exemplar.
- Aos meus queridos pais e irmãos, pelo que sou hoje e por terem sempre acreditado em meu potencial, dando-me o apoio necessário em todos os sentidos sem medir esforços. A nossa união é fundamental para o meu sucesso.
- Aos meus amigos Liles e Tina, pelo amparo na etapa mais crítica do doutorado que ficará guardado em minha memória para sempre. Sem eles dois, o doutorado não teria, sequer, iniciado. Agradeço a eles também os inúmeros momentos de companhia e descontração.
- Aos membros da banca de qualificação, Prof. Gauss Cordeiro, Profa. Teresa Ludermir e Prof. André Ponce de Leon, pelas valiosas críticas e sugestões.

Agradeço ao Prof. Gauss também pelo incentivo desde a graduação e à Profa. Teresa pela contribuição em redes neurais.

- À Profa. Cristina Raposo do Departamento de Estatística da UFPE, pelo apoio e conselhos e pela indicação do Programa de Pós-graduação em Ciência da Computação da UFPE.
- Aos professores deste programa de pós-graduação, pelos conhecimentos transmitidos que foram fundamentais para elaboração deste trabalho.
- A Eufrásio Lima Neto, enquanto colega de curso, pelas dicas e pelo auxílio e ao bolsista de iniciação científica Fernando Braga e Silva pela contribuição na implementação dos códigos em R.
- À minha enteada, Larissa, e à minha sobrinha, “Sá” Duda, por me fazerem sorrir mesmo em momentos de elevado estresse. A elas, eu peço desculpa pela reduzida dedicação durante estes anos.
- A todos os colegas da Fundação Joaquim Nabuco, em especial, a Ana Eliza e a Luis, pela compreensão, aos companheiros Wilson Fusco e Alexandre Zarias e à amiga Darcilene Gomes, pelas agradáveis e importantes conversas.
- Ao CNPq, pelo suporte financeiro no início do curso.
- Por fim, aos membros da banca examinadora da defesa de tese.

# Sumário

Dedicatória	ii
Agradecimentos	iii
Lista de tabelas	viii
Lista de figuras	xi
Resumo	1
Abstract	2
<b>1 Introdução e motivação</b>	<b>3</b>
1.1 Objetivos . . . . .	5
1.2 Organização da tese . . . . .	6
<b>2 Análise de dados simbólicos</b>	<b>8</b>
2.1 Dados simbólicos . . . . .	10
2.2 Séries temporais do tipo intervalo . . . . .	15
2.3 Alguns métodos estendidos para dados do tipo intervalo . . . . .	16
2.4 Considerações finais . . . . .	19
<b>I Métodos de ajuste independente</b>	<b>21</b>
<b>3 Métodos de ajuste independente</b>	<b>22</b>
3.1 Modelos de regressão para dados intervalares . . . . .	22
3.2 Modelos de ajuste independente para STI . . . . .	29

3.2.1	Modelos ARMA nos centros e nas amplitudes de STI . . . . .	31
3.2.2	Rede neural MLP nos centros e nas amplitudes de STI . . . . .	33
3.2.3	Modelo híbrido nos centros e nas amplitudes de STI . . . . .	34
3.2.4	Considerações finais . . . . .	37
<b>II Métodos de ajuste dependente</b>		<b>39</b>
<b>4</b>	<b>Métodos de ajuste dependente</b>	<b>40</b>
4.1	Metodologia adotada . . . . .	42
4.2	Técnicas descritivas . . . . .	45
4.2.1	Representação gráfica de STI . . . . .	45
4.2.2	Autocorrelação e autocovariância em STI . . . . .	48
4.3	Objetivos da previsão de STI . . . . .	55
4.4	Métodos de alisamento exponencial para STI . . . . .	58
4.4.1	Alisamento exponencial simples para STI . . . . .	59
4.4.2	Alisamento exponencial duplo para STI . . . . .	62
4.4.3	Método Holt–Winters para STI . . . . .	65
4.5	Modelos para STI . . . . .	74
4.5.1	Modelos iARMA . . . . .	77
4.5.2	Estimação irrestrita de modelos iARMA . . . . .	82
4.5.3	Estimação restrita de modelos iARMA . . . . .	89
4.5.4	Modelos iARIMA . . . . .	97
4.6	Métodos baseados em redes neurais para STI . . . . .	102
4.6.1	Redes neurais MLP para STI . . . . .	103
4.6.2	Modelo híbrido para STI . . . . .	109
4.7	Medidas de precisão de previsões para STI . . . . .	113
<b>5</b>	<b>Resultados experimentais</b>	<b>117</b>
5.1	Experimentos com STI sintéticas . . . . .	123
5.1.1	Avaliação da ACF <sup>I</sup> para STI sintéticas . . . . .	124
5.1.2	Comparação dos modelos no ajuste de STI sintéticas . . . . .	134
5.2	Aplicação a STI reais . . . . .	148

<b>6</b>	<b>Conclusões e trabalhos futuros</b>	<b>157</b>
6.1	Conclusões . . . . .	157
6.2	Trabalhos futuros . . . . .	162
<b>A</b>	<b>Tabelas e gráficos dos experimentos</b>	<b>164</b>
A.1	Tabelas e gráficos dos experimentos com STI sintéticas . . . . .	165
A.2	Tabelas e gráficos da aplicação a STI reais . . . . .	201
<b>B</b>	<b>Artigos publicados e submetidos</b>	<b>231</b>
	<b>Referências bibliográficas</b>	<b>233</b>

# Lista de Tabelas

2.1	Descrição de 600 aves através de três variáveis. . . . .	11
2.2	Descrição de três espécies de aves através de três variáveis. . . . .	11
2.3	Exemplo de uma base de dados clássicos. . . . .	12
2.4	Exemplo de uma base de dados simbólicos. . . . .	13
5.1	Nome e descrição dos métodos e modelos utilizados na previsão de STI. . . . .	119
5.2	Função de autocorrelação amostral para as séries $X_t^U$ , $X_t^L$ , $X_t^c$ e $\mathbf{I}_t$ . . . . .	133
5.3	Nome e descrição dos PGDs utilizados na geração de STI. . . . .	136
5.4	Número referente a cada modelo nas Figuras A.1 a A.18 e nas Figuras A.19 a A.33 apresentadas no Apêndice A. . . . .	140
5.5	Descrição das STI reais analisadas. . . . .	150
5.6	Período e tamanho das STI reais analisadas. . . . .	150
A.1	Resultados obtidos para o PGD-I com $n = 220$ ; conjunto de treinamento. . . . .	165
A.2	Resultados obtidos para o PGD-I com $n = 220$ ; 5 passos-à-frente. . . . .	167
A.3	Resultados obtidos para o PGD-I com $n = 220$ ; 20 passos-à-frente. . . . .	169
A.4	Resultados obtidos para o PGD-I com $n = 620$ ; conjunto de treinamento. . . . .	171
A.5	Resultados obtidos para o PGD-I com $n = 620$ ; 5 passos-à-frente. . . . .	173
A.6	Resultados obtidos para o PGD-I com $n = 620$ ; 20 passos-à-frente. . . . .	175
A.7	Resultados obtidos para o PGD-II com $n = 220$ ; conjunto de treinamento. . . . .	177
A.8	Resultados obtidos para o PGD-II com $n = 220$ ; 5 passos-à-frente. . . . .	179
A.9	Resultados obtidos para o PGD-II com $n = 220$ ; 20 passos-à-frente. . . . .	181

A.10 Resultados obtidos para o PGD-II com $n = 620$ ; conjunto de treinamento. . . . .	183
A.11 Resultados obtidos para o PGD-II com $n = 620$ ; 5 passos-à-frente. . . . .	185
A.12 Resultados obtidos para o PGD-II com $n = 620$ ; 20 passos-à-frente. . . . .	187
A.13 Resultados obtidos para o PGD-III com $n = 220$ ; conjunto de treinamento. . . . .	189
A.14 Resultados obtidos para o PGD-III com $n = 220$ ; 5 passos-à-frente. . . . .	191
A.15 Resultados obtidos para o PGD-III com $n = 220$ ; 20 passos-à-frente. . . . .	193
A.16 Resultados obtidos para o PGD-III com $n = 620$ ; conjunto de treinamento. . . . .	195
A.17 Resultados obtidos para o PGD-III com $n = 620$ ; 5 passos-à-frente. . . . .	197
A.18 Resultados obtidos para o PGD-III com $n = 620$ ; 20 passos-à-frente. . . . .	199
A.19 Resultados obtidos para a STI de preços de ações da Petrobras; conjunto de treinamento. . . . .	201
A.20 Resultados obtidos para a STI de preços de ações da Petrobras; 5 passos-à-frente. . . . .	203
A.21 Resultados obtidos para a STI de preços de ações da Petrobras; 20 passos-à-frente. . . . .	205
A.22 Resultados obtidos para a STI de preços de ações da Google; conjunto de treinamento. . . . .	207
A.23 Resultados obtidos para a STI de preços de ações da Google; 5 passos-à-frente. . . . .	209
A.24 Resultados obtidos para a STI de preços de ações da Google; 20 passos-à-frente. . . . .	211
A.25 Resultados obtidos para a STI de preços de ações da Gol; conjunto de treinamento. . . . .	213
A.26 Resultados obtidos para a STI de preços de ações da Gol; 5 passos-à-frente. . . . .	215
A.27 Resultados obtidos para a STI de preços de ações da Gol; 20 passos-à-frente. . . . .	217
A.28 Resultados obtidos para a STI de temperaturas no Rio de Janeiro; conjunto de treinamento. . . . .	219

A.29 Resultados obtidos para a STI de temperaturas no Rio de Janeiro; 5 passos-à-frente. . . . .	221
A.30 Resultados obtidos para a STI de temperaturas no Rio de Janeiro; 20 passos-à-frente. . . . .	223
A.31 Resultados obtidos para a STI de temperaturas em Nova Iorque; conjunto de treinamento. . . . .	225
A.32 Resultados obtidos para a STI de temperaturas em Nova Iorque; 5 passos-à-frente. . . . .	227
A.33 Resultados obtidos para a STI de temperaturas em Nova Iorque; 20 passos-à-frente. . . . .	229

# Lista de Figuras

3.1	Diagrama da metodologia híbrida de Zhang (2003) [94]. . . . .	36
4.1	Séries dos limites superior e inferior dos intervalos para a STI de temperaturas da China (gráfico à esquerda) e de preços de ações da Petrobras (gráfico à direita). . . . .	43
4.2	Séries dos centros (linha sólida) e das amplitudes (linha tracejada) dos intervalos para a STI de temperaturas da China (gráfico à esquerda) e de preços de ações da Petrobras (gráfico à direita). . . . .	43
4.3	Exemplo de concepção de um intervalo. . . . .	46
4.4	Séries temporais do tipo intervalo (gráficos à direita) obtidas a partir de séries temporais de dados clássicos (gráficos à esquerda). . . . .	47
4.5	Série dos intervalos de temperaturas mensais medidas em estação meteorológica na China. . . . .	49
4.6	Série dos intervalos de preços diários de ações da Petrobras. . . . .	49
4.7	Correlograma (gráfico inferior) da STI de temperaturas mensais medidas em estação meteorológica na China. . . . .	56
4.8	Correlograma (gráfico inferior) da STI de preços diários de ações da Petrobras. . . . .	57
4.9	Ajuste da STI de preços de ações da Petrobras via método AES <sup>I</sup> . . . . .	72
4.10	Ajuste da STI de preços de ações da Petrobras via método AED <sup>I</sup> . . . . .	72
4.11	Ajuste da STI de temperaturas na China via método HW <sup>I</sup> . . . . .	73
4.12	Diagrama da metodologia Box–Jenkins. . . . .	78
4.13	Ajuste da STI de preços de ações da Petrobras via modelo iAR(1). . . . .	88
4.14	Ajuste da STI de preços de ações da Petrobras via modelo iARMA(1,1). . . . .	88
4.15	Ajuste da STI de preços de ações da Petrobras via modelo iAR(1)-restrito. . . . .	98

4.16	Ajuste da STI de preços de ações da Petrobras via modelo iARMA(1,1)-restrito. . . . .	98
4.17	Estrutura de uma rede neural MLP com $p$ entradas, uma camada escondida com $q$ nodos e uma saída para previsão de séries temporais univariadas. . . . .	104
4.18	Estrutura da rede neural MLP <sup>I</sup> com $2p$ entradas, uma camada escondida com $q$ nodos e duas saídas para previsão de STI. . . . .	106
4.19	Ajuste da STI de preços de ações da Petrobras via rede neural MLP <sup>I</sup> . . . . .	108
4.20	Diagrama do modelo híbrido de Zhang (2003) para previsão de séries temporais univariadas. . . . .	111
4.21	Diagrama do modelo híbrido para previsão de STI. . . . .	113
5.1	Correlogramas em processos intervalares puramente aleatórios. . . . .	127
5.2	Correlogramas em processos intervalares com correlação de curto-prazo. . . . .	128
5.3	Correlogramas em processos intervalares iAR(2). . . . .	129
5.4	Correlogramas em processos intervalares com correlação negativa. . . . .	130
5.5	Correlogramas em processos intervalares com tendência. . . . .	131
5.6	Correlogramas em processos intervalares com sazonalidade. . . . .	132
5.7	STI real de preços diários de ações da Petrobras. . . . .	151
5.8	STI real de preços diários de ações da Google. . . . .	152
5.9	STI real de preços diários de ações da Gol. . . . .	152
5.10	STI real de temperaturas diárias no Rio de Janeiro. . . . .	153
5.11	STI real de temperaturas diárias em Nova Iorque. . . . .	153
A.1	Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com $n = 220$ ; conjunto de treinamento. . . . .	166
A.2	Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com $n = 220$ ; 5 passos-à-frente. . . . .	168
A.3	Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com $n = 220$ ; 20 passos-à-frente. . . . .	170
A.4	Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com $n = 620$ ; conjunto de treinamento. . . . .	172
A.5	Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com $n = 620$ ; 5 passos-à-frente. . . . .	174

A.6	Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com $n = 620$ ; 20 passos-à-frente. . . . .	176
A.7	Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com $n = 220$ ; conjunto de treinamento. . . . .	178
A.8	Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com $n = 220$ ; 5 passos-à-frente. . . . .	180
A.9	Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com $n = 220$ ; 20 passos-à-frente. . . . .	182
A.10	Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com $n = 620$ ; conjunto de treinamento. . . . .	184
A.11	Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com $n = 620$ ; 5 passos-à-frente. . . . .	186
A.12	Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com $n = 620$ ; 20 passos-à-frente. . . . .	188
A.13	Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com $n = 220$ ; conjunto de treinamento. . . . .	190
A.14	Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com $n = 220$ ; 5 passos-à-frente. . . . .	192
A.15	Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com $n = 220$ ; 20 passos-à-frente. . . . .	194
A.16	Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com $n = 620$ ; conjunto de treinamento. . . . .	196
A.17	Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com $n = 620$ ; 5 passos-à-frente. . . . .	198
A.18	Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com $n = 620$ ; 20 passos-à-frente. . . . .	200
A.19	Valores das medidas de precisão obtidos para a STI da Petrobras; conjunto de treinamento. . . . .	202
A.20	Valores das medidas de precisão obtidos para a STI da Petrobras; 5 passos-à-frente. . . . .	204
A.21	Valores das medidas de precisão obtidos para a STI da Petrobras; 20 passos-à-frente. . . . .	206

A.22	Valores das medidas de precisão obtidos para a STI da Google; conjunto de treinamento. . . . .	208
A.23	Valores das medidas de precisão obtidos para a STI da Google; 5 passos-à-frente. . . . .	210
A.24	Valores das medidas de precisão obtidos para a STI da Google; 20 passos-à-frente. . . . .	212
A.25	Valores das medidas de precisão obtidos para a STI da Gol; conjunto de treinamento. . . . .	214
A.26	Valores das medidas de precisão obtidos para a STI da Gol; 5 passos-à-frente. . . . .	216
A.27	Valores das medidas de precisão obtidos para a STI da Gol; 20 passos-à-frente. . . . .	218
A.28	Valores das medidas de precisão obtidos para a STI do Rio de Janeiro; conjunto de treinamento. . . . .	220
A.29	Valores das medidas de precisão obtidos para a STI do Rio de Janeiro; 5 passos-à-frente. . . . .	222
A.30	Valores das medidas de precisão obtidos para a STI do Rio de Janeiro; 20 passos-à-frente. . . . .	224
A.31	Valores das medidas de precisão obtidos para a STI de Nova Iorque; conjunto de treinamento. . . . .	226
A.32	Valores das medidas de precisão obtidos para a STI de Nova Iorque; 5 passos-à-frente. . . . .	228
A.33	Valores das medidas de precisão obtidos para a STI de Nova Iorque; 20 passos-à-frente. . . . .	230

# Resumo

Os dados simbólicos apresentam, em sua estrutura, formas interessantes para se transformar grandes bases de dados clássicos em novos conjuntos de dados de tamanho reduzido, facilitando a manipulação e proporcionando novas técnicas de análise dos mesmos. No entanto, mesmo com os recentes avanços promovidos por pesquisadores nesta área, o volume de técnicas de manipulação e, conseqüentemente, de análise de dados simbólicos (ADS) ainda é incipiente.

Uma série temporal do tipo intervalo (STI), no campo de dados simbólicos, pode ser definida como um conjunto de intervalos observados sequencialmente no tempo, em que cada intervalo é descrito por um vetor bidimensional com elementos em  $\mathbb{R}$  representados pelo limite superior e pelo limite inferior. O desenvolvimento de técnicas para previsão de STI é uma área de pesquisa muito promissora e os poucos resultados relatados na literatura surgiram muito recentemente.

Nesta tese, estendemos técnicas clássicas de análise de séries temporais para descrição, modelagem e previsão de STI no domínio de ADS. Neste contexto, nós apresentamos técnicas para descrição de uma STI, envolvendo cálculo de estatísticas sumárias e representação gráfica dos dados. Na modelagem, apresentamos métodos que consistem na explicação do processo gerador da STI a partir de certo modelo, bem como métodos de estimação de parâmetros e métodos para avaliação da qualidade do modelo, em termos do ajuste.

**Palavras-chave:** Análise de dados simbólicos, dados do tipo intervalo, análise de séries temporais, alisamento exponencial, modelos ARMA e ARIMA, redes neurais artificiais.

# Abstract

Symbolic data analysis is a relatively new field that provides a range of methods for analyzing complex datasets. However, even with recent advances promoted by researchers in this area, the volume of technical manipulation and, consequently, analysis of symbolic data (ADS) is still incipient.

Interval-valued time series (STI) in the field of symbolic data, can be defined as a set of intervals observed sequentially in time, where each interval is described as a two-dimensional vector with components in  $\mathbb{R}$  representing the upper boundary and the lower boundary.

In the present thesis, we extend classical techniques of time series analysis for modeling and prediction of STI in the field of ADS. In this context, we present techniques for description of STI, involving calculation of summary statistics and graphical representation of data. We present methods involving the explanation of the process generating the STI from a certain model and parameter estimation methods.

**Keywords:** Symbolic data analysis, interval-valued data, time series analysis, exponential smoothing, ARMA and ARIMA models, artificial neural networks.

# Capítulo 1

## Introdução e motivação

É cada vez mais comum a necessidade de analisarmos grandes bases de dados e de se extrair deles informações relevantes, ou seja, conhecimento. Este é um dos objetivos da abordagem exposta nesta tese que está incluída no que se denomina Análise de Dados Simbólicos (ADS). A ADS consiste em uma alternativa na mineração de grandes massas de dados, visto que ela possibilita a agregação/redução dos dados clássicos em uma estrutura complexa chamada de dados simbólicos. Os dados presentes numa base simbólica sumarizam as informações contidas em uma base de dados clássicos, apresentando-as de forma agregada. Assim, as variáveis contidas neste tipo de dados podem conter variação interna ou incertezas. Bock e Diday (2000) [11] apresentam de maneira sólida os principais conceitos da análise de dados simbólicos e os principais métodos estatísticos desenvolvidos para manipular dados desta natureza. Billard e Diday (2003) [9] atestam o crescimento de dados de natureza simbólica e alertam a necessidade do desenvolvimento de novas metodologias para o tratamento de informações dessa natureza.

Em análise de dados clássicos cada entrada da variável é representada por um único valor que pode ser qualitativo (categorias de um conjunto finito, por exemplo) ou quantitativo (valores são elementos de  $\mathbb{R}$ ). Entretanto, existem situações em que as informações são complexas para serem representadas através de dados desta natureza. Estas situações podem ser observadas quando as variáveis assumem mais que um simples valor para cada indivíduo. Como exemplo, podemos citar o tempo (hora inicial e hora final) usado por uma pessoa para estudar determinado assunto

que varia em diferentes dias, ou os meios de transporte utilizados em uma determinada cidade em um dado mês, que pode ser carro, ônibus, metrô, etc. No primeiro caso, as variáveis podem ser medidas como intervalos ( $[14\text{h};18\text{h}]$ ,  $[15\text{h};20\text{h}]$ , ...) e a segunda variável, em distribuição de frequência (30% de carro, 50% de ônibus, 20% de metrô). A necessidade de considerar dados que contêm informações que não podem ser representadas por modelos clássicos conduz ao desenvolvimento de ADS (Bock e Diday, 2000) [11].

Como caso particular de dados simbólicos, os dados do tipo intervalo aparecem quando os valores das variáveis são intervalos em  $\mathbb{R}$ . Eles são o caso mais simples de representar a variabilidade nos dados, tais como, temperatura mensal máxima e mínima (entre  $15^{\circ}\text{C}$  e  $25^{\circ}\text{C}$ ), preço de uma mercadoria em uma cidade (entre R\$ 5,00 e R\$ 10,00). Segundo Moore (1966) [73], no mundo real, observações e estimadores envolvem incertezas e, conseqüentemente, eles não representam os dados reais de forma exata. De acordo com o campo de análise intervalar introduzido por ele, se a precisão for necessária, os dados devem ser representados como os intervalos que incluem as quantidades reais.

Informalmente, uma série temporal do tipo intervalo, no campo de dados simbólicos, pode ser definida como uma lista de intervalos com valores em  $\mathbb{R}$  observados sequencialmente no tempo. Esta série temporal é obtida, geralmente, pela agregação espacial ou temporal de séries temporais de dados univalorados (dados clássicos). Referimos-nos a séries temporais de dados clássicos quando estamos falando de séries em que um único valor é observado em cada instante de tempo, por isso, ao longo do texto, denotaremos estas séries de séries temporais univaloradas. Como exemplo de série intervalar, citamos a amplitude dos preços diários de eletricidade em certa região, onde os intervalos observados são limitados pelos preços mínimos e máximos registrados durante o dia. No desenvolvimento de técnicas de análise de séries temporais do tipo intervalo, o instante em que é observado cada um dos intervalos deve ser registrado e usado na análise/modelagem da série. Desta forma, dado que temos uma série temporal do tipo intervalo, como resolver o problema de análise e previsão no contexto de ADS?

A maioria dos modelos de séries temporais usa dados de séries quantificados através de medidas pontuais (escalares). No entanto, em algumas situações reais, e.g., no mercado financeiro, muitas variáveis são delimitadas por intervalos para

um determinado período de tempo, o que pode tornar as medições pontuais inadequadas. Por exemplo, os valores de negociações de ações têm seus limites diários (ou semanais, ou mensais) representados pelo menor e maior valor, e varia em um determinado período de negociação – dia, semana ou mês. Representar as variações com simples medidas escalares, como o preço de fechamento, reflete apenas um determinado número (nível) em um momento específico, mas não reflete devidamente sua variabilidade durante o período de negociação. Além disso, pode ser mais conveniente para a tomada de decisão obter previsões sob a forma de uma amplitude ou de um intervalo, em vez de apenas valores pontuais ou escalares. Por exemplo, um intervalo previsto indicando que a taxa de câmbio no próximo mês estará no intervalo  $[1,956; 1,961]$  pode ser preferível a um simples valor numérico, digamos 1,957. Para abordar esta questão, o presente trabalho propõe um novo arcabouço de séries temporais que descreve séries intervalares e produz diretamente previsões em forma de intervalos.

A busca por extensões de métodos de análise estatística de dados clássicos para ADS vem crescendo devido à necessidade de se descobrir regularidades e resumir, de forma precisa, informações contidas em grandes conjuntos de dados. Neste contexto, esperamos com esta tese, não apenas estender a metodologia de estimação de modelos de séries temporais para análise de dados do tipo intervalo, mas também contribuir com novas metodologias baseadas em combinação de modelos, incluindo modelos de redes neurais artificiais, com a finalidade de melhorar a precisão na previsão de séries temporais intervalares.

## 1.1 **Objetivos**

O objetivo geral deste trabalho é estender técnicas de análise de séries temporais para descrição, modelagem e previsão de séries temporais do tipo intervalo no domínio de ADS. Neste contexto, nós apresentamos técnicas para descrição de uma série temporal do tipo intervalo, envolvendo cálculo de estatísticas sumárias em séries intervalares, representação gráfica dos dados e a detecção de características simples da série.

No contexto de modelagem de séries temporais do tipo intervalo (STI), o objetivo é apresentar métodos que consistem na explicação do processo gerador da STI a

partir de certo modelo, bem como métodos de estimação de parâmetros e métodos para avaliação da qualidade do modelo em termos do ajuste. Com isso, objetivamos contribuir com modelos capazes de realizar previsões precisas para intervalos futuros o que consiste em conhecer a evolução provável de uma série intervalar no futuro.

## 1.2 Organização da tese

A tese apresenta-se dividida em duas partes, Parte I e Parte II, e contém seis capítulos e dois apêndices. Na Parte I são apresentados os modelos que intitulamos de *Métodos de ajuste independente* e na Parte II são apresentados os modelos que intitulamos de *Métodos de ajuste dependente*.

Ainda neste Capítulo 1 tratamos, dentre outros tópicos, da apresentação do problema da pesquisa, da motivação que nos levou ao desenvolvimento deste trabalho, incluindo os objetivos da pesquisa.

No Capítulo 2, introduzimos a *Análise de Dados Simbólicos* (ADS) e alguns conceitos básicos sobre a ADS. Alguns conceitos e exemplos citados ao longo do capítulo foram extraídos do livro *Symbolic Data Analysis and the SODAS Software* editado por Diday e Noirhomme-Fraiture e publicado em 2008 [34]. Ainda no Capítulo 2 serão abordados alguns pontos principais necessários para entendimento da presente tese, incluindo a definição de STI. Também citamos alguns métodos já estendidos para dados simbólicos.

No Capítulo 3, apresentamos os primeiros modelos desenvolvidos para modelagem de dados de STI no contexto de ADS. A metodologia adotada no desenvolvimento destes modelos para modelagem e previsão de STI, a partir das séries dos centros e das amplitudes, segue em paralelo ao escopo principal desta tese.<sup>1</sup>

No Capítulo 4 encontra-se o escopo principal da tese, em que a metodologia adotada consiste no desenvolvimento de métodos baseados nos valores dos limites superiores e dos limites inferiores dos intervalos. Os modelos para previsão de STI desenvolvidos no Capítulo 4 fazem parte do que denominamos de classe de modelos dependentes. São apresentadas as principais técnicas de análise de séries temporais

---

<sup>1</sup>A metodologia principal adotada nesta tese para modelagem de STI baseia-se em modelos dependentes ajustados aos valores dos limites superiores e dos limites inferiores dos intervalos e será tratada na Parte II da tese.

para dados univariados e as respectivas extensões para STI. Mostramos a fundamentação teórica das técnicas adotadas. Ainda neste capítulo, apresentamos uma rede neural artificial para ajuste de STI e um método híbrido que usa conjuntamente modelos tradicionais e rede neural para previsão destas séries. A metodologia híbrida é inspirada no modelo de Zhang (2003) [94].

Os experimentos para validar as técnicas e modelos para STI obtidos neste trabalho são apresentados no Capítulo 5. Eles foram conduzidos através de simulação Monte Carlo com STI sintéticas e também em aplicações em STI reais. Os dados reais são compostos de três STI de preços diários de ações de três companhias de diferentes segmentos do mercado e de duas STI de temperaturas registradas durante o dia nas cidades de Nova Iorque e do Rio de Janeiro.

Finalmente, no Capítulo 6 apresentamos as conclusões, incluindo sugestões para trabalhos futuros a serem desenvolvidos.

## Capítulo 2

# Análise de dados simbólicos

Os métodos tradicionais de análise de dados foram desenvolvidos numa época em que a quantidade de informação disponível era infinitamente menor que a disponível atualmente. Além disso, apesar do alto poder de processamento de computadores atuais, o esforço computacional necessário para manipulação de grandes massas de dados ainda é um problema. Os dados simbólicos apresentam, em sua estrutura, formas interessantes para se transformar grandes bases de dados clássicos em novos conjuntos de dados de tamanho reduzido, facilitando a manipulação e proporcionando novas técnicas de análise dos mesmos. Uma base de dados simbólicos pode conter, em uma célula de sua matriz, informações expressas em forma de listas, de intervalos, de distribuições de frequência, de distribuições de probabilidade, diferentemente de uma base de dados clássica em que cada célula assume um valor único (células univaloradas). No entanto, mesmo com os recentes avanços promovidos por pesquisadores nesta área, o volume de técnicas de manipulação e, conseqüentemente, de análise de dados simbólicos ainda é incipiente.

Atualmente, as bases de dados contendo informações, tais como, dados pessoais de clientes, transações efetuadas, etc., são muito importantes para tomadas de decisão em empresas de qualquer ramo de atuação. Frequentemente estas bases são bastante grandes. Em geral, as bases contêm informações de indivíduos descritas, tipicamente, por variáveis categóricas ou numéricas e, obviamente, as variáveis numéricas podem posteriormente ser transformadas em categorias. Em *Análise de Dados Simbólicos* estas categorias são consideradas novas unidades estatísticas e o primeiro passo é descrever estas unidades levando-se em consideração a sua variação

interna. No entanto, o que se entende por “*variação interna*”? Por exemplo, a idade de um jogador de uma equipe de futebol é 32 anos, mas as idades dos jogadores desta equipe (considerada como uma categoria) variam entre 22 e 34 anos.

Um caso mais geral de dados simbólicos é considerar um processo de agrupamento aplicado em uma ampla base de dados com o objetivo de resumi-la. Cada grupo formado por este processo pode ser considerado como uma categoria, e, conseqüentemente, cada variável terá uma representação simbólica associada a uma variação interna em cada categoria. Como exemplo prático, podemos citar as bases de créditos financeiros. Devido ao aumento dos tipos de cartões de crédito no mercado e, conseqüentemente, ao elevado número de cartões disponíveis para usuários das mais diversas classes socioeconômicas, a base de dados com informações de transações realizadas por estes usuários é bastante grande. Em um caso como este, é fácil perceber que armazenar, por exemplo, o intervalo dos valores gastos por cada usuário em um determinado mês, e.g., de R\$ 20,00 a R\$ 300,00, é menos custoso que armazenar todos os valores gastos por este usuário neste mês. Assim como armazenar os nomes de todos os estabelecimentos onde o usuário realizou uma compra é mais custoso que armazenar as proporções de compras realizadas em cada ramo de estabelecimento, e.g., 15% em alimentação, 20% em vestuário, etc. Estes tipos de informações agrupadas já são utilizados por operadoras de cartões de créditos para traçar perfis de usuários e algumas até disponibilizam estas informações aos usuários.

Segundo Diday e Esposito (2003) [33], o objetivo geral da ADS é generalizar técnicas de mineração de dados e técnicas estatísticas aplicadas a dados padrão (clássicos) para o tratamentos de dados simbólicos. Com a ADS é possível estender os métodos tradicionais de análise exploratória de dados às variáveis simbólicas que, como será discutido mais adiante, podem assumir valores como subconjuntos de categorias, intervalos ou distribuições de frequências, contrastando com a análise clássica, onde as variáveis podem assumir apenas um único número, categoria ou nível.

A principal vantagem da ADS em face à análise de dados clássicos reside no fato de ser possível analisar as bases de dados complexas existentes atualmente, que raras vezes preenchem os requisitos necessários à análise clássica. Além disso, a agregação dos dados iniciais sob a forma de dados simbólicos permite, dentre outros

aspectos, preservar a confidencialidade dos microdados.

## 2.1 Dados simbólicos

Os dados simbólicos surgem com a necessidade de analisar informações presentes em bases de dados de estrutura complexa, consonantes com as características das grandes bases de dados existentes atualmente. Os dados simbólicos são apresentados em tabelas, nas quais em cada célula não existe apenas um valor, mas vários, os quais podem ser, por exemplo, ponderados, apresentar dependência hierárquica, ou ainda, estruturados com base em taxonomias entre categorias. Cada célula pode ainda conter um intervalo ou uma distribuição de frequências, como casos particulares das possibilidades descritas.

Nesta seção, nós apresentamos a transição de uma tabela de dados padrão (dados clássicos) para uma tabela de dados simbólicos. Esta situação é ilustrada por um exemplo mostrando que um indivíduo simples pode ser definido a partir de variáveis numéricas padrão ou categóricas, porém, grupo de indivíduos necessita de descrições simbólicas, a fim de conservar a sua variação interna. Apresentamos as definições e propriedades das unidades básicas, como indivíduos e categorias. Um processo de generalização que produz uma descrição simbólica é apresentado. Em seguida, são apresentados brevemente alguns tipos de variáveis consideradas na ADS. São fornecidas algumas aplicações gerais no paradigma de ADS.

Extrair conhecimento de grandes bases de dados é fundamental para a tomada de decisão. Na prática, o objetivo é, muitas vezes, extrair novos conhecimentos a partir de uma base de dados usando uma tabela de dados padrão, em que as entradas são um conjunto de unidades descritas por um conjunto finito de variáveis quantitativas ou categóricas. Em Diday e Noirhomme-Fraiture (2008) [34] é citado que muita informação pode ser perdida quando as unidades são apresentadas através de tabelas de dados clássicos e mostra uma nova estrutura de dados e novas ferramentas para a extração de conhecimentos úteis obtidas a partir destas bases de dados padrão.

Consideremos o seguinte exemplo extraído de Diday e Noirhomme-Fraiture (2008) [34]: Suponha que temos uma base de dados padrão com informações referentes à espécie, capacidade de vôo e tamanho de 600 aves observadas em uma ilha. Observe as características destas aves descritas através de três variáveis na Tabela 2.1. Agora,

se considerarmos que as novas unidades são as espécies de aves na ilha (que são um exemplo do que são chamados de unidades de “alto nível”), respostas diferentes para as mesmas perguntas são obtidas, uma vez que o número de pássaros voando pode ser diferente do número de espécies voando. Para ilustrar este exemplo simples com alguns dados, a Tabela 2.2 descreve as três espécies de aves observadas na ilha: há 100 avestruzes, 100 pinguins e 400 andorinhas. Assim,  $2/3$  das aves (400 andorinhas) estava voando e  $1/3$  não. No entanto, se considerarmos a unidade de alto nível espécie, as frequências para a variável ‘voando’ são o reverso das anteriores, i.e.,  $1/3$  das espécies (andorinhas) estava voando e  $2/3$  das espécies (avestruzes e pinguins) não. Isto significa que os pontos de vista ‘micro’ (as aves) e ‘macro’ (as espécies) podem fornecer resultados totalmente distintos, como a frequência de pássaros voando na ilha que é  $2/3$ , mas a frequência das espécies que voam é  $1/3$ .

Tabela 2.1: Descrição de 600 aves através de três variáveis.

Ave	Espécie	Voando	Tamanho
001	Pinguim	Não	80
$\vdots$	$\vdots$	$\vdots$	$\vdots$
599	Andorinha	Sim	70
600	Avestruz	Não	125

Tabela 2.2: Descrição de três espécies de aves através de três variáveis.

Espécie	Voando	Tamanho	Migração
Pinguim	{Não}	[70; 95]	[100% sim, 0% não]
Andorinha	{Sim}	[60; 85]	[90% sim, 10% não]
Avestruz	{Não}	[85; 160]	[0% sim, 100% não]

Observe na Tabela 2.2 que os valores fornecidos pela variável “tamanho” não são simples números (escalares), mas intervalos. As variáveis tipo intervalo configuram-se em uma forma de “dados simbólicos” envolvendo um tipo diferente de variável a partir de dados padrão. Outros tipos de variáveis simbólicas podem ser adotados com a finalidade de caracterizar as unidades de segundo nível, tais como, a variável ‘migração’ que simboliza que 90% das andorinhas migraram, todos os pinguins migraram e nenhum avestruz migrou. Esta última variável é chamada em ADS de variável simbólica do tipo distribuição de frequências.

Consideremos agora outro exemplo em que temos um conjunto de dados consistindo de registros médicos em um país. Suponha que para cada indivíduo, serão

registradas variáveis referentes à localização geográfica, tais como região (Norte, Nordeste, Sul, ...), cidade (Recife, Salvador, ...), zona rural (sim, não), variáveis demográficas, tais como, gênero, estados civil, idade, número de filhos, plano de saúde, etc., variáveis relacionadas com a saúde, como peso, altura, tipo sanguíneo e outras variáveis são registradas. Note que, em cada variável citada, apenas um valor ou categoria pode ser registrado para cada indivíduo. Um exemplo de um conjunto de tais dados (clássicos) pode ser visto na Tabela 2.3.

Tabela 2.3: Exemplo de uma base de dados clássicos.

Indivíduo	Região	Cidade	Rural	Sexo	Peso	...
1	Sul	Porto Alegre	Não	M	76,2	...
2	Nordeste	Salvador	Não	F	45,8	...
3	Norte	Manaus	Sim	M	69,1	...
4	Sul	Curitiba	Não	F	54,0	...
5	Sudoeste	Niterói	Não	M	60,9	...
6	Norte	Belém	Sim	M	71,0	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Seja  $p$  o número de variáveis registradas para o indivíduo  $i \in \Omega = \{1, \dots, n\}$ , onde obviamente  $p$  e  $n$  podem ser extremamente grande. Seja  $Y_j$ ,  $j = 1, \dots, p$ , representando a  $j$ -ésima variável e  $y_{ij}$  um valor particular assumido para a variável  $Y_j$  para o  $i$ -ésimo indivíduo em um conjunto de dados clássicos. Agora, considere  $Y = (y_{ij})$  como a matriz  $(n \times p)$  de entradas do conjunto de dados. Seja  $\mathcal{Y}_j$  o domínio de  $Y_j$ .

Variáveis podem ser quantitativas, por exemplo, idade  $\mathcal{Y}_{idade} = \{x \geq 0\}$  como uma variável aleatória contínua ou anos de estudo  $\mathcal{Y}_{escolaridade} = \{0, 1, 2, \dots\}$  como uma variável aleatória discreta. Podem ser categóricas, por exemplo,  $\mathcal{Y}_{cidade} = \{\text{Recife, Salvador, ...}\}$ . Variáveis relacionadas com doenças podem ser registradas como categorias de uma variável com domínio  $\mathcal{Y} = \{\text{cardíaca, câncer, ...}\}$ , ou, como é mais provável, como uma variável indicadora, por exemplo,  $Y = \text{câncer}$  com domínio  $\mathcal{Y} = \{\text{Sim, Não}\}$  ou  $\{0, 1\}$  ou com outros códigos indicando os níveis de estágios da doença. Deste modo, para o registro dos muito possíveis tipos de câncer, cada tipo pode ser representado por uma variável  $Y$ , ou por uma categoria da variável câncer.

Note que a natureza precisa presente nas descrições das variáveis de uma base de

dados clássicos faz com que cada valor  $y_{ij}$  assuma apenas uma única característica ou valor numérico. Por exemplo, idade de um indivíduo  $Y_{idade} = 21$  ou a cidade onde reside  $Y_{cidade} = \text{Recife}$ . Assim, um indivíduo representado por dados clássicos é um simples ponto no espaço  $p$ -dimensional. Em contraste, em uma base de dados simbólicos, cada observação, denotada por  $\xi_{ij}$ , não está restrita a um simples valor. A variável ‘peso’, por exemplo, pode ser registrada de forma intervalar ( $[40, 1; 79, 0]$ ,  $[50, 2; 100, 7]$ , ...) representando os pesos de integrantes de uma família ou grupo de indivíduos, ou até de um indivíduo que no momento não sabe seu peso exato e o relata através de um intervalo. A variável pressão sanguínea pode ser registrada por valores, por exemplo,  $\xi_{ij} = [8; 12]$ .

Tabela 2.4: Exemplo de uma base de dados simbólicos.

Unidade	Diabetes	Automóvel	Peso	...
1	[3/4 diabetes; 1/4 não diabetes]	{ <i>Ford, Chevrolet</i> }	[76,2; 98,0]	...
2	[2/5 diabetes, 3/5 não diabetes]	{ <i>Honda, Ford</i> }	[35,8; 91,2]	...
3	[1/5 diabetes, 4/5 não diabetes]	{ <i>Toyota, Fiat, Volvo</i> }	[49,1; 85,2]	...
4	[1/8 diabetes, 7/8 não diabetes]	{ <i>Chevrolet</i> }	[54,0; 100,2]	...
5	[2/5 diabetes, 3/5 não diabetes]	{ <i>Ford, Fiat</i> }	[40,9; 85,9]	...
6	[1/7 diabetes, 6/7 não diabetes]	{ <i>Volvo</i> }	[61,0; 115,1]	...
⋮	⋮	⋮	⋮	⋮

Suponha agora outra variável representando a marca do automóvel que um indivíduo possui. Para certa revendedora, poderíamos obter  $\xi_{ij} = \{\text{Chevrolet, Fiat, Ford}\}$ . Esse tipo de variável é chamado de simbólica multivalorada. Um terceiro tipo de variável simbólica é representado por variáveis multi-estado com frequências, probabilidades, ou pesos associados aos seus possíveis valores. Por exemplo, se em um grupo de pessoas, trinta possuem diabetes e apenas dez não possuem, então a variável representando a propensão a diabetes desse grupo seria expressa por  $\xi_{ij} = \{3/4 \text{ diabetes, } 1/4 \text{ não-diabetes}\}$ . De maneira mais geral,  $\xi_{ij}$  pode assumir, por exemplo, um histograma, uma distribuição de frequência ou uma distribuição de probabilidade. Na Tabela 2.4 apresentamos um exemplo de uma base de dados simbólicos. A quarta coluna (Peso) desta tabela apresenta um exemplo de dados do tipo intervalo.

Um processo de generalização é aplicado a um conjunto de indivíduos, com a

finalidade de produzir uma ‘descrição simbólica’. Por exemplo, o conceito de ‘andorinha’ é descrito na Tabela 2.2 pelo vetor descrição  $d = (\{\text{Sim}\}, [60;85], [90\% \text{ sim}, 10\% \text{ não}])$ . O processo de generalização deve levar em consideração a variação interna da descrição dos indivíduos pertencentes ao conjunto de indivíduos que se descreve. Por exemplo, as 400 andorinhas na ilha têm tamanhos que variam entre 60 e 85. Outra variável poderia também ser considerada: a variável ‘cor’. Portanto, a cor dos avestruzes pode ser branca ou preta (que exprime uma variação da cor dos avestruzes nesta ilha entre branca e preta), a cor dos pinguins é sempre preta e branca (que não expressam qualquer variação, mas um conjunto válido para todas as aves desta espécie), e a cor das andorinhas é preta e branca ou cinzenta. Esta variação leva a um novo tipo de variável definida sobre o conjunto de conceitos, pois o valor dessa variável para um conceito pode ser um único valor numérico ou categórico, mas também um intervalo, um conjunto de valores ordinais ou categóricos que podem ou não ser ponderados, uma distribuição, etc. Uma vez que esses valores não são meramente numéricos ou categóricos, eles foram denominados *dados simbólicos*. As variáveis associadas a estes dados são chamadas *variáveis simbólicas*.

O significado de dados simbólicos, tais como intervalos, é muito importante para determinar a forma de estender técnicas estatísticas, análise de dados ou mineração de dados a esses dados. Por exemplo, se dispomos de um equipamento capaz de medir a altura exata de um determinado jogador de futebol, podemos ter certeza da altura do jogador e, então, podemos afirmar que é 182 centímetros. Mas, se não tivermos certeza, podemos dizer com alguma incerteza que sua altura pertence ao intervalo  $\mathbf{I}_1 = [180; 185]$ . É por esta razão que podemos dizer que  $\mathbf{I}_1$  é uma ‘altura com incerteza’. Por sua vez, se considerarmos a variável aleatória  $X$  associada à altura de membros da equipe de futebol, podemos associar com  $X$  vários tipos de dados simbólicos, como a distribuição das alturas, o intervalo de alturas ou uma amostra delas. Se representamos esta variável pelo intervalo de alturas de jogadores da equipe  $\mathbf{I}_2 = [180; 185]$ , podemos notar que  $\mathbf{I}_1 = \mathbf{I}_2$ , mas seus significados são completamente distintos. Isso vem do fato de que o intervalo  $\mathbf{I}_1$  exprime a ‘incerteza’ dada pela nossa própria expectativa e subjetividade, enquanto  $\mathbf{I}_2$  exprime a ‘variabilidade’ da altura dos jogadores na equipe. Ao considerar a altura com incerteza de todos os jogadores da equipe, obtemos um cenário de dados onde os indivíduos são os jogadores e as variáveis associadas são um intervalo (a altura com

incerteza) para cada jogador.

Ao considerar a variável aleatória definida pela altura dos jogadores de uma equipe em uma determinada competição, temos um quadro de dados aleatórios em que os indivíduos (de nível superior) são as equipes e a variável ‘altura’ associa uma variável aleatória a cada equipe. Podemos então criar uma tabela de dados simbólicos que contém, em cada célula associada às alturas de uma determinada equipe, uma distribuição de densidade (ou um histograma), induzida pela variável aleatória definida pela altura dos jogadores desta equipe. Qualquer uma dessas distribuições de densidade associada com uma equipe expressa a variabilidade das alturas dos jogadores desta equipe. Na análise dos dados simbólicos existe o interesse particular no estudo da variabilidade dessas distribuições de densidade (ver, e.g., Diday, 1995) [31].

## 2.2 Séries temporais do tipo intervalo

Estatísticas usuais e análise de dados clássicos tratam, em geral, indivíduos descritos por variáveis nas formas padrão (numéricas, categóricas) que assumem um único valor para um dado indivíduo. Como apresentado antes, variáveis simbólicas generalizam este paradigma clássico proporcionando novas formas de descrever indivíduos através de dados simbólicos, por exemplo, através de intervalos ou histogramas.

Muitas vezes os dados numéricos manipulados em estatística e mineração de dados têm origem em dados intervalares. De fato, toda medida numérica pode ser expressa como um intervalo  $[x \pm \delta]$ , em que,  $x$  é o resultado da medida e  $\delta$  é a imprecisão do instrumento de medição. No contexto de ADS, uma variável tipo intervalo  $X$  (Bock e Diday, 2000) [11] é uma correspondência definida a partir de  $\Omega$  (conjunto de indivíduos) em  $\mathbb{R}$ , tal que, para cada  $k \in \Omega$ ,  $X(k) = [a; b] \in \mathfrak{S}$ , no qual  $\mathfrak{S} = \{[a; b] \mid a, b \in \mathbb{R}, a \leq b\}$  é o conjunto de intervalos fechados definido a partir de  $\mathbb{R}$ . Bock e Diday (2000) [11] apresentaram os dois exemplos a seguir de variável quantitativa intervalar.

**Exemplo 2.1** (Variável quantitativa intervalar). Seja  $E$  o conjunto de estudantes participantes de um curso e seja  $X$  a variável que representa o tempo (em horas) de

estudo do estudante por dia e  $X(i)$  o tempo de estudo do  $i$ -ésimo estudante. Algumas respostas possíveis são  $X(1) = [0; 2]$ ,  $X(2) = [1, 5; 4, 6]$ , ...

**Exemplo 2.2** (Variável quantitativa intervalar obtida por agregação). Considere que  $E = \{C_1, \dots, C_m\}$  denota o conjunto de  $m$  classes de uma escola e que  $Z(i)$  é a altura do  $i$ -ésimo estudante (em metros). Uma classe  $C_k$  é agora caracterizada pela amplitude das alturas:  $X = [\min; \max]$ . A resposta possível para uma classe ginásial é  $X(C_k) = [1, 50; 1, 74]$ .

Quando dados simbólicos do tipo intervalo são coletados em uma sequência ordenada no tempo, dizemos que temos uma *série temporal do tipo intervalo*. Séries deste tipo podem ser observadas em domínios variados. Por exemplo, em economia (rendimento mensal de uma família medido pelo menor e maior salário do domicílio, preço diário de ações de uma companhia que pode ser expressa pelo maior e menor valor de negociação em um dia), engenharia (variação da corrente elétrica num dado ponto expressa pela menor e maior intensidade num dado dia), medicina (pressão arterial diastólica e sistólica), meteorologia (pluviosidade mínima e máxima no mês em um determinado local e umidade relativa do ar também medida pela mínima e máxima registrada em um dado mês).

## 2.3 Alguns métodos estendidos para dados do tipo intervalo

Desde os primeiros textos publicados sobre ADS (Diday, 1987a, 1987b) [29]-[30], vários trabalhos foram publicados, incluindo as coleções de trabalhos sobre ADS editadas por Bock e Diday (2000) [11], por Billard e Diday (2006) [10] e, recentemente, a coleção editada por Diday e Noirhomme-Fraiture (2008) [34]. Diversos trabalhos importantes que oferecem uma síntese sobre o assunto podem ser mencionados, entre eles, Billard e Diday (2003) [9], Diday e Esposito (2003) [33] e Diday (2005) [32]. Várias técnicas estatísticas e de mineração de dados já foram estendidas para dados simbólicos. No trabalho de Billard e Diday (2003) [9], eles forneceram uma revisão dos métodos atualmente disponíveis para analisar esses dados e concluíram que ainda há escassez de metodologias disponíveis para dados simbólicos e,

por isso, chamaram a atenção para uma enorme necessidade de desenvolvimento de novas ferramentas simbólicas.

Para o tratamento de dados do tipo intervalo no contexto da ADS, Billard (2004) [6] apresentou uma medida de correlação para o caso de variáveis simbólicas intervalares. Gioia e Lauro (2006a) [42] introduziram métodos para analisar a dependência entre variáveis simbólicas do tipo intervalo. Em trabalho recente, Guo et al. (2009) [46] definiram dados simbólicos do tipo intervalo não-uniforme e forneceram um conjunto de estatísticas descritivas univariadas e bivariadas utilizando este novo conceito. Com base em estudo empírico sobre a distribuição não-uniforme de dados intervalares, eles obtiveram formas para o cálculo da média e variância amostral. Além disso, fórmulas para a covariância e para o coeficiente de correlação entre duas variáveis intervalares não-uniforme foram fornecidas com base na função de distribuição conjunta empírica.

No domínio de análise multivariada de dados, alguns procedimentos já foram apresentados para tratamento de dados simbólicos intervalares. Cazes et al. (1997) [19] estenderam a análise de componentes principais (ACP) para objetos caracterizados por variáveis multivaloradas descrevendo a variação ou imprecisão nos dados. Chouakria (1998) [21] estendeu dois métodos de análise fatorial para dados de natureza intervalar. Além da ACP, ele estendeu a análise de correspondências múltiplas para tratamento de dados simbólicos do tipo intervalo. Irpino (2006) [54] introduziu uma abordagem para a ACP em que um tipo especial de intervalo é descrito em função do tempo. Na sua abordagem, cada observação é caracterizada por um intervalo de valores com um valor inicial e um valor final em cada período de observação como, por exemplo, o preço de negociação de abertura e de fechamento de uma ação no mercado financeiro em um determinado dia. Gioia e Lauro (2006b) [43] apresentaram uma extensão da ACP com o objetivo obter uma visualização de dados intervalares representados por retângulos, em um espaço de menor dimensão. Eles utilizaram instrumentos da álgebra de intervalos para adaptar o método clássico para a situação em que uma matriz de dados do tipo intervalo é considerada.

Técnicas de agrupamento (*clustering*) para dados simbólicos intervalares começaram a ser exploradas mais recentemente. Brito (2002) [13] apresentou um método para agrupamento hierárquica e piramidal de um conjunto de dados simbólicos, em

que os indivíduos são descritos por variáveis simbólicas de vários tipos, incluindo variáveis intervalares, levando em conta a variabilidade ou incerteza presente nos dados. De Souza e De Carvalho (2004) [28] introduziram métodos de agrupamento para dados simbólicos intervalares que representam uma extensão do algoritmo de nuvens dinâmicas usado em dados usuais. De Carvalho et al. (2006a) [27] propuseram um método dinâmico de partição para dados intervalares baseado na distância de *Hausdorff*, em que cada padrão é representado por um vetor de intervalos. Eles concluíram que o algoritmo de agrupamento utilizando a distância adaptativa apresentada é capaz de reconhecer grupos de diferentes formas e tamanhos. De Carvalho et al. (2006b) [25] consideraram um método de agrupamento para dados simbólicos intervalares através da abordagem dinâmica usando a medida de distância  $L_2$ . De Carvalho e Lechevallier (2009a) [?] apresentaram novos métodos dinâmicos para particionamento de dados simbólicos intervalares baseados em distâncias adaptativas *city-block* e de *Hausdorff*.

Billard e Diday (2000) [7] foram os primeiros a propor um modelo de regressão linear para dados simbólicos de natureza intervalar. A abordagem proposta por eles consiste em ajustar o modelo de regressão nos pontos médios dos intervalos e, então, aplicar o modelo estimado sob os limites inferiores e superiores dos intervalos das variáveis independentes. Posteriormente, eles apresentaram outra abordagem que ajusta um modelo de regressão linear no caso de dados simbólicos do tipo histograma (Billard e Diday, 2002) [8]. Lima Neto e De Carvalho (2008) [63] introduziram uma nova abordagem para ajuste de modelos de regressão linear a dados simbólicos do tipo intervalo, em que os modelos são ajustados ao ponto médio e à amplitude de variáveis intervalares. Lima Neto e De Carvalho (2010) [64] apresentam outra abordagem para ajustar modelo de regressão linear ao ponto médio e à amplitude de variáveis intervalares, usando restrições no vetor de parâmetros com a finalidade de assegurar coerência matemática aos valores estimados dos limites dos intervalos.

No domínio de séries temporais, Arroyo e Maté (2006) [2] fornecem medidas de precisão para STI baseadas em distâncias para dados intervalares, tais como, as distâncias de *Ichino-Yaguchi* e de *Hausdorff*. Arroyo et al. (2007) [4] apresentaram em uma conferência métodos de alisamento exponencial adaptados para STI. Os autores apresentam as computações utilizadas baseadas em aritmética de intervalos (Moore, 1966) [73] e afirmam que este é um campo separado da ADS. Com esta

aritmética, eles definem o intervalo médio para o conjunto de intervalos, em que cada intervalo é definido através de duas coordenadas: o limite superior e o limite inferior, ou, equivalentemente, o centro e a amplitude. Os mesmos autores de Arroyo et al. (2007) [4] apresentaram em Roque et al. (2007) [79] um novo tipo de MLP baseado em aritmética de intervalos, em que as entradas e as saídas da rede neural são dadas por intervalos, porém, os pesos são univalorados. Eles denominaram a rede proposta por iMLP.

Em trabalho recente, García-Ascanio e Maté (2010) [39] promovem uma comparação entre modelos vetores autorregressivos (VAR) e redes neurais iMLP na previsão de demanda de energia elétrica. Eles confrontaram dois modelos VAR, um considerando as séries dos centros e das amplitudes dos intervalos e outro considerando as séries dos limites superiores e inferiores, com a rede iMLP ajustada às séries dos centros e das amplitudes. Os seus resultados revelaram que o modelo VAR fornece melhores resultados que a rede iMLP na previsão de demanda de energia elétrica.

Em outro trabalho recente, Arroyo e Maté (2009) [3] formularam novos modelos para previsão de séries temporais simbólicas com dados do tipo histograma sob o argumento de que existem situações em que os histogramas são mais informativos que um valor agregado, como a média. Eles desenvolveram uma adaptação do algoritmo  $k$ -NN usando as métricas de distância de *Mallows* e de *Wasserstein* para medir as dissimilaridades entre sequências de histogramas e para efetuar previsões de séries do tipo histograma.

## 2.4 Considerações finais

No domínio de modelos de regressão linear, nós também temos uma contribuição em que introduzimos um método robusto que considera como critério de minimização para estimação dos parâmetros do modelo a soma dos desvios absolutos (Maia e De Carvalho, 2008) [69]. O modelo de regressão apresentado é uma extensão do modelo de regressão da mediana para dados simbólicos intervalares e será apresentado em detalhes no Capítulo 3.

Os modelos desenvolvidos pelos autores citados na seção anterior com a finalidade de prever STI são baseados em aritmética de intervalos (Moore, 1966) [73],

domínio não explorado nesta tese.

Recentemente, nós apresentamos alguns métodos direcionados à extensão de técnicas de análise de séries temporais para o caso particular de dados simbólicos do tipo intervalo (Maia et al., 2006, 2008, 2008b) [66]-[67]-[70]. Unindo a área de modelagem e previsão de STI com a área da ADS, apresentamos duas abordagens para previsão de séries temporais considerando variáveis simbólicas do tipo intervalo. Os dois métodos que apresentamos consideram dois modelos independentes ajustados sobre os pontos médios e sobre as amplitudes dos intervalos. O primeiro método apresentado baseia-se no modelo auto-regressivo integrado média-móvel (ARIMA) e o segundo, em um modelo híbrido usando ARIMA e redes neurais artificiais. Estes métodos também serão descritos em detalhes no Capítulo 3.

As técnicas e os modelos estendidos para STI abrem um novo campo de pesquisa no contexto da ADS.

# Parte I

## Métodos de ajuste independiente

## Capítulo 3

# Métodos de ajuste independente

Neste capítulo, nós apresentamos os modelos desenvolvidos com a finalidade de ajustar STI do ponto de vista da ADS que fazem parte dos artigos Maia et al. (2006, 2008) [66]-[67] e estão inseridos no que denominamos de classe de *métodos de ajuste independente*. A razão para este título vem do fato de modelarmos duas séries – a série dos centros dos intervalos e a série das amplitudes dos intervalos – através de dois modelos independentes. Os detalhes serão dados adiante.

Uma discussão inicial sobre os modelos de regressão para dados de natureza intervalar faz-se necessária para caracterizar as origens da ideia adotada nos modelos apresentados em Maia et al. (2006, 2008) [66]-[67]. Assim sendo, inicialmente apresentamos brevemente os modelos de regressão propostos por Billard e Diday (2000) [7], em seguida, apresentamos os modelos propostos por Lima Neto e De Carvalho (2008) [63] e Maia e De Carvalho (2008) [69] e, por fim, os modelos de Lima Neto e De Carvalho (2010) [64]. O modelo de Billard e Diday (2000) [7] é baseado, unicamente, na modelagem dos pontos médios dos intervalos e os modelos desenvolvidos pelos demais autores são baseados na modelagem dos centros e das amplitudes dos intervalos e formam a base para a metodologia adotada no desenvolvimento dos modelos de Maia et al. (2006, 2008) [66]-[67] para ajuste de STI.

### 3.1 Modelos de regressão para dados intervalares

Os primeiros modelos de regressão propostos para modelagem de dados simbólicos com variáveis de natureza intervalar foram apresentados por Billard e Diday (2000)

[7]. O método desenvolvido pelos autores consiste em ajustar um único modelo de regressão linear sob os pontos médios (centros) dos intervalos e, então, aplicar esse modelo sob os valores dos limites inferiores e dos limites superiores dos intervalos das variáveis independentes para prever o limite inferior e o limite superior da variável dependente. Além disso, o método de Billard e Diday (2000) [7] consiste em minimizar a soma dos quadrados dos erros dos pontos médios dos intervalos.

Para definir o modelo de Billard e Diday (2000) [7], seja  $E = \{e_1, \dots, e_n\}$  um conjunto de exemplos descritos por  $k + 1$  variáveis simbólicas do tipo intervalo,  $X_1, \dots, X_k$  e  $Y$ . Cada exemplo  $e_i \in E$  ( $i = 1, \dots, n$ ) é representado por um vetor intervalar  $\mathbf{z}_i = (\mathbf{x}_i, y_i)$ ,  $\mathbf{x}_i = (x_{i1}, \dots, x_{ik})$ , em que  $x_{ij} = [a_{ij}, b_{ij}] \in \mathfrak{S} = \{[a, b] : a, b \in \mathbb{R}, a \leq b\}$  ( $j = 1, \dots, k$ ) e  $y_i = [y_i^L, y_i^U] \in \mathfrak{S}$  são os valores observados de  $X_j$  e  $Y$ , respectivamente. Este modelo considera a relação entre a variável intervalar dependente  $Y$  e as variáveis intervalares independentes  $X_1, \dots, X_k$  através da forma linear

$$\begin{aligned} y_i^L &= \beta_0 + \beta_1 a_{i1} + \dots + \beta_k a_{ik} + \epsilon_i^L, \\ y_i^U &= \beta_0 + \beta_1 b_{i1} + \dots + \beta_k b_{ik} + \epsilon_i^U, \end{aligned}$$

e, conseqüentemente, define a função objetivo como a soma de quadrados dos erros,

$$\sum_{i=1}^n (y_i^L - \beta_0 - \beta_1 a_{i1} - \dots - \beta_k a_{ik} + y_i^U - \beta_0 - \beta_1 b_{i1} - \dots - \beta_k b_{ik})^2, \quad (3.1)$$

que representa a soma dos quadrados dos erros dos limites inferior e superior.

Neste modelo de Billard e Diday (2000) [7], a estimativa de mínimos quadrados dos parâmetros  $\beta_0, \dots, \beta_k$ , que minimiza a expressão (3.1), pode ser obtida da forma

$$\hat{\boldsymbol{\beta}} = ((\mathbf{X}^c)^\top \mathbf{X}^c)^{-1} (\mathbf{X}^c)^\top \mathbf{y}^c,$$

em que  $\mathbf{X}^c = ((\mathbf{x}_1^c)^\top, \dots, (\mathbf{x}_n^c)^\top)^\top$  é uma matriz de posto completo  $k + 1 \leq n$ , com  $(\mathbf{x}_i^c)^\top = (1, x_{i1}^c, \dots, x_{ik}^c)$  e  $\mathbf{y}^c = (\mathbf{y}_1^c, \dots, \mathbf{y}_n^c)^\top$  e, por sua vez,  $x_{ij}^c = (a_{ij} + b_{ij})/2$  e  $y_i^c = (y_i^L + y_i^U)/2$  são, respectivamente, os pontos médios dos intervalos  $x_{ij}$  e  $y_i$ .

Dado um novo exemplo  $e$  descrito pelo vetor  $\mathbf{z} = (\mathbf{x}, y)$ , em que  $\mathbf{x} = (x_1, \dots, x_k)$  com  $x_j = [a_j, b_j]$  ( $j = 1, \dots, k$ ), então, o valor  $y = [y^L, y^U]$  de  $Y$  será predito por  $\hat{y} = [\hat{y}^L, \hat{y}^U]$  da seguinte forma:

$$\hat{y}^L = (\mathbf{x}^L)^\top \hat{\boldsymbol{\beta}} \quad \text{e} \quad \hat{y}^U = (\mathbf{x}^U)^\top \hat{\boldsymbol{\beta}}, \quad (3.2)$$

em que  $(\mathbf{x}^L)^\top = (1, a_1, \dots, a_k)$  e  $(\mathbf{x}^U)^\top = (1, b_1, \dots, b_k)$ .

Em suma, o método proposto por Billard e Diday (2000) [7] consiste em ajustar um modelo de regressão linear aos pontos médios da variável dependente do tipo intervalo  $Y$  sobre os pontos médios das variáveis independentes do tipo intervalo  $X_j$  e, então, utilizar os parâmetros estimados para obter predições do limite superior e do limite inferior do intervalo, como na Equação (3.2).

Em seguida, Lima Neto e De Carvalho (2008) [63], Maia e De Carvalho (2008) [69] e, posteriormente, Lima Neto e De Carvalho (2010) [64] propuseram novos modelos de regressão linear para dados intervalares que levam em consideração as informações contidas nos centros e nas amplitudes dos intervalos, diferentemente do método de Billard e Diday (2000) [7] que leva em consideração apenas as informações contida nos centros dos intervalos. As abordagens diferem do método de Billard e Diday (2000) [7] também em outros aspectos. Primeiro, com relação aos critérios de minimização do erro adotados. Lima Neto e De Carvalho (2008) [63] mostram que as estimativas dos parâmetros encontradas mediante outro critério de minimização resultam em melhor desempenho de predição. Maia e De Carvalho (2008) [69] estenderam a regressão da mediana condicional empregando como critério a soma dos desvios absolutos e mostraram a utilidade do modelo para o caso de dados intervalares na presença de valores atípicos. Por sua vez, Lima Neto e De Carvalho (2010) [64] incorporaram restrições aos parâmetros do modelo visando garantir coerência matemática aos valores preditos dos limites dos intervalos. Eles consideraram que ao prever um intervalo é necessário que a estimativa do limite inferior seja menor ou igual à estimativa do limite superior e apresentaram um método com a finalidade de garantir esta coerência matemática nas predições.

Para definir o modelo apresentado por Lima Neto e De Carvalho (2008) [63], consideremos  $E = \{e_1, \dots, e_n\}$  um conjunto de exemplos descritos por  $k+1$  variáveis simbólicas do tipo intervalo,  $X_1, \dots, X_k$  e  $Y$ . Sejam  $Y^c$  e  $X_j^c$  ( $j = 1, 2, \dots, k$ ) variáveis quantitativas que representam os valores dos centros (pontos médios) dos intervalos das variáveis simbólicas  $Y$  e  $X_j$  ( $j = 1, 2, \dots, k$ ), respectivamente. Além disso, sejam  $Y^r$  e  $X_j^r$  ( $j = 1, 2, \dots, k$ ) variáveis quantitativas que representam os valores das metades das amplitudes (meia-amplitude ou meio-*range*) dos intervalos das variáveis simbólicas  $Y$  e  $X_j$  ( $j = 1, 2, \dots, k$ ), respectivamente.

De acordo com os elementos descritos acima, Lima Neto e De Carvalho (2008)

[63] representam cada exemplo  $e_i \in E$  ( $i = 1, \dots, n$ ) através de dois vetores,  $\mathbf{w}_i = (\mathbf{x}_i^c, y_i^c)$  e  $\mathbf{r}_i = (\mathbf{x}_i^r, y_i^r)$ , com  $\mathbf{x}_i^c = (x_{i1}^c, \dots, x_{ik}^c)$  e  $\mathbf{x}_i^r = (x_{i1}^r, \dots, x_{ik}^r)$ , em que  $x_{ij}^c = (a_{ij} + b_{ij})/2$ ,  $x_{ij}^r = (b_{ij} - a_{ij})/2$ ,  $y_i^c = (y_i^L + y_i^U)/2$  e  $y_i^r = (y_i^U - y_i^L)/2$  são, respectivamente, os valores observados de  $X_j^c$ ,  $X_j^r$ ,  $Y^c$  e  $Y^r$ .

Desta forma, considerando  $Y^c$  e  $Y^r$  as variáveis dependentes e  $X_j^c$  e  $X_j^r$  ( $j = 1, 2, \dots, p$ ) um conjunto de variáveis independentes, a relação adotada é de acordo com a seguinte estrutura de regressão linear:

$$\begin{aligned} y_i^c &= \beta_0^c + \beta_1^c x_{i1}^c + \dots + \beta_p^c x_{ik}^c + \epsilon_i^c, \\ y_i^r &= \beta_0^r + \beta_1^r x_{i1}^r + \dots + \beta_p^r x_{ik}^r + \epsilon_i^r. \end{aligned}$$

Então, Lima Neto e De Carvalho (2008) [63] denotaram a soma de quadrados dos erros da forma

$$\sum_{i=1}^n (y_i^c - \beta_0^c - \beta_1^c x_{i1}^c - \dots - \beta_k^c x_{ik}^c)^2 + \sum_{i=1}^n (y_i^r - \beta_0^r - \beta_1^r x_{i1}^r - \dots - \beta_k^r x_{ik}^r)^2,$$

que representa a soma dos quadrados dos erros dos centros mais a soma dos quadrados dos erros do meio-range. Observe que a metodologia empregada neste modelo de regressão considera vetores independente de parâmetros para predizer o centro e o meio-range dos intervalos.

Sejam as matrizes  $\mathbf{X}^c = ((\mathbf{x}_1^c)^\top, \dots, (\mathbf{x}_n^c)^\top)^\top$  e  $\mathbf{X}^r = ((\mathbf{x}_1^r)^\top, \dots, (\mathbf{x}_n^r)^\top)^\top$  de posto completo  $k + 1 \leq n$ , com  $(\mathbf{x}_i^c)^\top = (1, x_{i1}^c, \dots, x_{ik}^c)$ ,  $(\mathbf{x}_i^r)^\top = (1, x_{i1}^r, \dots, x_{ik}^r)$  e sejam os vetores  $\mathbf{y}^c = (\mathbf{y}_1^c, \dots, \mathbf{y}_n^c)^\top$  e  $\mathbf{y}^r = (\mathbf{y}_1^r, \dots, \mathbf{y}_n^r)^\top$ , as estimativas de mínimos quadrados dos parâmetros  $\boldsymbol{\beta}^c = (\beta_0^c, \beta_1^c, \dots, \beta_k^c)$  e  $\boldsymbol{\beta}^r = (\beta_0^r, \beta_1^r, \dots, \beta_k^r)$  são obtidas, respectivamente, por

$$\hat{\boldsymbol{\beta}}^c = ((\mathbf{X}^c)^\top \mathbf{X}^c)^{-1} (\mathbf{X}^c)^\top \mathbf{y}^c \quad \text{e} \quad \hat{\boldsymbol{\beta}}^r = ((\mathbf{X}^r)^\top \mathbf{X}^r)^{-1} (\mathbf{X}^r)^\top \mathbf{y}^r.$$

Assim, dado um novo exemplo  $e$  representado pelos vetores  $\mathbf{z} = (\mathbf{x}, y)$ ,  $\mathbf{w} = (\mathbf{x}^c, y^c)$  e  $\mathbf{r} = (\mathbf{x}^r, y^r)$ , em que  $\mathbf{x} = (x_1, \dots, x_k)$ , com  $x_j = [a_j, b_j]$ ,  $\mathbf{x}^c = (1, x_1^c, \dots, x_k^c)$  e  $\mathbf{x}^r = (1, x_1^r, \dots, x_k^r)$ , o valor do intervalo  $y = [y^L, y^U]$  de  $Y$  é predito considerando os valores preditos  $\hat{y}^c$  de  $Y^c$  e  $\hat{y}^r$  de  $Y^r$  através das expressões

$$\hat{y}^L = \hat{y}^c - \hat{y}^r \quad \text{e} \quad \hat{y}^U = \hat{y}^c + \hat{y}^r,$$

em que,  $\hat{y}^c = (\mathbf{x}^c)^\top \hat{\boldsymbol{\beta}}^c$  e  $\hat{y}^r = (\mathbf{x}^r)^\top \hat{\boldsymbol{\beta}}^r$ .

Nos experimentos de Lima Neto e De Carvalho (2008) [63], eles compararam seu modelo com o modelo de Billard e Diday (2000) [7] e demonstraram a importância de levar em consideração informação contida nas amplitudes das variáveis simbólicas do tipo intervalo. Lima Neto e De Carvalho (2008) [63] concluíram que seu modelo apresentou desempenho superior no ajuste de dados intervalares.

Ainda no campo da regressão linear para dados intervalares, nós apresentamos um modelo robusto que considera a soma dos desvios absolutos como critério de minimização para estimação dos parâmetros do modelo (Maia e De Carvalho, 2008) [69]. O nosso modelo é inspirado no modelo de Lima Neto e De Carvalho (2008) [63] e também está inserido na classe de *modelos de ajuste independente*. Trata-se de uma extensão do modelo de regressão da mediana para dados simbólicos intervalares e, então, é uma alternativa ao modelo de Lima Neto e De Carvalho (2008) [63].

Em Maia e De Carvalho (2008) [69], nosso propósito foi obter estimativas de parâmetros do modelo de regressão mais robustas via minimização da *soma dos desvios absolutos*. Este critério implica em dar menores pesos aos valores extremos. De maneira análoga ao modelo proposto por Lima Neto e De Carvalho (2008) [63], nós ajustamos dois modelos de regressão linear, especificamente, aos pontos médios e às amplitudes dos intervalos assumidos pelas variáveis.

De forma idêntica ao modelo anterior, considerando  $Y^c$  e  $Y^r$  as variáveis dependentes e  $X_j^c$  e  $X_j^r$  ( $j = 1, 2, \dots, p$ ) um conjunto de variáveis independentes, nós consideramos a seguinte estrutura de regressão linear:

$$\begin{aligned} y_i^c &= \beta_0^c + \beta_1^c x_{i1}^c + \dots + \beta_p^c x_{ik}^c + \epsilon_i^c, \\ y_i^r &= \beta_0^r + \beta_1^r x_{i1}^r + \dots + \beta_p^r x_{ik}^r + \epsilon_i^r. \end{aligned}$$

No entanto, o critério de minimização adotado é dado pela soma dos desvios absolutos,

$$\sum_{i=1}^n |y_i^c - \beta_0^c - \beta_1^c x_{i1}^c - \dots + \beta_p^c x_{ip}^c| + \sum_{i=1}^n |y_i^r - \beta_0^r - \beta_1^r x_{i1}^r - \dots + \beta_p^r x_{ip}^r|.$$

Como resultado importante, é possível obter estimativas de parâmetros que minimizam a expressão acima usando um *algoritmo simplex* desenvolvido por Barrodale e Roberts (1974) [5] para regressão- $\ell_1$ . Este algoritmo é uma modificação do método simplex de programação linear aplicado à formulação do problema  $\ell_1$ . De acordo

com Buchinsky (1997) [17], a representação da regressão- $\ell_1$  como um problema de programação linear tem algumas implicações importantes, tais como, a garantia de que os coeficientes de regressão são estimados em um número finito de iterações simplex e a robustez do vetor de coeficientes estimados na presença de observações extremas na variável dependente.

No nosso modelo, analogamente ao modelo de Lima Neto e De Carvalho (2008) [63], para um novo exemplo  $e$  representado pelos vetores  $\mathbf{z} = (\mathbf{x}, y)$ ,  $\mathbf{w} = (\mathbf{x}^c, y^c)$  e  $\mathbf{r} = (\mathbf{x}^r, y^r)$ , em que  $\mathbf{x} = (x_1, \dots, x_k)$ , com  $x_j = [a_j, b_j]$ ,  $\mathbf{x}^c = (1, x_1^c, \dots, x_k^c)$  e  $\mathbf{x}^r = (1, x_1^r, \dots, x_k^r)$ , o valor do intervalo  $y = [y^L, y^U]$  de  $Y$  é predito considerando os valores preditos  $\hat{y}^c$  de  $Y^c$  e  $\hat{y}^r$  de  $Y^r$  através das expressões

$$\hat{y}^L = \hat{y}^c - \hat{y}^r \quad \text{e} \quad \hat{y}^U = \hat{y}^c + \hat{y}^r,$$

em que,  $\hat{y}^c = (\mathbf{x}^c)^\top \hat{\boldsymbol{\beta}}^c$  e  $\hat{y}^r = (\mathbf{x}^r)^\top \hat{\boldsymbol{\beta}}^r$ .

Nós concluímos em Maia e De Carvalho (2008) [69] que o modelo da mediana condicional forneceu melhores resultados quando comparado ao modelo de Lima Neto e De Carvalho (2008) [63] na presença de valores extremos na variável dependente.

Por fim, nós descrevemos brevemente o modelo de Lima Neto e De Carvalho (2010) [64]. Eles apresentaram uma nova abordagem para ajustar modelo de regressão linear a dados intervalares usando restrições sob forma de inequações no vetor de parâmetros. Estas restrições têm por finalidade assegurar coerência matemática aos valores estimados dos limites dos intervalos, ou seja, garantir que os valor de  $\hat{y}_i^L$  seja sempre inferior ou igual aos valor de  $\hat{y}_i^U$ , para todo  $i$ . A ideia é considerar restrições de desigualdade apenas sobre os parâmetros  $\hat{\boldsymbol{\beta}}^r$ , responsável por garantir que os valores estimados de  $\hat{y}^r$  sejam sempre maiores ou iguais a zero, o que implica que  $\hat{y}_i^L$  seja sempre inferior ou igual a  $\hat{y}_i^U$ . Os parâmetros  $\hat{\boldsymbol{\beta}}^c$  utilizados para estimar o ponto médio dos intervalos são irrestritos.

Sejam  $Y^c$  e  $Y^r$  as variáveis dependentes e  $X_j^c$  e  $X_j^r$  ( $j = 1, 2, \dots, k$ ) as variáveis independentes, o modelo de regressão linear restrito apresentado por Lima Neto e De Carvalho (2010) [64] é representado de acordo com a seguinte estrutura de regressão linear:

$$\begin{aligned} y_i^c &= \beta_0^c + \beta_1^c x_{i1}^c + \dots + \beta_k^c x_{ik}^c + \epsilon_i^c, \\ y_i^r &= \beta_0^r + \beta_1^r x_{i1}^r + \dots + \beta_k^r x_{ik}^r + \epsilon_i^r, \end{aligned}$$

com as restrições  $\beta_j^r \geq 0$ ,  $j = 0, \dots, k$ .

Nota-se que o modelo ajustado sobre os centros dos intervalos não tem restrições e que o modelo de regressão ajustado sobre as metades das amplitudes dos intervalos está sujeito a restrições de desigualdade com o objetivo de garantir  $\hat{y}_i^r \geq 0$  e, conseqüentemente,  $\hat{y}_i^U \geq \hat{y}_i^L$ .

Desta maneira, o objetivo do método de Lima Neto e De Carvalho (2010) [64] é minimizar a soma dos quadrados dos erros,

$$\sum_{i=1}^n (y_i^c - \beta_0^c - \beta_1^c x_{i1}^c - \dots - \beta_k^c x_{ik}^c)^2 + \sum_{i=1}^n (y_i^r - \beta_0^r - \beta_1^r x_{i1}^r - \dots - \beta_k^r x_{ik}^r)^2,$$

sujeito a  $\beta_j^r \geq 0$ ,  $j = 0, \dots, k$ .

Se  $\mathbf{X}^c$  tem posto completo  $k + 1 \leq n$ , as estimativas de mínimos quadrados do vetor de parâmetros  $\beta^c$  são dadas de forma análoga ao modelo de Lima Neto e De Carvalho (2008) [63], ou seja,

$$\hat{\beta}^c = ((\mathbf{X}^c)^\top \mathbf{X}^c)^{-1} (\mathbf{X}^c)^\top \mathbf{y}^c,$$

visto que não requer restrições nos parâmetros. Entretanto, o modelo sobre as metades das amplitudes dos intervalos deve ser ajustado respeitando as restrições  $\beta_j^r \geq 0$ ,  $j = 0, \dots, k$ . Para resolver este problema, Lima Neto e De Carvalho (2010) [64] consideraram uma solução apresentada por Lawson e Hanson (1974) [61]. Esta solução consiste em um algoritmo que implementa uma solução de Judge e Takayama (1966) [55] através de técnicas iterativas de programação quadrática. A ideia do algoritmo é identificar valores incompatíveis com as restrições e torná-los positivos por meio de processo de re-ponderação. A solução empregada assegura a positividade das estimativas de mínimos quadrados do vetor de parâmetros  $\beta^r$ , se  $\mathbf{X}^r$  tem posto completo  $k + 1 \leq n$ . Lima Neto e De Carvalho (2010) [64] apresentam o algoritmo em detalhes.

Como no modelo anterior, dado um novo exemplo  $e$  representado pelos vetores  $\mathbf{z} = (\mathbf{x}, y)$ ,  $\mathbf{w} = (\mathbf{x}^c, y^c)$  e  $\mathbf{r} = (\mathbf{x}^r, y^r)$ , o valor do intervalo  $y = [y^L, y^U]$  de  $Y$  é predito considerando os valores preditos  $\hat{y}^c$  de  $Y^c$  e  $\hat{y}^r$  de  $Y^r$  através das expressões

$$\hat{y}^L = \hat{y}^c - \hat{y}^r \quad \text{e} \quad \hat{y}^U = \hat{y}^c + \hat{y}^r,$$

em que,  $\hat{y}^c = (\mathbf{x}^c)^\top \hat{\beta}^c$  e  $\hat{y}^r = (\mathbf{x}^r)^\top \hat{\beta}^r$ .

Dos experimentos de Lima Neto e De Carvalho (2010) [64], eles concluíram que os modelos de regressão irrestritos de Billard e Diday (2000) [7] e de Lima Neto e De Carvalho (2008) [63] exibem algumas previsões invertidas, i.e., com  $\hat{y}_i^U < \hat{y}_i^L$ . Obviamente, este problema não é observado nas previsões obtidas via o modelo de regressão restrito. Eles concluíram também que o desempenho do modelo restrito foi superior ao modelo de Billard e Diday (2000) [7] e similar ao modelo de Lima Neto e De Carvalho (2008) [63], sendo, por vezes, superado por este último. Assim, os autores sugerem o uso do método restrito como uma alternativa ao modelo de Lima Neto e De Carvalho (2008) [63] quando este fornece previsão invertida.

## 3.2 Modelos de ajuste independente para STI

A característica mais importante de dados de séries temporais é que as observações consecutivas são dependentes e é de interesse analisar e modelar esta dependência. Esta é uma diferença em relação aos dados para análise de regressão, em que a independência ou não-correlação entre as observações deve existir. Desta forma, nos modelos de regressão a ordem das observações é irrelevante para a análise, enquanto que nos modelos de séries temporais a ordem dos dados é crucial. Como os modelos de regressão para dados intervalares foram desenvolvidos para analisar observações independentes, a modelagem de STI requer o uso de técnicas específicas.

Para tratar o problema de séries temporais com dados do tipo intervalo, nós consideramos que em cada instante de tempo,  $t = 1, 2, \dots, n$ , um intervalo é descrito por um vetor bidimensional com elementos em  $\mathbb{R}$  representados pelo limite superior (*upper bound*) e pelo limite inferior (*lower bound*), respectivamente,  $X_t^U$  e  $X_t^L$ , com  $X_t^L \leq X_t^U$ . Assim, uma série temporal do tipo intervalo é dada por

$$\mathbf{I}_t = \{[X_t^L; X_t^U] \mid X_t^L, X_t^U \in \mathbb{R}, X_t^L \leq X_t^U\}_{t=-\infty}^{\infty}.$$

Equivalentemente, no desenvolvimento dos modelos apresentados nesta seção, nós consideramos que o intervalo no instante  $t$  é definido por um vetor bidimensional  $\langle X_t^c, X_t^r \rangle$  representado pelo seu ponto médio (centro),  $X_t^c$ , e sua meia-amplitude (raio),  $X_t^r$ , em que  $X_t^c = (X_t^U + X_t^L)/2$  e  $X_t^r = (X_t^U - X_t^L)/2$ . Assim, uma STI observada pode ser representada como

$$\langle X_1^c, X_1^r \rangle, \langle X_2^c, X_2^r \rangle, \dots, \langle X_n^c, X_n^r \rangle,$$

em que  $n$  denota o número de intervalos observados na série temporal, i.e., o tamanho da amostra.

Neste caso, um intervalo  $\mathbf{I}_t = \langle X_t^c, X_t^r \rangle$  se reduz a um ponto se  $X_t^r = 0$  e ele é dito ser um intervalo degenerado. Logo,  $\mathbf{I}_t$  é dito ser uma *série temporal do tipo intervalo* se existe ao menos um  $\mathbf{I}_t$  ( $t = 1, 2, \dots, n$ ), tal que  $X_t^r > 0$ .

Seguindo a metodologia empregada em Lima Neto e De Carvalho (2008) [63], Maia e De Carvalho (2008) [69] e Lima Neto e De Carvalho (2010) [64], em que os modelos são ajustados de modo independente aos centros e às amplitudes dos intervalos, nós estendemos modelos de séries temporais para tratar dados de STI. Os modelos apresentados a seguir são integrantes da nossa publicação “*Forecasting models for interval-valued time series*” publicado no periódico *Neurocomputing* em 2008.

Em Maia et al. (2008) [67], nós apresentamos, inicialmente, metodologias para modelagem de STI baseadas em modelos autorregressivos médias móveis (ARMA). Box e Jenkins (1976) [12] desenvolveram estas técnicas para modelagem de séries temporais univariadas (dados clássicos), até então, a principal técnica de modelagem de séries temporais univariadas. Os modelos ARMA são utilizados sobre pressuposição de que o valor atual da variável é assumida ser uma função linear de observações passadas.

O segundo modelo para STI que apresentamos em Maia et al. (2008) [67] é uma metodologia baseada em redes neurais *multilayer perceptron* (MLP), visto que esta metodologia tende a comportar-se bem em séries temporais com estruturas complexas. Kaastra e Boyd (1996) [56] forneceram uma introdução geral de como modelos de redes neural artificiais podem ser desenvolvidas para modelar séries temporais econômicas e financeiras univariadas. Estes modelos são constantemente confrontados com os modelos de Box–Jenkins na literatura. Por exemplo, Sharda e Patil (1990) [84] encontraram que redes neurais ajustam séries temporais tão bem quanto os modelos de Box–Jenkins.

Em seguida, nós adaptamos uma metodologia híbrida através da combinação do modelo ARMA com redes neurais MLP, baseada na proposta de Zhang (2003) [94], com a finalidade de ajustar STI. Algumas razões nos levaram a considerar este método híbrido. Primeiro, na prática é muito difícil determinar quando uma

série temporal é gerada por um processo linear ou por um processo não-linear. Frequentemente, as séries temporais apresentam, simultaneamente, padrões lineares e não-lineares, i.e., elas raramente são puramente lineares ou puramente não-lineares. Segundo, é sabido que não existe um método único que seja melhor em todas as situações. A tarefa de determinar quando um método particular é mais eficiente que outro não é tão simples. Terceiro, métodos que combinam modelos lineares com redes neurais tendem a ter melhor desempenho na modelagem de séries com estruturas complexas de autocorrelação. Exemplos de sucesso obtido através de combinação de modelos de Box–Jenkins e redes neurais são encontrados em Su et al. (1997) [86], Taskaya-Temizel e Ahmad (2005) [87] e Tseng et al. (2002) [89].

O ponto comum nas metodologias que nós propomos em Maia et al. (2008) [67] para STI é o fato de que os modelos consistem em ajustar duas séries temporais: a série dos pontos médios dos intervalos,  $X^c$ , e a série das metades das amplitudes dos intervalos,  $X^r$ , que são representadas, respectivamente, por

$$X_t^c = \frac{X_t^U + X_t^L}{2} \quad \text{e} \quad X_t^r = \frac{X_t^U - X_t^L}{2}, \quad t = 1, 2, \dots, n. \quad (3.3)$$

### 3.2.1 Modelos ARMA nos centros e nas amplitudes de STI

Dadas as duas séries temporais na expressão (3.3),  $X^c$  e  $X^r$ , o modelo AR para STI que nós apresentamos em Maia et al. (2008) [67] consiste em ajustar estas séries tendo como variáveis predictoras valores passados obtidos dos intervalos passados destas próprias variáveis. Assim, no modelo AR para STI os valores atuais dos centros e das metades das amplitudes dos intervalos são representados pelos modelos  $\text{AR}(p^c)$  e  $\text{AR}(p^r)$ , respectivamente,

$$\begin{aligned} X_t^c &= \theta_0^c + \sum_{i=1}^{p^c} \phi_i^c X_{t-i}^c + \epsilon_t^c, \\ X_t^r &= \theta_0^r + \sum_{i=1}^{p^r} \phi_i^r X_{t-i}^r + \epsilon_t^r, \end{aligned}$$

em que  $\theta_0^c$ ,  $\theta_0^r$ ,  $\phi_i^c$  e  $\phi_i^r$  são os parâmetros desconhecidos do modelo a serem estimados,  $p^c$  e  $p^r$  são, respectivamente, a ordem do modelo AR para a série dos centros e a ordem do modelo AR para a série das metades das amplitudes e os termos  $\epsilon_t^c$  e  $\epsilon_t^r$  são os erros de cada modelo.

Analogamente, o modelo ARMA que apresentamos em Maia et al. (2008) [67] consiste em ajustar as duas séries temporais  $X^c$  e  $X^r$ . No entanto, a estrutura do modelo ARMA para STI é dada através dos modelos ARMA( $p^c, q^c$ ) e ARMA( $p^r, q^r$ ), respectivamente,

$$\begin{aligned} X_t^c &= \theta_0^c + \sum_{i=1}^{p^c} \phi_i^c X_{t-i}^c + \epsilon_t^c + \sum_{j=1}^{q^c} \theta_j^c \epsilon_{t-j}^c, \\ X_t^r &= \theta_0^r + \sum_{i=1}^{p^r} \phi_i^r X_{t-i}^r + \epsilon_t^r + \sum_{j=1}^{q^r} \theta_j^r \epsilon_{t-j}^r, \end{aligned}$$

compostos pela combinação linear de valores passados da própria série e erros passados. Nos modelos acima,  $\theta_0^c$ ,  $\phi_i^c$ ,  $\theta_j^c$  e  $\theta_0^r$ ,  $\phi_i^r$ ,  $\theta_j^r$  são os parâmetros desconhecidos do modelo a serem estimados,  $p^c$  e  $p^r$  são, respectivamente, a ordem autorregressiva do modelo ARMA para a série dos centros e a ordem autorregressiva do modelo ARMA para a série das metades das amplitudes,  $q^c$  e  $q^r$  são, respectivamente, a ordem médias móveis do modelo ARMA para a série dos centros e a ordem médias móveis do modelo ARMA para a série das metades das amplitudes e os termos  $\epsilon_t^c$  e  $\epsilon_t^r$  são os erros de cada modelo.

Nos modelos AR e ARMA apresentados para modelagem de STI através de dois modelos independentes, a previsão do limite superior do intervalo,  $\widehat{X}_t^U$ , e a previsão do limite inferior do intervalo,  $\widehat{X}_t^L$ , são obtidas através da combinação das previsões do ponto médio e da metade da amplitude fornecidas pelo modelo AR ou ARMA da seguinte forma:

$$\widehat{X}_t^U = \widehat{X}_t^c + \widehat{X}_t^r \quad \text{e} \quad \widehat{X}_t^L = \widehat{X}_t^c - \widehat{X}_t^r,$$

em que  $\widehat{X}_t^c$  e  $\widehat{X}_t^r$  representam os valores previstos do ponto médio e da metade da amplitude, respectivamente.

Procedimentos de estimação de parâmetros, assim como, mecanismos de identificação da ordem dos modelos AR e ARMA empregados no tratamento de séries temporais univariadas podem ser utilizados para estimação e identificação destes modelos para modelagem de STI. Nos modelos AR e ARMA para STI apresentados, o exame da adequação dos modelos ajustados aos centros e às amplitudes dos intervalos, por exemplo, através de análise dos resíduos, também pode ser verificado via métodos aplicados em séries temporais univariadas.

### 3.2.2 Rede neural MLP nos centros e nas amplitudes de STI

Quando a condição de linearidade na relação funcional do modelo é relaxada, vários modelos podem ser usados para previsão de séries temporais com a finalidade de capturar a estrutura não-linear. Modelos tipicamente utilizados nestas condições são as redes neurais artificiais (RNA). As redes neurais são capazes de aproximar com eficiência diversas formas de não-linearidade nos dados e, diferente dos modelos ARMA, elas não requerem quaisquer pressuposições sobre a forma do modelo.

Existe uma grande variedade de tipos de RNA, no entanto, a mais popular delas é a conhecida por *multilayer perceptron* (MLP). Em especial, redes MLP com uma única camada escondida são frequentemente utilizadas na modelagem e previsão de séries temporais.

Na rede neural MLP que usamos na modelagem da STI de modo independente, as séries temporais dos centros e das metades das amplitudes dadas na expressão (3.3) são modeladas uma por vez. Na rede referente ao ajuste dos centros dos intervalos, a relação entre a saída da rede,  $X_t^c$ , e as correspondentes entradas,  $X_{t-1}^c, X_{t-2}^c, \dots, X_{t-p^c}^c$ , é dada por

$$X_t^c = \alpha_0^c + \sum_{j=1}^{q^c} \alpha_j^c \cdot g\left(\beta_{0j}^c + \sum_{i=1}^{p^c} \beta_{ij}^c X_{t-i}^c\right) + \epsilon_t^c,$$

e no ajuste das metades das amplitudes, a relação entre a saída da rede,  $X_t^r$ , e as correspondentes entradas,  $X_{t-1}^r, X_{t-2}^r, \dots, X_{t-p^r}^r$ , é dada por

$$X_t^r = \alpha_0^r + \sum_{j=1}^{q^r} \alpha_j^r \cdot g\left(\beta_{0j}^r + \sum_{i=1}^{p^r} \beta_{ij}^r X_{t-i}^r\right) + \epsilon_t^r,$$

em que  $\alpha^c$ 's,  $\alpha^r$ 's,  $\beta^c$ 's e  $\beta^r$ 's são os parâmetros do modelo (pesos) a serem ajustados,  $p^c$  e  $p^r$  são os números de entradas em cada rede,  $q^c$  e  $q^r$  são os números de nodos na camada escondida de cada rede, e  $g$  é a função de ativação adotada nos nodos escondidos.

Observa-se que a rede neural MLP com uma camada escondida tem uma parte similar ao modelo AR referente ao atributo a ser computado pela função de ativação  $g$ . Esta função é encarregada de realizar um mapeamento não-linear de observações passadas para o valor futuro.

Existem vários tipos de funções de ativação que podem ser usados na camada escondida da rede MLP, porém, a função logística,

$$g(x) = \frac{1}{1 + e^{-x}},$$

é a mais frequentemente utilizada. A principal razão para a popularidade desta função está no fato de que sua primeira derivada (requerida durante o processo de treinamento da rede) assume uma forma simples.

Na rede neural MLP para ajuste de STI, a previsão do limite superior do intervalo,  $\hat{X}_t^U$ , e a previsão do limite inferior do intervalo,  $\hat{X}_t^L$ , são obtidas através da combinação das previsões do ponto médio e da metade da amplitude, da forma

$$\hat{X}_t^U = \hat{X}_t^c + \hat{X}_t^r \quad \text{e} \quad \hat{X}_t^L = \hat{X}_t^c - \hat{X}_t^r,$$

em que  $\hat{X}_t^c$  e  $\hat{X}_t^r$  representam os valores previstos do ponto médio e da metade da amplitude, respectivamente.

Apesar da vantagem da rede MLP não requerer pressuposições como, e.g., estacionariedade da série temporal a ser modelada e da capacidade de modelar características complexas da série, ela apresenta algumas desvantagens, por exemplo: a rede MLP requer um número grande de observações na amostra, devido ao grande número de parâmetros (pesos) a serem ajustados (Dorffner et al., 1994) [36].

### 3.2.3 Modelo híbrido nos centros e nas amplitudes de STI

Modelos ARMA e RNA são utilizados com sucesso em seus respectivos domínios linear e não-linear. Entretanto, o uso de modelos ARMA para problemas complexos não-lineares pode não ser adequado, assim como, o uso de RNA para modelar problemas lineares pode produzir resultados insatisfatórios (Zhang, 2003) [94]. O uso de metodologias que combinam modelos de diferentes domínios pode proporcionar bons resultados devido ao fato de diferentes padrões nos dados poderem ser capturados. Assim, uma metodologia híbrida que seja capaz de modelar simultaneamente processos lineares e não-lineares é uma boa estratégia para o uso prático. Por exemplo, Wedding e Cios (1996) [91] propuseram um modelo híbrido combinando rede neural do tipo função de base radial (RBF) e modelos ARMA e Hansen e Nelson (1997) [50] obtiveram sucesso na combinação de redes neurais com modelos

de Box–Jenkins. Su et al. (1997) [86] usaram um modelo híbrido com a finalidade de ajustar séries temporais que apresentam tendência e demonstraram que este modelo produz melhores previsões que o modelo de Box–Jenkins ou que a rede neural individualmente.

O modelo híbrido que consideramos em Maia et al. (2008) [67] é baseado na metodologia proposta por Zhang (2003) [94]. Na metodologia de Zhang (2003) [94], ele considerou que uma série temporal  $X_t$  é composta de uma estrutura de autocorrelação linear,  $L_t$ , e de uma componente não-linear,  $N_t$ , i.e.,

$$X_t = L_t + N_t.$$

O modelo híbrido de Zhang (2003) [94] consiste de ajustar a série temporal em dois passos. Primeiro, um modelo ARMA é usado para modelar a estrutura linear da série e, assim, obtemos a previsão linear  $\widehat{L}_t$ . Então, a série dos resíduos do modelo ARMA,

$$e_t = X_t - \widehat{L}_t,$$

deve conter informações sobre a não-linearidade da série  $X_t$ . No segundo passo, a série  $e_t$  é modelada através da rede neural MLP com a finalidade de capturar a relação não-linear da série  $X_t$ . Assim, a rede MLP usa, como entradas, as séries defasadas dos resíduos, i.e.,

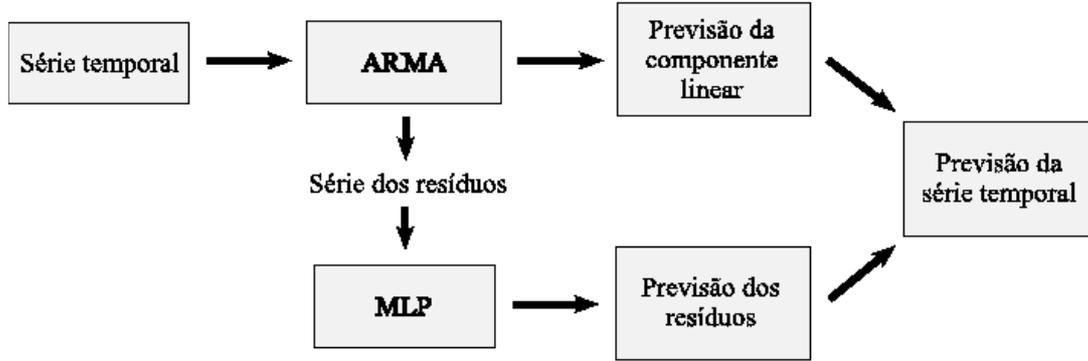
$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-p}) + \epsilon_t,$$

em que  $f$  é a função não-linear determinada pela rede neural e  $\epsilon_t$  é o erro. A previsão dos resíduos obtida via rede MLP é denotada como  $\widehat{N}_t$  e, então, a previsão combinada fornecida pelo modelo híbrido é dada por

$$\widehat{X}_t = \widehat{L}_t + \widehat{N}_t.$$

A Figura 3.1 mostra o diagrama com a estrutura da metodologia híbrida proposta por Zhang (2003) [94]. Observe que os três retângulos na parte superior do diagrama equivalem à modelagem ARMA tradicional e os dois retângulos inferiores são o segundo passo do modelo híbrido, em que a rede MLP é aplicada para modelar a série dos resíduos do modelo ARMA. O último retângulo à direita é a previsão da série obtida através da soma das previsões das componentes linear e não-linear.

Figura 3.1: Diagrama da metodologia híbrida de Zhang (2003) [94].



Nós aplicamos esta metodologia híbrida para modelar, independentemente, as séries dos centros e das metades das amplitudes dos intervalos,  $X_t^c$  e  $X_t^r$ , respectivamente. Consideramos que cada uma destas duas séries é composta de uma componente linear e uma componente não-linear,

$$\begin{aligned} X_t^c &= L_t^c + N_t^c, \\ X_t^r &= L_t^r + N_t^r. \end{aligned}$$

No primeiro passo da metodologia híbrida adaptada para modelagem de STI, o modelo ARMA é aplicado e, então, computamos as previsões das componentes lineares da série dos centros e da série das metades das amplitudes,  $\hat{L}_t^c$  e  $\hat{L}_t^r$ , respectivamente. Logo, a série dos resíduos do modelo ARMA para  $X_t^c$  e a série dos resíduos do modelo ARMA para  $X_t^r$  são, respectivamente,

$$\begin{aligned} e_t^c &= X_t^c - \hat{L}_t^c, \\ e_t^r &= X_t^r - \hat{L}_t^r. \end{aligned}$$

No segundo passo, as séries  $e_t^c$  e  $e_t^r$  são ajustadas via rede MLP e as previsões da componente não-linear da série dos centros,  $\hat{N}_t^c$ , e da componente não-linear da série das metades das amplitudes,  $\hat{N}_t^r$ , são obtidas.

Finalmente, computadas as previsões da componente linear e da componente não-linear para as séries dos centros e das amplitudes, as previsões  $\hat{X}_t^U$  e  $\hat{X}_t^L$  são obtidas da seguinte forma:

$$\begin{aligned} \hat{X}_t^U &= \hat{X}_t^c + \hat{X}_t^r = \hat{L}_t^c + \hat{N}_t^c + \hat{L}_t^r + \hat{N}_t^r = (\hat{L}_t^c + \hat{L}_t^r) + (\hat{N}_t^c + \hat{N}_t^r), \\ \hat{X}_t^L &= \hat{X}_t^c - \hat{X}_t^r = \hat{L}_t^c + \hat{N}_t^c - \hat{L}_t^r - \hat{N}_t^r = (\hat{L}_t^c - \hat{L}_t^r) + (\hat{N}_t^c - \hat{N}_t^r), \end{aligned}$$

em que  $\widehat{X}_t^c$  e  $\widehat{X}_t^r$  representam os valores previstos do ponto médio e da metade da amplitude, respectivamente.

### 3.2.4 Considerações finais

Neste capítulo, nós apresentamos métodos para modelagem e previsão de séries temporais do tipo intervalo no contexto da ADS. Os métodos foram apresentados nos artigos Maia et al. (2006, 2008) [66]-[67]. Os métodos são baseados nos modelos ARMA, em redes neurais MLP e em uma metodologia híbrida proposta por Zhang (2003) [94]. Os modelos consistem em ajustar a série dos pontos médios dos intervalos e a série das metades das amplitudes dos intervalos de modo independente. A previsão do limite superior e a previsão do limite inferior da STI são obtidas através da combinação das previsões do ponto médio e da metade da amplitude do intervalo.

Nós demonstramos a utilidade dos modelos apresentados através de experimentos envolvendo dados de STI sintéticos e envolvendo duas STI reais correspondentes aos dados de temperaturas mensais medidas em uma estação meteorológica na China e aos dados de preços diários de ações da companhia de Petróleo Brasileiro (Petrobras).

Nós usamos redes neurais MLP com uma camada escondida para modelar a série dos pontos médios e a série das metades das amplitudes de modo independente. No segundo passo da metodologia híbrida também foram usadas redes MLP com uma única camada escondida para modelar a série dos resíduos do modelo ARMA. As redes foram treinadas usando o algoritmo de aprendizagem do gradiente conjugado. A seleção do número de entradas da rede neural, bem como a seleção do número de nodos na camada escondida foi realizada através de experimentação. Várias redes MLP com diferentes números de entradas e diferentes números de nodos escondidos foram treinadas e cada rede foi avaliada no conjunto de treinamento usando um número razoável de inicializações de pesos selecionados aleatoriamente para reduzir o problema de soluções em mínimos locais (Pretchelt, 1994) [77]. A topologia da rede neural com melhor desempenho no conjunto de treinamento foi selecionada para efetuar previsões da STI.

A identificação da ordem do modelo AR e do modelo ARMA para ajuste das

séries dos centros e das metades das amplitudes dos intervalos foi realizada através do processo de minimização do *Akaike Information Criterion* (AIC) e os parâmetros foram estimados via método de máxima verossimilhança.

Os modelos foram avaliados em STI no conjunto de treinamento e em dois conjuntos de teste, um com 6 passos-à-frente e outro com 12 passos-à-frente. Os experimentos com STI sintéticas em simulação Monte Carlo e com STI reais demonstraram que os métodos desenvolvidos para modelagem de STI exibem desempenhos satisfatórios e são alternativas úteis para previsão de STI.

## Parte II

# Métodos de ajuste dependente

## Capítulo 4

# Métodos de ajuste dependente

Na Parte I da tese, nós apresentamos modelos de regressão e de séries temporais para dados simbólicos do tipo intervalo ajustados de modo independente às séries descritas pelos pontos médios e pelas metades das amplitudes dos intervalos. Os modelos que foram apresentados fazem parte do que denominamos de classe de *métodos de ajuste independente*.

Como será visto ao longo deste capítulo, os modelos apresentados na Parte I – *Métodos de ajuste independente* – podem ser escritos como casos especiais dos métodos de ajuste dependente.

Nos métodos que são apresentados no presente capítulo, nós optamos por considerar os próprios valores observados dos limites superiores e dos limites inferiores, em vez de utilizar os valores das séries dos centros e das amplitudes dos intervalos. Os métodos que iniciaremos a apresentar neste capítulo fazem parte de outra classe de métodos que nós denominamos por *métodos de ajuste dependente*. A razão para este título será facilmente entendida adiante.

Duas razões nos levaram a optar pelo desenvolvimento de métodos com base nos valores dos limites superior e inferior do intervalo. A primeira e principal razão vem do fato de que os valores dos limites superior e inferior são valores realmente observados nos dados da amostra, i.e., os valores máximo e mínimo de uma amostra sempre fazem parte do conjunto de valores que a compõe. Este fato não é sempre verdadeiro para o valor do ponto médio e para o valor da amplitude do intervalo, pois, como foi mostrado no Capítulo 3, estes valores são obtidos através de operações matemáticas sobre os dados da amostra, especificamente, sobre os

valores observados dos limites superior e inferior.

Para ilustrar a primeira razão, vamos considerar um exemplo simples. Consideremos que algumas ações de uma determinada companhia foram negociadas oito vezes durante o dia e os preços formam o conjunto  $\mathcal{P} = \{\text{R\$ } 16,00, \text{R\$ } 18,00, \text{R\$ } 15,00, \text{R\$ } 17,00, \text{R\$ } 19,00, \text{R\$ } 22,00, \text{R\$ } 21,00, \text{R\$ } 20,00\}$ . Estes valores podem ser simbolizados através do intervalo  $[\text{R\$ } 15,00; \text{R\$ } 22,00]$  que expressa a variação de preços de negociação da ação no dia representados pelo menor e pelo maior valor de negociação. Observemos que tanto o limite superior,  $p_{\text{superior}} = \text{R\$ } 22,00$ , quanto o limite inferior do intervalo,  $p_{\text{inferior}} = \text{R\$ } 15,00$ , são realizações verídicas no conjunto de preços, i.e., em notação matemática,  $p_{\text{superior}} \in \mathcal{P}$  e  $p_{\text{inferior}} \in \mathcal{P}$ . Por outro lado, o ponto médio deste intervalo,  $p_{\text{centro}} = \text{R\$ } 18,50$ , e a amplitude,  $p_{\text{amplitude}} = \text{R\$ } 7,00$ , não são observações reais do conjunto  $\mathcal{P}$ , i.e.,  $p_{\text{centro}} \notin \mathcal{P}$  e  $p_{\text{amplitude}} \notin \mathcal{P}$ .

A segunda razão para uso dos valores dos limites superior e inferior do intervalo vem do fato de que, estatisticamente falando, uma grande parcela da variabilidade da série dos limites superiores pode ser explicada pela trajetória da série dos limites inferiores e vice-versa. Em geral, espera-se que a correlação entre a série do limite superior e a série do limite inferior seja maior que a correlação entre as séries do ponto médio e da amplitude dos intervalos. Esta razão é crucial para o desenvolvimento de modelos de ajuste dependente com base nos valores dos limites superior e inferior dos intervalos e é considerada decisiva para pretermos o ponto médio e a amplitude dos intervalos.

Ilustramos a segunda razão para uso dos valores das séries dos limites superiores e dos limites inferiores, em vez das séries dos centros e das amplitudes, considerando duas STI reais. A primeira STI corresponde aos dados de temperaturas mensais medidas em uma estação meteorológica na China,<sup>1</sup> a outra corresponde aos dados de preços diários de ações da Companhia de Petróleo Brasileiro S.A. (Petrobras).<sup>2</sup>

Na Figura 4.1, podemos ver as trajetórias das séries dos limites superiores e dos limites inferiores dos intervalos para a STI de temperaturas da China (gráfico à esquerda) e de preços de ações da Petrobras (gráfico à direita). Em seguida, a Figura 4.2 mostra as trajetórias das séries dos centros e das amplitudes dos intervalos

---

<sup>1</sup>Série temporal extraída da *Long-Term Instrumental Climatic Database of the People's Republic of China* e disponível em <http://dss.ucar.edu/datasets/ds578.5/data/> (Acesso em fevereiro de 2007).

<sup>2</sup>Série temporal disponível em <http://finance.yahoo.com/q/hp?s=PBR-A> (Acesso em março de 2009).

para a STI de temperaturas da China (gráfico à esquerda) e de preços de ações da Petrobras (gráfico à direita). Através de exame visual nos gráficos de cada STI individualmente, podemos perceber que as trajetórias das séries dos limites superiores são similares às trajetórias das séries dos limites inferiores (Figura 4.1). Esta similaridade não é observada nas trajetórias das séries dos centros e das amplitudes (Figura 4.2). Os coeficientes de correlação linear entre a série dos limites superiores e a série dos limites inferiores para as STI de temperaturas da China e de ações da Petrobras são muito próximos da unidade, respectivamente, 0,991 e 0,986. Por outro lado, os coeficientes de correlação linear entre a série dos centros e a série das amplitudes destas STI são menores, respectivamente, 0,129 e -0,266.

Devidos às duas razões consideradas acima, desenvolver métodos em que informações do limite inferior possam ser utilizadas adicionalmente para analisar o limite superior e vice-versa pode ser uma estratégia mais apropriada e eficiente do que modelos que considerem a dependência entre a série dos centros e a série das amplitudes dos intervalos.

Em suma, a ideia principal das técnicas e modelos propostos neste capítulo para STI é considerar que a trajetória do limite superior pode ser afetada por realizações do limite inferior e que a trajetória do limite inferior pode ser afetada por realizações do limite superior. Sendo assim, nos modelos propostos neste capítulo, a previsão de um dos limites usa informações relativas ao outro limite da série. Esta é a razão para o título de *modelos de ajuste dependente*.

## 4.1 Metodologia adotada

Nos métodos desenvolvidos para modelagem de dados intervalares apresentados no Capítulo 3, incluindo os modelos para STI que propomos em Maia et al. (2006, 2008) [66]-[67], a metodologia adotada usa a representação do intervalo dada pelo centro e pela metade da amplitude (ou raio), i.e.,  $\mathbf{I}_t = \langle X_t^c, X_t^r \rangle$ , em que  $X_t^c = (X_t^U + X_t^L)/2$  e  $X_t^r = (X_t^U - X_t^L)/2$ .

O procedimento adotado nesta Parte II da tese para produzir extensões de métodos univalorados de séries temporais para análise de STI é diferente do descrito no Capítulo 3. Basicamente, o procedimento consiste da substituição do ponto (escalar)  $y_t$ , nos métodos usuais, pelo vetor intervalo  $\mathbf{I}_t$ . Consideramos também que

Figura 4.1: Séries dos limites superior e inferior dos intervalos para a STI de temperaturas da China (gráfico à esquerda) e de preços de ações da Petrobras (gráfico à direita).

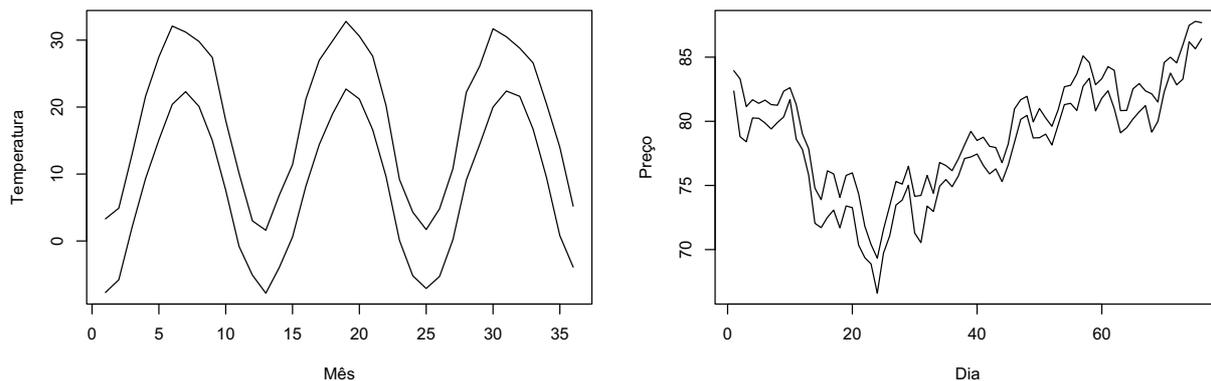
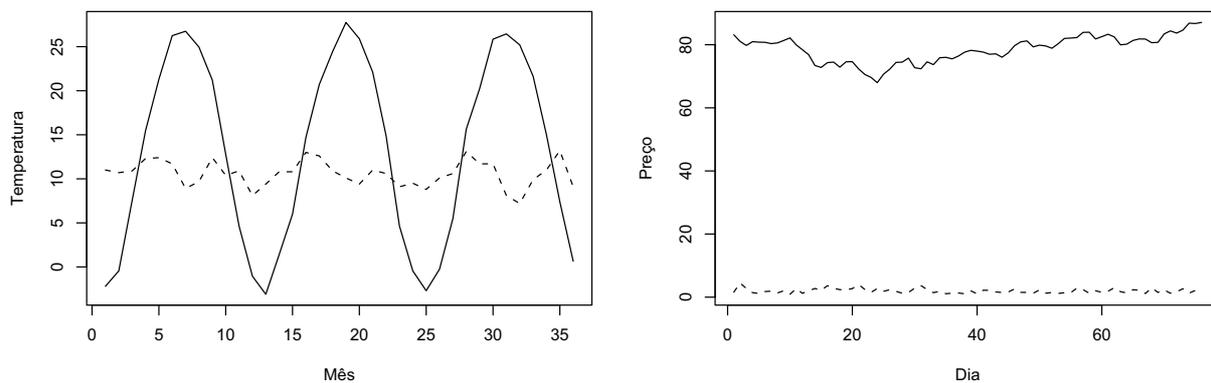


Figura 4.2: Séries dos centros (linha sólida) e das amplitudes (linha tracejada) dos intervalos para a STI de temperaturas da China (gráfico à esquerda) e de preços de ações da Petrobras (gráfico à direita).



em cada instante de tempo,  $t = 1, 2, \dots, n$ , um intervalo é descrito por um vetor bi-dimensional com elementos em  $\mathbb{R}$  representados pelo limite superior (*upper bound*) e pelo limite inferior (*lower bound*), respectivamente,  $X_t^U$  e  $X_t^L$ , com  $X_t^L \leq X_t^U$ . Assim, uma série temporal do tipo intervalo é dada por

$$\mathbf{I}_t = \{[X_t^L; X_t^U] \mid X_t^L, X_t^U \in \mathbb{R}, X_t^L \leq X_t^U\}_{t=-\infty}^{\infty}.$$

Assim, uma STI observada é representada como

$$[X_1^L; X_1^U], [X_2^L; X_2^U], \dots, [X_n^L; X_n^U],$$

em que  $n$  denota o número de intervalos observados da séries temporal, i.e., o tamanho da amostra. Em particular, denotamos um intervalo observado no tempo  $t$  por  $\mathbf{I}_t$  e representamos por um vetor  $(X_t^U, X_t^L)^\top$ , ou seja,

$$\mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix}.$$

No caso acima, um intervalo  $\mathbf{I}_t = [X_t^L; X_t^U]$  se reduz a um ponto se  $X_t^L = X_t^U$  e ele é dito ser um intervalo degenerado. Logo,  $\mathbf{I}_t$  é dito ser uma *série temporal do tipo intervalo* se existe ao menos um objeto  $\mathbf{I}_t$  ( $t = 1, 2, \dots, n$ ), tal que  $[X_t^L; X_t^U]$  é um intervalo com  $X_t^U > X_t^L$ .

Além de adotar o intervalo  $\mathbf{I}_t$  no lugar do escalar  $y_t$ , na extensão de métodos univalorados de previsão para modelar STI, outro aspecto foi considerado. Os parâmetros dos modelos clássicos de previsão são dados por escalares, enquanto que nos métodos propostos para STI, nós decidimos empregar matrizes  $(2 \times 2)$  de parâmetros. Estas matrizes assumem papel fundamental para modelagem de STI, pois é através dos seus elementos que é possível a inclusão de informações do limite inferior no modelo relativo ao limite superior e vice-versa.

Vale lembrar que séries temporais univaloradas são aquelas em que um único valor é observado em cada instante de tempo. No presente capítulo, veremos que, devido ao procedimento de substituir um ponto  $y_t$  pelo vetor intervalo  $\mathbf{I}_t = [X_t^L; X_t^U]$ , os métodos propostos podem ser vistos como generalizações de métodos clássicos.

Na próxima seção, iniciaremos a apresentação dos novos métodos para análise de STI. Quando necessário, no início de cada seção, o método que será explorado é apresentado brevemente no contexto de dados univalorados. Isto é importante

para auxiliar na compreensão de cada método e para verificarmos as semelhanças entre o método univariado e o método estendido para STI. Em seguida, apresentamos as considerações feitas para realizar a extensão destes métodos para tratar STI e, por fim, apresentamos a formulação dos métodos resultantes e exemplos da aplicabilidade.

## 4.2 Técnicas descritivas

As técnicas descritivas de análise são procedimentos simples que permitem avaliar propriedades das séries temporais. Nesta seção, apresentamos algumas técnicas de caráter descritivo que podem ser úteis na análise de STI. Apresentamos uma forma de representar graficamente uma STI através dos valores observados do limite superior e do limite inferior. Em seguida, nós propomos uma forma de medir a autocorrelação e a autocovariância amostrais em STI e, conseqüentemente, o gráfico das autocorrelações chamado de correlograma para séries intervalares. Adicionalmente, propomos novas expressões para a correlação e para a covariância amostrais em variáveis simbólicas do tipo intervalo.

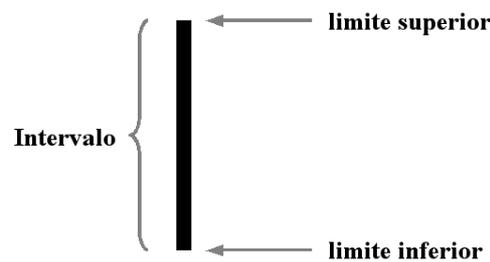
### 4.2.1 Representação gráfica de STI

A representação gráfica dos dados sequencialmente ao longo do tempo é fundamental e pode revelar padrões de comportamento importantes na série temporal. Com este gráfico podemos fazer um exame visual com relação à natureza da série temporal e observar componentes presentes nos dados. Tendências de crescimento ou decréscimo, padrões cíclicos, alterações estruturais ou descontinuidades, observações extremas, dentre outras, são características, geralmente, de fácil identificação visual. A ausência ou presença de tais componentes pode ajudar na seleção do potencial modelo para produzir melhores previsões. Sendo assim, observar o gráfico temporal dos dados deve ser sempre o primeiro passo, antecedendo qualquer outra investigação na análise de séries temporais. Isto não é diferente na análise de STI.

Antes de apresentar a representação gráfica de STI, na Figura 4.3, nós mostramos como um intervalo é composto graficamente. Cada segmento de reta vertical no

gráfico da STI representa um intervalo  $\mathbf{I}_t = [X_t^L; X_t^U]$  observado em um dado instante de tempo  $t$ . Os extremos inferior e superior de cada segmento de reta correspondem, respectivamente, ao valor do limite inferior e ao valor do limite superior do intervalo observado, como detalhado na Figura 4.3. O comprimento do segmento de reta representa a amplitude do intervalo e dá indícios – podemos não conhecer os dados que compõem o intervalo – da variabilidade dos dados neste instante de tempo.

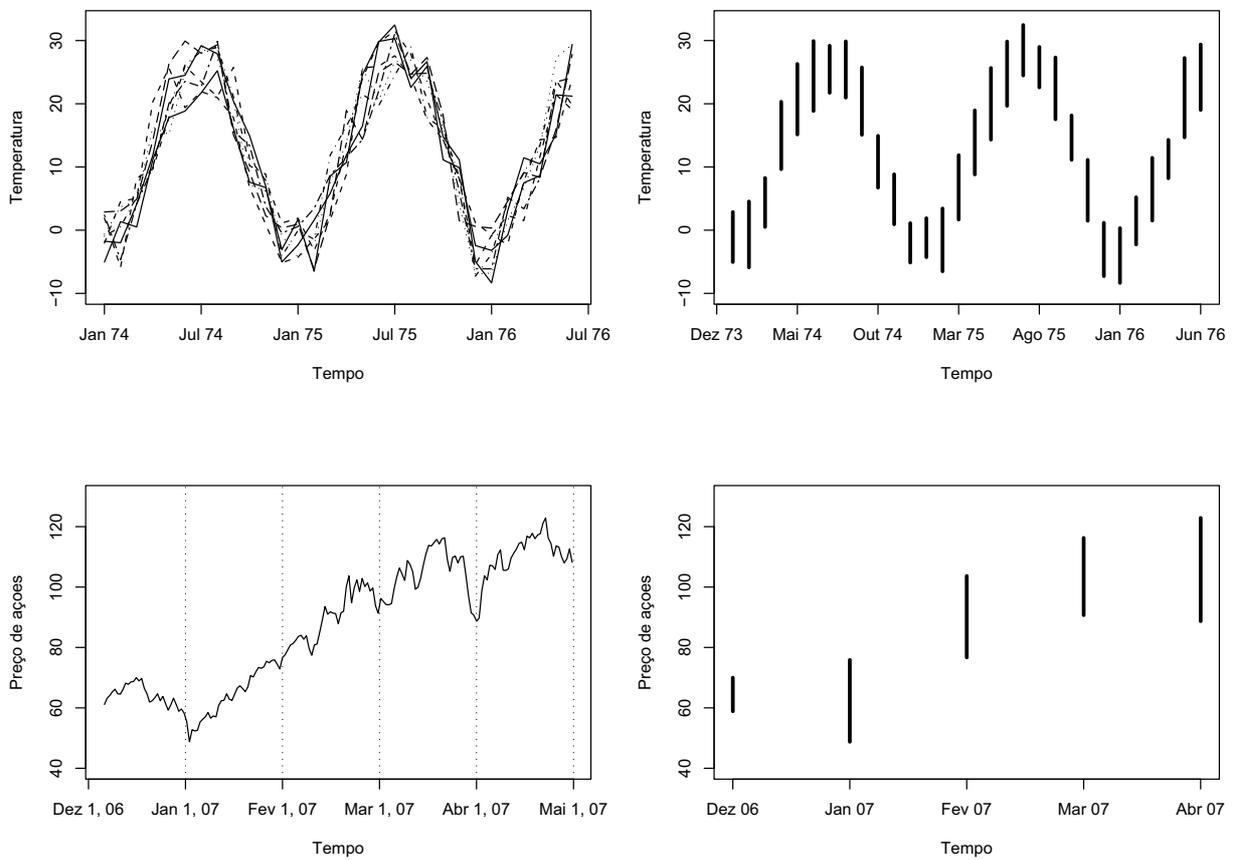
Figura 4.3: Exemplo de concepção de um intervalo.



No gráfico temporal dos dados está a primeira diferença entre a análise clássica de séries temporais e a análise de STI. Para entender melhor esta diferença, na Figura 4.4 ilustramos duas situações típicas onde STI podem ser derivadas de séries temporais de dados univariados. O gráfico na parte superior à esquerda da Figura 4.4 representa várias séries univariadas de temperaturas observadas em várias estações meteorológicas na China. Cada curva neste gráfico representa uma série de temperaturas em uma determinada estação meteorológica. No lado direito deste gráfico está a STI obtida tomando-se os valores mínimos e máximos observados em cada instante de tempo nas séries univariadas das temperaturas medidas em cada estação. Os dois gráficos inferiores correspondem à série dos preços de ações da Petrobras e à STI obtida tomando-se o maior valor de negociação mensal da ação e o menor valor de negociação mensal da ação. Os dois gráficos no lado direito da Figura 4.4 são as representações gráficas de ambas as STI.

**Exemplo 4.1** (Representação gráfica de STI). Para exemplificar a técnica de representação gráfica de STI, nós consideramos às STI de intervalos de temperaturas mensais na China e de intervalos de preços diários de negociações de ações da Petrobras. Estas STI estão representadas, respectivamente, nas Figuras 4.5 e 4.6. Algumas características podem ser observadas nos dois gráficos. Por exemplo, para

Figura 4.4: Séries temporais do tipo intervalo (gráficos à direita) obtidas a partir de séries temporais de dados clássicos (gráficos à esquerda).



a STI das temperaturas na China, o gráfico indica a presença de comportamento cíclico com duração de 12 meses similar à sazonalidade em séries temporais univariadas. No gráfico dos preços de ações da Petrobras, nota-se tendência decrescente até o 24º dia e inversão da tendência após este período. As amplitudes dos intervalos das temperaturas são relativamente constantes, enquanto que na STI das ações da Petrobras, os intervalos têm maior variabilidade nos comprimentos.

### 4.2.2 Autocorrelação e autocovariância em STI

Nesta seção nós propomos expressões para cálculo da autocorrelação amostral e da autocovariância amostral em séries temporais do tipo intervalo. Adicionalmente, são propostas novas formas para obtenção da correlação e da covariância amostrais, conseqüentemente, da variância amostral, para variáveis simbólicas do tipo intervalo. Expressões para cálculo da correlação e da covariância amostrais para variáveis intervalares já foram apresentadas em outros trabalhos, e.g., em Billiard e Diday (2002) [8] e em Gioia e Lauro (2005) [44]. No entanto, nesta tese, expressões da autocorrelação e da autocovariância amostral são apresentadas pela primeira vez no contexto de ADS, especificamente, para analisar STI.

Uma importante ferramenta para identificar propriedades de uma série temporal consiste de uma quantidade chamada autocorrelação amostral. Esta quantidade é utilizada para avaliar o grau de dependência serial nos dados. Uma prática comum em análise de séries temporais é analisar a função de autocorrelação amostral (ACF) com a finalidade de identificar modelos univariados que representem esta dependência nos dados.

Em análise de dados univariados, a ideia da autocorrelação para séries temporais é similar ao coeficiente de correlação usual. Consideremos  $n$  pares de observações das variáveis aleatórias  $X$  e  $Y$ , então, uma maneira de obter o coeficiente de correlação amostral,  $r$ , em dados clássicos é

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (4.1)$$

Figura 4.5: Série dos intervalos de temperaturas mensais medidas em estação meteorológica na China.

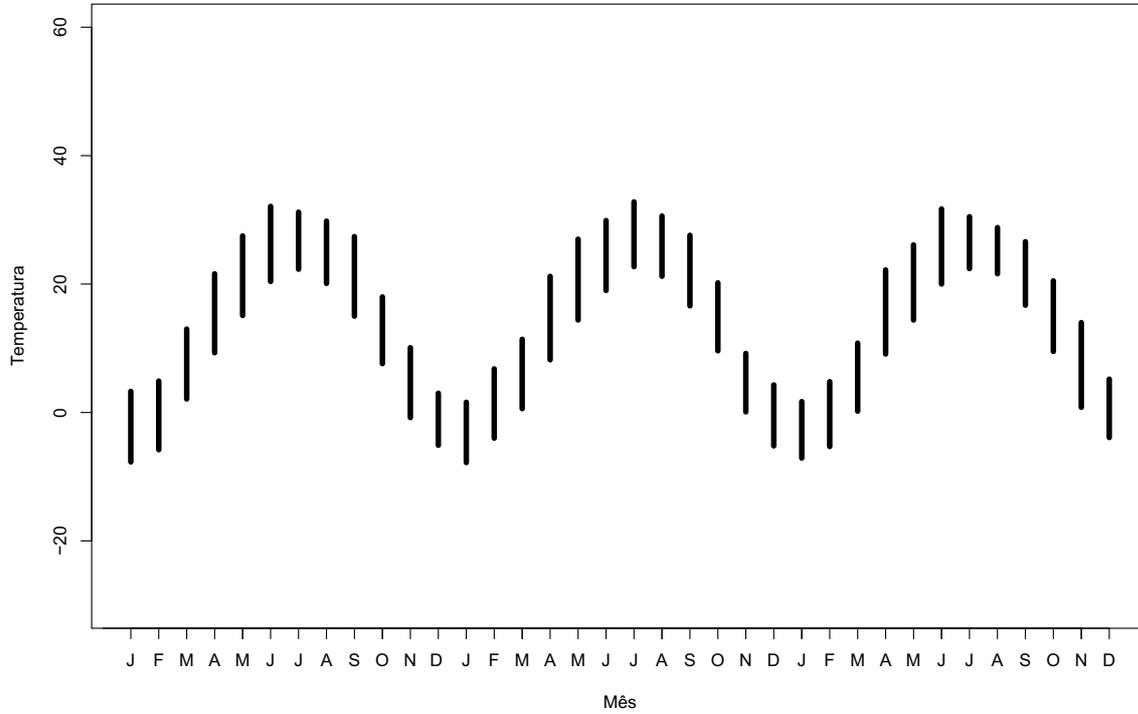
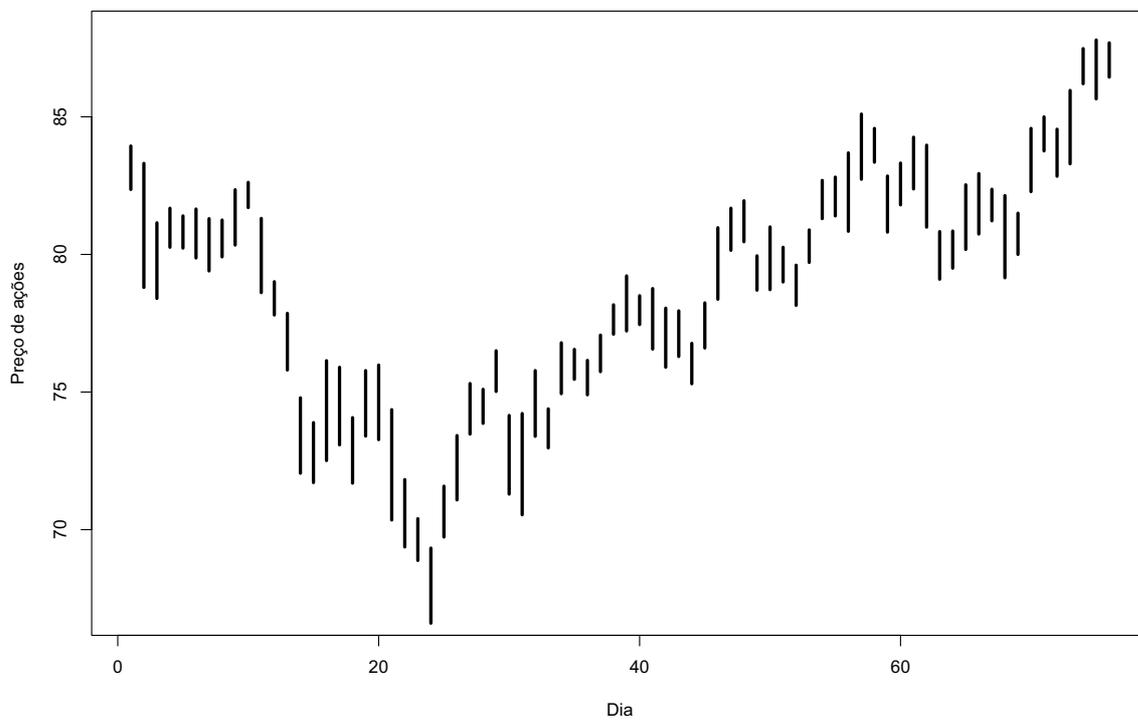


Figura 4.6: Série dos intervalos de preços diários de ações da Petrobras.



No entanto, em séries temporais, o interesse é mensurar a correlação entre observações de uma mesma variável em diferentes horizontes de tempo, ou seja, medir a correlação entre observações defasadas  $k$  períodos de tempo. Quando o interesse é avaliar a dependência linear entre uma série temporal  $y_t$  e a série de seus valores passados  $y_{t-k}$ , o conceito de correlação é generalizado para autocorrelação.

A equação (4.1) pode ser usada para calcular a autocorrelação entre observações de séries temporais defasadas de  $k$  períodos de tempo da seguinte forma

$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}. \quad (4.2)$$

Estendendo os coeficientes de correlação (4.1) e de autocorrelação (4.2) amostrais para dados simbólicos do tipo intervalo, podemos medir, respectivamente, a correlação entre duas variáveis intervalares e a autocorrelação entre intervalos defasados de  $k$  períodos de tempo. Dadas duas amostras de duas variáveis simbólicas do tipo intervalo,  $X_1^I, X_2^I, \dots, X_n^I$  e  $Y_1^I, Y_2^I, \dots, Y_n^I$ , em que  $X_i^I = (X_i^U, X_i^L)^\top$  e  $Y_i^I = (Y_i^U, Y_i^L)^\top$ , e sejam os intervalos médios amostrais

$$\bar{X}^I = \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n X_i^U \\ \sum_{i=1}^n X_i^L \end{bmatrix} \quad \text{e} \quad \bar{Y}^I = \begin{bmatrix} \bar{Y}^U \\ \bar{Y}^L \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n Y_i^U \\ \sum_{i=1}^n Y_i^L \end{bmatrix},$$

em que  $\bar{X}^U$  e  $\bar{Y}^U$  são as médias amostrais dos limites superiores das variáveis  $X^I$  e  $Y^I$ , respectivamente e  $\bar{X}^L$  e  $\bar{Y}^L$  são as médias amostrais dos limites inferiores;  $X_i^U$  e  $X_i^L$  são, respectivamente, o limite superior e o limite inferior do  $i$ -ésimo intervalo de  $X^I$  e  $Y_i^U$  e  $Y_i^L$  são, respectivamente, o limite superior e o limite inferior do  $i$ -ésimo intervalo de  $Y^I$ . Definimos, então, o coeficiente de correlação amostral para dados intervalares,  $r^I$ , como

$$r^I = \frac{\sum_{i=1}^n (X_i^I - \bar{X}^I)^\top (Y_i^I - \bar{Y}^I)}{\sqrt{\sum_{i=1}^n (X_i^I - \bar{X}^I)^\top (X_i^I - \bar{X}^I) \sum_{i=1}^n (Y_i^I - \bar{Y}^I)^\top (Y_i^I - \bar{Y}^I)}}. \quad (4.3)$$

Expandindo a expressão (4.3), da forma

$$r^I = \frac{\sum_{i=1}^n \left( \begin{bmatrix} X_i^U \\ X_i^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)^\top \left( \begin{bmatrix} Y_i^U \\ Y_i^L \end{bmatrix} - \begin{bmatrix} \bar{Y}^U \\ \bar{Y}^L \end{bmatrix} \right)}{\sqrt{\sum_{i=1}^n \left( \begin{bmatrix} X_i^U - \bar{X}^U \\ X_i^L - \bar{X}^L \end{bmatrix}^\top \begin{bmatrix} X_i^U - \bar{X}^U \\ X_i^L - \bar{X}^L \end{bmatrix} \right) \sum_{i=1}^n \left( \begin{bmatrix} Y_i^U - \bar{Y}^U \\ Y_i^L - \bar{Y}^L \end{bmatrix}^\top \begin{bmatrix} Y_i^U - \bar{Y}^U \\ Y_i^L - \bar{Y}^L \end{bmatrix} \right)}}$$

chegamos a

$$r^I = \frac{\sum_{i=1}^n (X_i^U - \bar{X}^U)(Y_i^U - \bar{Y}^U) + \sum_{i=1}^n (X_i^L - \bar{X}^L)(Y_i^L - \bar{Y}^L)}{\sqrt{\sum_{i=1}^n \left( (X_i^U - \bar{X}^U)^2 + (X_i^L - \bar{X}^L)^2 \right) \sum_{i=1}^n \left( (Y_i^U - \bar{Y}^U)^2 + (Y_i^L - \bar{Y}^L)^2 \right)}}$$

■

Assim como em dados clássicos, em que a correlação amostral entre dois vetores iguais é  $r = 1$ , em dados intervalares, se  $X_i^I = Y_i^I$ , para  $i = 1, 2, \dots, n$ , a correlação amostral é  $r^I = 1$ . Considerando o caso extremo  $X_i^I = -Y_i^I$ , para  $i = 1, 2, \dots, n$ , de modo similar à correlação em dados clássicos, temos que  $r^I = -1$ .

No contexto de STI, dada uma sequência  $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n$  de  $n$  intervalos, definimos, a partir de (4.2), a autocorrelação amostral,  $r_k^I$ , como a sequência de valores

$$r_k^I = \frac{\sum_{t=1}^{n-k} (\mathbf{I}_t - \bar{\mathbf{I}})^\top (\mathbf{I}_{t+k} - \bar{\mathbf{I}})}{\sum_{t=1}^n (\mathbf{I}_t - \bar{\mathbf{I}})^\top (\mathbf{I}_t - \bar{\mathbf{I}})}, \quad (4.4)$$

em que,  $\mathbf{I}_t$  representa o  $t$ -ésimo intervalo e  $\bar{\mathbf{I}}$  é o intervalo médio amostral, respectivamente,

$$\mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} \quad \text{e} \quad \bar{\mathbf{I}} = \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \sum_{t=1}^n X_t^U \\ \sum_{t=1}^n X_t^L \end{bmatrix}.$$

Expandindo a representação matricial (4.4), ou seja, resolvendo

$$r_k^I = \frac{\sum_{t=1}^{n-k} \left( \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)^\top \left( \begin{bmatrix} X_{t+k}^U \\ X_{t+k}^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)}{\sum_{t=1}^n \left( \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)^\top \left( \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)},$$

a função de autocorrelação amostral para STI (ACF<sup>I</sup>) proposta é expressa como a sequência

$$r_k^I = \frac{\sum_{t=1}^{n-k} (X_t^U - \bar{X}^U)(X_{t+k}^U - \bar{X}^U) + \sum_{t=1}^{n-k} (X_t^L - \bar{X}^L)(X_{t+k}^L - \bar{X}^L)}{\sum_{t=1}^n (X_t^U - \bar{X}^U)^2 + \sum_{t=1}^n (X_t^L - \bar{X}^L)^2}. \quad (4.5)$$

■

É fácil ver que o coeficiente de autocorrelação proposto para STI mantém a propriedade de correlação perfeita da autocorrelação usual, i.e., para  $k = 0$ , a autocorrelação  $r_k^I$  é igual a 1.

As expressões dos coeficientes de correlação e de autocorrelação amostrais para dados do tipo intervalo foram construídas, simplesmente, através da substituição de um ponto  $y_t$  por um vetor intervalo  $\mathbf{I}_t$ , nas estatísticas usuais para dados univariados. Assim, os coeficientes propostos podem ser vistos com generalizações dos coeficientes clássicos.

Com base na mesma ideia, estendemos as funções de covariância e de autocovariância amostrais usadas em dados clássicos, respectivamente,

$$c = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad \text{e} \quad c_k = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x}),$$

para analisar dados simbólicos do tipo intervalo. A covariância amostral entre duas amostras de duas variáveis do tipo intervalo,  $X_1^I, X_2^I, \dots, X_n^I$  e  $Y_1^I, Y_2^I, \dots, Y_n^I$  é dada por

$$c^I = \frac{1}{2n} \sum_{i=1}^n (X_i^I - \bar{X}^I)^\top (Y_i^I - \bar{Y}^I)$$

$$\begin{aligned}
&= \frac{1}{2n} \sum_{i=1}^n \left( \begin{bmatrix} X_i^U \\ X_i^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)^\top \left( \begin{bmatrix} Y_i^U \\ Y_i^L \end{bmatrix} - \begin{bmatrix} \bar{Y}^U \\ \bar{Y}^L \end{bmatrix} \right) \\
&= \frac{1}{2n} \sum_{i=1}^n (X_i^U - \bar{X}^U)(Y_i^U - \bar{Y}^U) + \frac{1}{2n} \sum_{i=1}^n (X_i^L - \bar{X}^L)(Y_i^L - \bar{Y}^L)
\end{aligned}$$

■

e a autocovariância amostral em STI com defasagem de  $k$  períodos de tempo,

$$\begin{aligned}
c_k^I &= \frac{1}{2n} \sum_{t=1}^{n-k} (\mathbf{I}_t - \bar{\mathbf{I}})^\top (\mathbf{I}_{t+k} - \bar{\mathbf{I}}) \\
&= \frac{1}{2n} \sum_{t=1}^{n-k} \left( \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)^\top \left( \begin{bmatrix} X_{t+k}^U \\ X_{t+k}^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right) \\
&= \frac{1}{2n} \sum_{t=1}^{n-k} (X_t^U - \bar{X}^U)(X_{t+k}^U - \bar{X}^U) + \frac{1}{2n} \sum_{t=1}^{n-k} (X_t^L - \bar{X}^L)(X_{t+k}^L - \bar{X}^L).
\end{aligned}$$

■

O denominador  $2n$ , em vez de  $n$ , deve-se ao fato de estarmos realizando a soma de dois vetores, ambos de tamanho  $n$ . Com isso, as expressões para STI propostas equivalem às expressões usadas em dados clássicos se aplicarmos a séries univariadas, i.e., intervalos com limite superior igual ao limite inferior.

Note que a autocorrelação para STI pode ser estimada através do quociente da covariância de ordem  $k$  e a covariância de ordem 0,

$$r_k^I = \frac{c_k^I}{c_0^I}.$$

Como em dados univariados, nós definimos a variância amostral para dados do tipo intervalo,  $s_1^2$ , como sendo a covariância de ordem  $k = 0$ , i.e.,

$$\begin{aligned}
s_1^2 = c_0^I &= \frac{1}{2n} \sum_{t=1}^n (\mathbf{I}_t - \bar{\mathbf{I}})^\top (\mathbf{I}_t - \bar{\mathbf{I}}) \\
&= \frac{1}{2n} \sum_{t=1}^n \left( \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right)^\top \left( \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} - \begin{bmatrix} \bar{X}^U \\ \bar{X}^L \end{bmatrix} \right) \\
&= \frac{1}{2n} \sum_{t=1}^n (X_t^U - \bar{X}^U)^2 + \frac{1}{2n} \sum_{t=1}^n (X_t^L - \bar{X}^L)^2.
\end{aligned}$$



Em dados clássicos, é mostrado que, para processos puramente aleatórios com variância finita, a autocorrelação amostral,  $r_k$ , é assintoticamente normalmente distribuída com média  $-1/n$  e desvio-padrão  $1/\sqrt{n}$  para  $k \neq 0$ ; propriedades de  $r_k$  podem ser vistas em Kendall, Stuart e Ord (1983) [57]. Consequentemente, limites de confiança aproximados para  $r_k$  são dados por  $-1/n \pm \Phi^{-1}(1 - \frac{\alpha}{2})/\sqrt{n}$ , em que  $\Phi^{-1}(1 - \frac{\alpha}{2})$  representa a função quantil da distribuição normal padrão que acumula probabilidade  $1 - \frac{\alpha}{2}$ . Usualmente, os limites de 95% de confiança ( $\alpha = 0,05$ ) são  $-1/n \pm 1,96/\sqrt{n}$ , visto que 1,96 é o quantil que acumula probabilidade 0,975 na distribuição normal padrão. A princípio, consideramos que os limites de confiança para as autocorrelações amostrais em STI,  $r_k^I$ , são representados pelos limites apresentados para  $r_k$ .

Na prática, para auxiliar na identificação de características na série temporal, observa-se o gráfico com os  $k$  primeiros coeficientes de autocorrelação em função de  $k$ , em que  $k$  é a ordem de defasagem da série. Este gráfico de  $r_k^I$  contra  $k$  é chamado de correlograma.

**Exemplo 4.2** (Correlograma para STI). Para ilustrar o correlograma para STI, voltamos aos exemplos das séries dos intervalos de temperaturas na China e dos intervalos de preços de ações da Petrobras apresentados na Seção 4.2.1. As Figuras 4.7 e 4.8 apresentam os correlogramas para estas duas STI, respectivamente. As autocorrelações foram computadas através da estatística  $r_k^I$  para as 30 primeiras defasagens ( $k = 0, 1, \dots, 29$ ); as linhas horizontais tracejadas representam os limites de 95% de confiança de  $r_k$ .

Conforme as regras de interpretação do correlograma adotadas em séries temporais univariadas, também podemos associar certos padrões do correlograma para STI como determinadas características da STI. Nos correlogramas das Figuras 4.7 e 4.8 são constatadas, por exemplo, evidências de não estacionariedade nas STI.<sup>3</sup> No correlograma da STI de temperaturas na China (Figura 4.7) observa-se o comportamento oscilatório das autocorrelações similar ao comportamento de correlogramas em séries univariadas sazonais e no correlograma da STI de preços de ações da

---

<sup>3</sup>O conceito de estacionariedade para STI será apresentado mais adiante na Seção 4.5

Petrobras (Figura 4.8) observa-se o decaimento lento das autocorrelações que pode indicar presença de tendência na STI.

Em geral, o interesse maior é computar a função de autocorrelação amostral para os resíduos de um modelo de previsão ajustado à série temporal. A finalidade desta prática é avaliar se os resíduos são correlacionados no tempo, ou seja, se os resíduos apresentam correlação serial. Se esta característica for detectada nos resíduos, há indícios de que modelo ajustado não capturou toda a correlação serial presente nos dados e, então, outro modelo pode ser mais apropriado. Esta funcionalidade do correlograma será explorada mais adiante.

Nesta seção, nós propomos formas de medir a autocorrelação e a autocovariância amostrais em STI, denominadas  $r_k^I$  e  $c_k^I$ , respectivamente. Com base na estatística  $r_k^I$ , nós apresentamos o gráfico das autocorrelações chamado de correlograma para STI. O objetivo deste correlograma é proporcionar uma ferramenta que auxilie na identificação, a partir da avaliação de padrões do gráfico da  $ACF^I$ , de características presentes em STI. Adicionalmente, além das estatísticas  $r_k^I$  e  $c_k^I$ , propomos novas expressões para a correlação, para a variância e para a covariância amostrais em variáveis simbólicas do tipo intervalo denominadas, respectivamente, de  $r^I$ ,  $s^2$  e  $c^I$ .

### 4.3 Objetivos da previsão de STI

Como na análise de séries temporais univariadas, na ADS o principal objetivo da análise de STI é desenvolver modelos matemáticos que forneçam descrições plausíveis para os dados da amostra e, assim, proporcionar um ajuste estatístico que descreva características dos dados que parecem flutuar de forma aleatória ao longo do tempo. Nós assumimos que uma STI pode ser definida como um conjunto de variáveis intervalares indexadas de acordo com a ordem em que são obtidas no tempo. Ou seja, consideramos uma STI como uma sequência de variáveis intervalares,  $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n$ , em que a variável  $\mathbf{I}_1$  denota o intervalo assumido pela primeira vez na STI, a variável  $\mathbf{I}_2$  denota o intervalo para o segundo período, e assim sucessivamente. Nesta tese, uma coleção de variáveis intervalares,  $\{\mathbf{I}_t\}$ , indexadas por  $t$  é referida como um processo estocástico intervalar. Um intervalo  $\mathbf{I}_t$  é referido como uma realização do processo estocástico intervalar.

Figura 4.7: Correlograma (gráfico inferior) da STI de temperaturas mensais medidas em estação meteorológica na China.

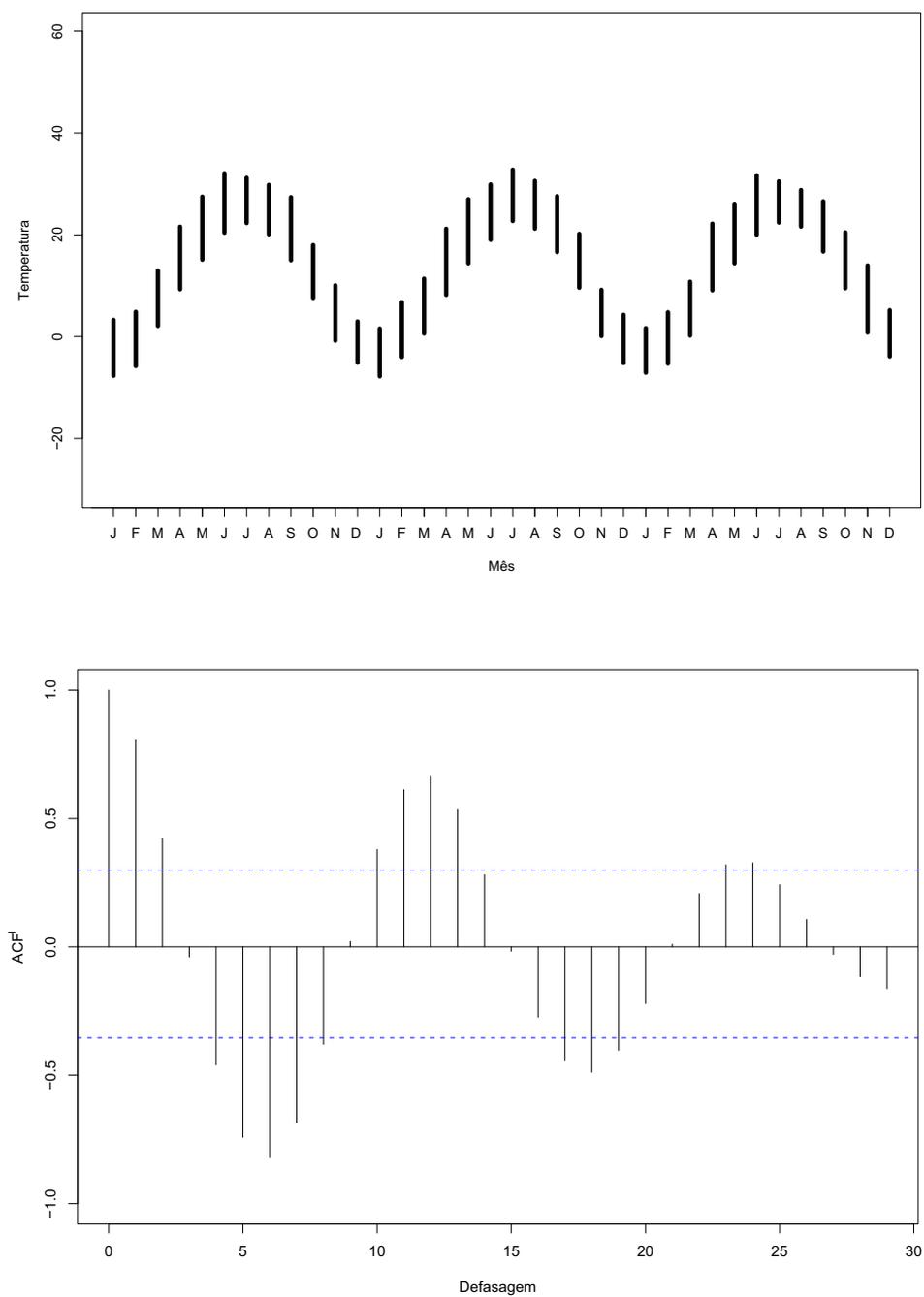
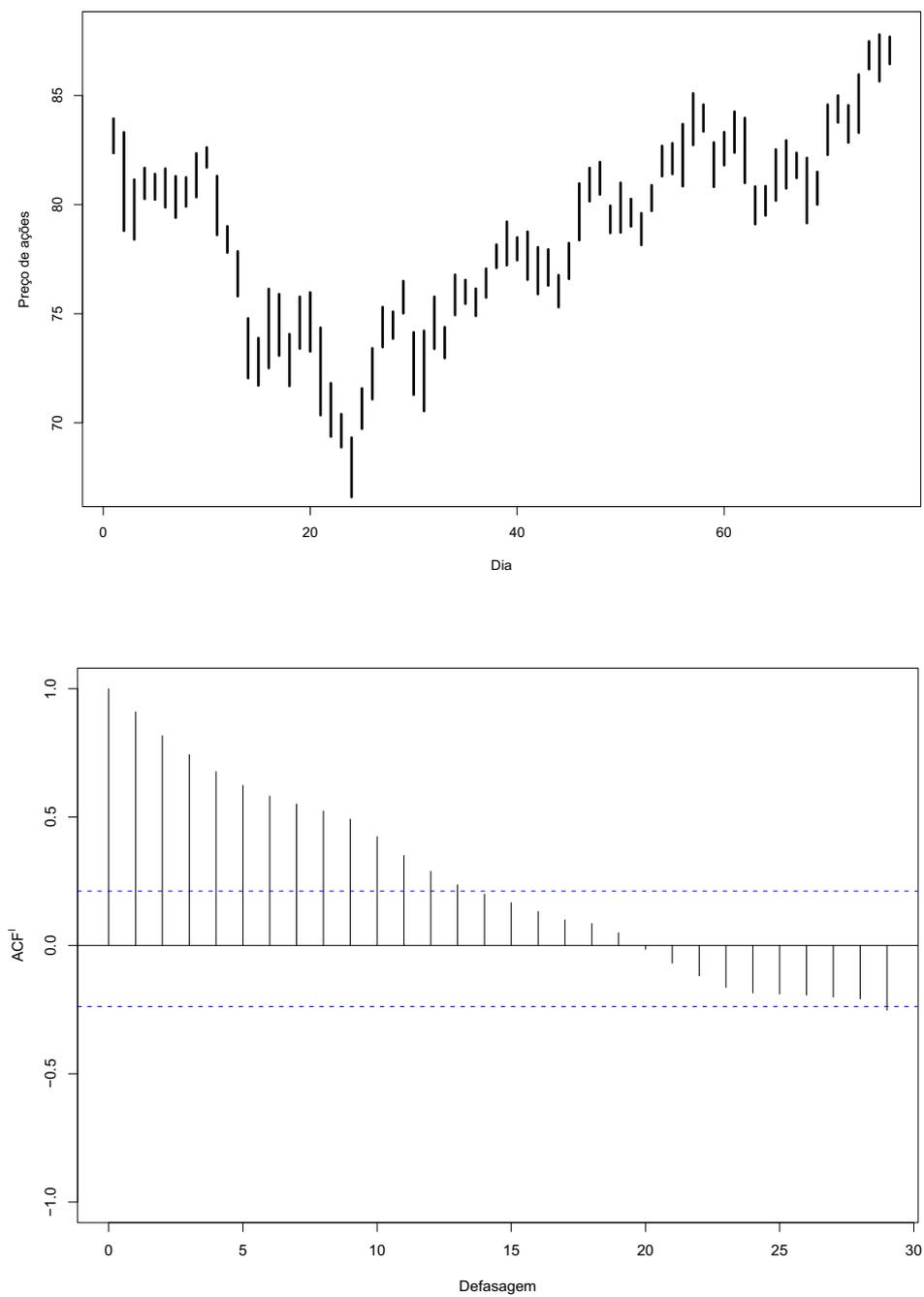


Figura 4.8: Correlograma (gráfico inferior) da STI de preços diários de ações da Petrobras.



Com base em definições que adotamos no contexto da ADS, nós desenvolvemos modelos capazes de descrever STI. Apresentamos a formulação matemática de cada modelo proposto e descrevemos os procedimentos adotados para estimação de parâmetros. Os procedimentos de estimação envolvem métodos clássicos de otimização numérica, em que as estimativas são obtidas a partir da minimização direta da função objetivo, utilizando processos iterativos. Para alguns modelos nós apresentamos procedimentos de estimação com restrições para garantir a coerência matemática dos intervalos previstos.

## 4.4 Métodos de alisamento exponencial para STI

Os métodos de previsão baseiam-se na ideia de que observações passadas contêm informações sobre o padrão de comportamento da série temporal. Sendo assim, o propósito destes métodos é distinguir tal padrão de qualquer ruído contido nas observações e, então, usar este padrão para prever valores futuros da série (Morettin e Toloí, 2006) [74].

Uma das principais técnicas de previsão de séries temporais é a denominada por métodos de alisamento exponencial (AE). Estes métodos foram apresentados originalmente com os trabalhos de Brown (1959) [15], Holt (1959, publicado novamente em 2004) [53] e Winters (1960) [92]. Uma discussão detalhada sobre os métodos AE é encontrada em Gardner (1985, 2006) [40]-[41]. Estes métodos se tornaram muito populares e atraentes devido à simplicidade e à fácil implementação computacional. Além destes pontos positivos, os métodos de AE mostram razoável precisão global na previsão de séries temporais, como ilustrado em Makridakis et al. (1998) [71] e Kotsialos et al. (2005) [59].

Os métodos de AE são classificados como univariadas, pois nenhuma covariável é utilizada para prever a série, ou seja, os dados são assumidos serem auto-explicativos. Em suma, AE trata-se de uma média ponderada de valores passados do processo observado, em que os pesos atribuídos aos valores passados decrescem à medida que as observações são mais antigas. Em outras palavras, as observações mais recentes da série recebem pesos relativamente maiores na previsão que as observações mais antigas.

### 4.4.1 Alisamento exponencial simples para STI

O termo alisamento exponencial é aplicado genericamente a uma variedade de métodos que consistem de realizar previsões com fórmulas simples de atualização e podem incorporar mudanças no nível local, tendência e sazonalidade. A versão mais simples destes métodos, muito utilizado para previsões a curto prazo, é conhecida por *alisamento exponencial simples* (AES). O AES computa a previsão um-passo-à-frente através de fórmula equivalente à soma geométrica de observações passadas, ou seja,

$$\hat{y}_{n+1} = a_0 y_n + a_1 y_{n-1} + \dots, \quad (4.6)$$

em que  $\{a_j\}$  são os pesos associados às defasagens da série que decrescem geometricamente a uma taxa constante, isto é,  $a_0 > a_1 > \dots$ . Com isso, observações mais recentes recebem peso maior que observações mais distantes no passado. No modelo considerado, os pesos são dados por

$$a_j = \alpha(1 - \alpha)^j, \quad j = 0, 1, \dots,$$

em que  $0 \leq \alpha \leq 1$  é chamada de *constante de alisamento* ou *parâmetro de amortecimento*. Assim, podemos reescrever a previsão um-passo-à-frente (4.6) da seguinte forma

$$\hat{y}_{n+1} = \alpha y_n + \alpha(1 - \alpha)y_{n-1} + \alpha(1 - \alpha)^2 y_{n-2} + \dots, \quad (4.7)$$

Na prática, (4.7) é expressa na forma de equação recursiva. Colocando  $(1 - \alpha)$  em evidência, temos

$$\hat{y}_{n+1} = \alpha y_n + (1 - \alpha)\hat{y}_n. \quad (4.8)$$

Mostramos até agora como obter a previsão um-passo-à-frente através do AES. As previsões dentro da amostra, são obtidas de modo análogo. Estabelecendo a previsão inicial  $\hat{y}_2 = y_1$  e definindo  $e_t = y_t - \hat{y}_t$ , usamos a expressão na forma de equação recursiva (4.8) para obter as previsões no tempo  $t + 1$ , i.e.,

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t. \quad (4.9)$$

Com base na metodologia discutida na Seção 4.1, estendemos o método de alisamento exponencial simples para prever STI. A metodologia consiste em substituir o ponto  $y_t$  pelo vetor intervalo  $\mathbf{I}_t = [X_t^L; X_t^U]$  e os parâmetros escalares por matrizes

( $2 \times 2$ ) de parâmetros. Desta forma, a partir da expressão (4.9) definimos o método de AES para STI (AES<sup>I</sup>) da seguinte forma

$$\widehat{\mathbf{I}}_{t+1} = \mathcal{A}\mathbf{I}_t + (\mathbf{I} - \mathcal{A})\widehat{\mathbf{I}}_t, \quad (4.10)$$

em que  $\mathcal{A}$  é a matriz ( $2 \times 2$ ) de constantes de alisamento e  $\mathbf{I}$  é uma matriz identidade ( $2 \times 2$ ), ou seja,

$$\mathcal{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \quad \text{e} \quad \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Para manter a restrição imposta à constante de alisamento  $\alpha$  no método clássico, nós assumimos que os elementos da matriz  $\mathcal{A}$  também são restritos ao intervalo  $(0, 1)$ , i.e.,  $0 \leq \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \leq 1$ .

Assim, expandindo a expressão matricial (4.10), podemos mostrar que o método AES<sup>I</sup> pode ser escrito através de um vetor bidimensional de equações,

$$\begin{aligned} \widehat{\mathbf{I}}_{t+1} = \begin{bmatrix} \widehat{X}_{t+1}^U \\ \widehat{X}_{t+1}^L \end{bmatrix} &= \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} + \begin{bmatrix} 1 - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & 1 - \alpha_{22} \end{bmatrix} \begin{bmatrix} \widehat{X}_t^U \\ \widehat{X}_t^L \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{11}X_t^U + \alpha_{12}X_t^L \\ \alpha_{22}X_t^L + \alpha_{21}X_t^U \end{bmatrix} + \begin{bmatrix} (1 - \alpha_{11})\widehat{X}_t^U - \alpha_{12}\widehat{X}_t^L \\ (1 - \alpha_{22})\widehat{X}_t^L - \alpha_{21}\widehat{X}_t^U \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{11}X_t^U + (1 - \alpha_{11})\widehat{X}_t^U + \alpha_{12}e_t^L \\ \alpha_{22}X_t^L + (1 - \alpha_{22})\widehat{X}_t^L + \alpha_{21}e_t^U \end{bmatrix}, \end{aligned} \quad (4.11)$$

■

com  $e_t^U = X_t^U - \widehat{X}_t^U$  e  $e_t^L = X_t^L - \widehat{X}_t^L$  representando, respectivamente, o erro de previsão do limite superior e o erro de previsão do limite inferior no tempo  $t$ . Observe que a previsão um-passo-à-frente de um limite do intervalo usa informação do outro limite através de uma proporção  $\alpha_{ij}$ ,  $i \neq j$ , do erro de previsão. Observe, por exemplo, que quanto maior é a constante de alisamento  $\alpha_{12}$ , maior é a influência do limite inferior na previsão do limite superior.

Os valores de  $\alpha_{ij}$ , para  $i = 1, 2$  e  $j = 1, 2$ , devem ser especificados de modo a refletir a influência de intervalos observados no passado nos intervalos previstos. Note que, quanto mais próximas da unidade são os elementos da diagonal principal da matriz  $\mathcal{A}$ , o método proposto produzirá previsões que dependem de intervalos

mais recentes. Se os elementos da diagonal secundária de  $\mathcal{A}$  são nulos, i.e.,  $\alpha_{12} = \alpha_{21} = 0$ , tem-se modelos independentes, ou seja, a previsão de um dos limites independe de informações do outro limite. No caso particular, se  $\mathcal{A}$  é uma matriz identidade, reduzimos a um processo passeio aleatório do tipo intervalo, i.e.,  $\widehat{\mathbf{I}}_{t+1} = \mathbf{I}_t$ . O conceito de passeio aleatório do tipo intervalo será formalizado na Seção 4.5.

No método AES para dados univalorados, a constante de alisamento  $\alpha$  pode ser estimada a partir dos dados, minimizando o critério da soma de quadrados dos erros de previsão,  $\sum_{t=2}^n (y_t - \widehat{y}_t)^2$ . No método AES<sup>I</sup>, usando  $\mathbf{I}_t$  em vez de  $y_t$  e alguma álgebra matricial, chegamos ao critério da soma de quadrados dos erros de previsão para STI,

$$\begin{aligned} \mathcal{S}_2(\mathcal{A}) &= \sum_{t=2}^n (\mathbf{I}_t - \widehat{\mathbf{I}}_t)^\top (\mathbf{I}_t - \widehat{\mathbf{I}}_t) \\ &= \sum_{t=2}^n \begin{bmatrix} X_t^U - \widehat{X}_t^U \\ X_t^L - \widehat{X}_t^L \end{bmatrix}^\top \begin{bmatrix} X_t^U - \widehat{X}_t^U \\ X_t^L - \widehat{X}_t^L \end{bmatrix} \\ &= \sum_{t=2}^n (X_t^U - \widehat{X}_t^U)^2 + \sum_{t=2}^n (X_t^L - \widehat{X}_t^L)^2. \end{aligned}$$

Este é o critério usado como função objetivo a ser minimizada no método AES<sup>I</sup> para estimação da matriz de parâmetros  $\mathcal{A}$ .

Usando as equações dadas em (4.11), reescrevemos a função objetivo para  $i = 1, 2$  e  $j = 1, 2$  como

$$\begin{aligned} \mathcal{S}_2(\alpha_{ij}) &= \sum_{t=2}^n \left( X_t^U - \alpha_{i1} X_{t-1}^U - (1 - \alpha_{i1}) \widehat{X}_{t-1}^U - \alpha_{i2} e_{t-1}^L \right)^2 \\ &\quad + \sum_{t=2}^n \left( X_t^L - \alpha_{i2} X_{t-1}^L - (1 - \alpha_{i2}) \widehat{X}_{t-1}^L - \alpha_{i1} e_{t-1}^U \right)^2. \end{aligned}$$

O índice 2 em  $\mathcal{S}_2(\alpha_{ij})$  deve-se ao fato de utilizarmos a norma  $l_2$  (mínimos quadrados) dos resíduos. Em contrapartida, podemos usar como função objetivo a norma  $l_1$  dos resíduos,  $\mathcal{S}_1(\mathcal{A}) = \sum_{t=2}^n |\mathbf{I}_t - \widehat{\mathbf{I}}_t|$ . Isto também é válido para os modelos que serão apresentados em seções seguintes.

A partir da função objetivo  $\mathcal{S}_2(\alpha_{ij})$ , podemos expressar formalmente a estimação

da matriz  $\mathcal{A}$  como um *problema de mínimos quadrados linear*<sup>4</sup> com restrições, especificamente,

$$\begin{aligned} \text{Minimizar} \quad & \mathcal{S}_2(\alpha_{ij}), \\ \text{sujeito a} \quad & 0 \leq \alpha_{ij} \leq 1. \end{aligned}$$

#### 4.4.2 Alisamento exponencial duplo para STI

O segundo método de alisamento exponencial estendido para séries temporais do tipo intervalo é o método de *alisamento exponencial duplo* (AED). Em dados univariados, este método é adotado quando a série exibe tendência. A diferença para o AES é que no AED duas componentes são atualizadas em cada instante de tempo: *nível e tendência*.

No AED supõe-se que o nível estimado no tempo  $t$ ,  $N_t$ , é a média ponderada do valor observado no tempo  $t$  e a soma dos valores estimados do nível e da tendência no tempo  $t - 1$  como segue

$$\widehat{N}_t = \alpha y_t + (1 - \alpha)(\widehat{N}_{t-1} + \widehat{T}_{t-1}). \quad (4.12)$$

A tendência é, então, estimada no tempo  $t$  como a média ponderada de  $\widehat{N}_t - \widehat{N}_{t-1}$  e a tendência estimada no tempo  $t - 1$ , ou seja,

$$\widehat{T}_t = \beta(\widehat{N}_t - \widehat{N}_{t-1}) + (1 - \beta)\widehat{T}_{t-1}. \quad (4.13)$$

Os parâmetros de alisamento  $\alpha$  e  $\beta$  têm suporte entre 0 e 1 e podem ser estimados como no método de alisamento exponencial simples, minimizando a soma de quadrados dos erros de previsão. Neste método, o parâmetro  $\alpha$  controla o alisamento para estimação do nível e o parâmetro  $\beta$  controla o alisamento da tendência.

Para resolver as equações (4.12) e (4.13) de forma recursiva faz-se necessário fornecer os valores iniciais para o nível e para a tendência. Uma escolha usual é fazer

$$\widehat{N}_2 = y_2 \quad \text{e} \quad \widehat{T}_2 = y_2 - y_1.$$

---

<sup>4</sup>Considere uma função de mínimos quadrados  $\sum_k [y_k - f(\mathbf{y}_k, \theta)]^2$  que soma os quadrados das discrepâncias entre os valores observados e os valores preditos pelo modelo. Esta função pode ser expressa da forma  $\sum_k e_k^2(\theta)$ , em que  $e_k(\theta) = y_k - f(\mathbf{y}_k, \theta)$ . Em problemas de modelagem de dados, se  $f(\mathbf{y}_k, \theta)$  é função linear de  $\theta$ , então, os resíduos  $e_k(\theta)$  também é linear e o problema de minimização é denominado *problema de mínimos quadrados linear* (ver Nocedal e Wright, 1999, Capítulo 10) [76].

Então, (4.12) e (4.13) podem ser resolvidas sucessivamente para  $\widehat{N}_t$  e  $\widehat{T}_t$ ,  $t = 2, \dots, n$  e as previsões são dadas por

$$\widehat{y}_{t+1} = \widehat{N}_t + \widehat{T}_t. \quad (4.14)$$

A extensão do método AED para STI, denominado método AED<sup>I</sup>, segue metodologia análoga à adotada na extensão do método AES. Seja  $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n$  uma STI, chamamos de  $\widehat{\mathbf{N}}_t^{\mathbf{I}}$  o nível intervalar estimado e de  $\widehat{\mathbf{T}}_t^{\mathbf{I}}$  a tendência intervalar estimada para esta STI. No método AED<sup>I</sup> estas componentes são representadas na forma matricial

$$\begin{aligned} \widehat{\mathbf{N}}_t^{\mathbf{I}} &= \mathcal{A}\mathbf{I}_t + (\mathbf{I} - \mathcal{A})(\widehat{\mathbf{N}}_{t-1}^{\mathbf{I}} + \widehat{\mathbf{T}}_{t-1}^{\mathbf{I}}), \\ \widehat{\mathbf{T}}_t^{\mathbf{I}} &= \mathcal{B}(\widehat{\mathbf{N}}_t^{\mathbf{I}} - \widehat{\mathbf{N}}_{t-1}^{\mathbf{I}}) + (\mathbf{I} - \mathcal{B})\widehat{\mathbf{T}}_{t-1}^{\mathbf{I}}, \end{aligned}$$

em que

$$\widehat{\mathbf{N}}_t^{\mathbf{I}} = \begin{bmatrix} \widehat{N}_t^U \\ \widehat{N}_t^L \end{bmatrix} \quad \text{e} \quad \widehat{\mathbf{T}}_t^{\mathbf{I}} = \begin{bmatrix} \widehat{T}_t^U \\ \widehat{T}_t^L \end{bmatrix}$$

e as matrizes de parâmetros de alisamento,  $\mathcal{A}$  e  $\mathcal{B}$ , são

$$\mathcal{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \quad \text{e} \quad \mathcal{B} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}.$$

$\mathbf{I}$  é uma matriz  $(2 \times 2)$  identidade. Assim, reescrevemos  $\widehat{\mathbf{N}}_t^{\mathbf{I}}$  e  $\widehat{\mathbf{T}}_t^{\mathbf{I}}$  através das equações

$$\begin{aligned} \widehat{\mathbf{N}}_t^{\mathbf{I}} &= \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} + \begin{bmatrix} 1 - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & 1 - \alpha_{22} \end{bmatrix} \begin{bmatrix} \widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U \\ \widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{11}X_t^U + \alpha_{12}X_t^L \\ \alpha_{22}X_t^L + \alpha_{21}X_t^U \end{bmatrix} + \begin{bmatrix} (1 - \alpha_{11})(\widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U) - \alpha_{12}(\widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L) \\ (1 - \alpha_{22})(\widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L) - \alpha_{21}(\widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U) \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{11}X_t^U + (1 - \alpha_{11})(\widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U) + \alpha_{12}(X_t^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L) \\ \alpha_{22}X_t^L + (1 - \alpha_{22})(\widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L) + \alpha_{21}(X_t^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U) \end{bmatrix} \end{aligned} \quad (4.15)$$

e

$$\widehat{\mathbf{T}}_t^{\mathbf{I}} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{N}_t^U - \widehat{N}_{t-1}^U \\ \widehat{N}_t^L - \widehat{N}_{t-1}^L \end{bmatrix} + \begin{bmatrix} 1 - \beta_{11} & -\beta_{12} \\ -\beta_{21} & 1 - \beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{T}_{t-1}^U \\ \widehat{T}_{t-1}^L \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} \beta_{11}(\widehat{N}_t^U - \widehat{N}_{t-1}^U) + \beta_{12}(\widehat{N}_t^L - \widehat{N}_{t-1}^L) \\ \beta_{22}(\widehat{N}_t^L - \widehat{N}_{t-1}^L) + \beta_{21}(\widehat{N}_t^U - \widehat{N}_{t-1}^U) \end{bmatrix} + \begin{bmatrix} (1 - \beta_{11})\widehat{T}_{t-1}^U - \beta_{12}\widehat{T}_{t-1}^L \\ (1 - \beta_{22})\widehat{T}_{t-1}^L - \beta_{21}\widehat{T}_{t-1}^U \end{bmatrix} \\
&= \begin{bmatrix} \beta_{11}(\widehat{N}_t^U - \widehat{N}_{t-1}^U) + (1 - \beta_{11})\widehat{T}_{t-1}^U + \beta_{12}(\widehat{N}_t^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L) \\ \beta_{22}(\widehat{N}_t^L - \widehat{N}_{t-1}^L) + (1 - \beta_{22})\widehat{T}_{t-1}^L + \beta_{21}(\widehat{N}_t^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U) \end{bmatrix}. \quad (4.16)
\end{aligned}$$

■

No método AED<sup>I</sup>, a interpretação das matrizes de parâmetros de alisamento  $\mathcal{A}$  e  $\mathcal{B}$  é similar à interpretação dos parâmetros escalares do método AED para dados univariados. As matrizes  $\mathcal{A}$  e  $\mathcal{B}$  regulam o grau de alisamento do método. Se  $\mathcal{A}$  é uma matriz identidade (i.e.,  $\alpha_{11} = \alpha_{22} = 1$  e  $\alpha_{12} = \alpha_{21} = 0$ ), as curvas ajustadas não apresentam qualquer alisamento sobre o nível, pois  $\widehat{\mathbf{N}}_t^{\mathbf{I}} = \mathbf{I} \cdot \mathbf{I}_t = \mathbf{I}_t$ . Por outro lado, se  $\mathcal{A}$  é uma matriz nula (i.e.,  $\alpha_{11} = \alpha_{22} = \alpha_{12} = \alpha_{21} = 0$ ), as curvas ajustadas sobre o nível são completamente alisadas, porém não apresentam variação para esta componente. A matriz de parâmetros  $\mathcal{B}$  controla o alisamento da tendência. Além disso, os parâmetros  $\alpha_{12}$  e  $\beta_{12}$  são responsáveis pela introdução de informações relativas ao limite inferior no ajuste do limite superior, enquanto que os parâmetros  $\alpha_{21}$  e  $\beta_{21}$  introduzem informações do limite superior no ajuste do limite inferior.

A partir das equações (4.15) e (4.16), realizamos as previsões de intervalos da forma

$$\widehat{\mathbf{I}}_{t+1} = \widehat{\mathbf{N}}_t^{\mathbf{I}} + \widehat{\mathbf{T}}_t^{\mathbf{I}}. \quad (4.17)$$

A estimação das matrizes de parâmetros de alisamento propostas do método AED para série do tipo intervalo,  $\mathcal{A}$  e  $\mathcal{B}$ , é realizada analogamente à estimação no método AES<sup>I</sup>, através da minimização de um critério. No método AED<sup>I</sup> a função objetivo é dada por

$$\begin{aligned}
\mathcal{R}_2(\mathcal{A}, \mathcal{B}) &= \sum_{t=3}^n (\mathbf{I}_t - \widehat{\mathbf{I}}_t)^\top (\mathbf{I}_t - \widehat{\mathbf{I}}_t) \\
&= \sum_{t=3}^n \begin{bmatrix} X_t^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U \\ X_t^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L \end{bmatrix}^\top \begin{bmatrix} X_t^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U \\ X_t^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L \end{bmatrix} \\
&= \sum_{t=3}^n (X_t^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U)^2 + \sum_{t=3}^n (X_t^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L)^2.
\end{aligned}$$

Usando as expressões (4.15), (4.16) e (4.17) em  $\mathcal{R}_2(\mathcal{A}, \mathcal{B})$ , temos

$$\begin{aligned} \mathcal{R}_2(\alpha_{ij}, \beta_{ij}) &= \sum_{t=3}^n \left( X_t^U - \alpha_{11} X_{t-1}^U - (1 - \alpha_{11})(\widehat{N}_{t-2}^U + \widehat{T}_{t-2}^U) - \alpha_{12}(X_{t-1}^L - \widehat{N}_{t-2}^L - \widehat{T}_{t-2}^L) \right. \\ &\quad \left. - \beta_{11}(\widehat{N}_{t-1}^U - \widehat{N}_{t-2}^U) - (1 - \beta_{11})\widehat{T}_{t-2}^U - \beta_{12}(\widehat{N}_{t-1}^L - \widehat{N}_{t-2}^L - \widehat{T}_{t-2}^L) \right)^2 \\ &+ \sum_{t=3}^n \left( X_t^L - \alpha_{22} X_{t-1}^L - (1 - \alpha_{22})(\widehat{N}_{t-2}^L + \widehat{T}_{t-2}^L) - \alpha_{21}(X_{t-1}^U - \widehat{N}_{t-2}^U - \widehat{T}_{t-2}^U) \right. \\ &\quad \left. - \beta_{22}(\widehat{N}_{t-1}^L - \widehat{N}_{t-2}^L) - (1 - \beta_{22})\widehat{T}_{t-2}^L - \beta_{21}(\widehat{N}_{t-1}^U - \widehat{N}_{t-2}^U - \widehat{T}_{t-2}^U) \right)^2, \end{aligned}$$

com  $i = 1, 2$  e  $j = 1, 2$ . Então, a estimação de  $\mathcal{A}$  e  $\mathcal{B}$  resume-se ao problema de mínimos quadrados com restrições

$$\begin{aligned} \text{Minimizar} \quad & \mathcal{R}_2(\alpha_{ij}, \beta_{ij}), \\ \text{sujeito a} \quad & 0 \leq \alpha_{ij}, \beta_{ij} \leq 1. \end{aligned}$$

#### 4.4.3 Método Holt–Winters para STI

Os métodos de alisamento exponencial podem ser generalizados de forma a tratar variações sazonais nos dados. Esta generalização foi apresentada originalmente em Holt (1959, publicado novamente em 2004) [53] e Winters (1960) [92] e é denominada de método *Holt–Winters* (HW). Algumas referências designam este método como método de *alisamento exponencial triplo*, contudo, nós adotamos a primeira denominação.

A maneira assumida no método Holt–Winters para tratar a sazonalidade presente na série temporal foi efetuar a inclusão de um novo índice às fórmulas de atualização adotadas no método AED. Como no método AED, suponha que  $N_t$  e  $T_t$  denotem, respectivamente, o nível e a tendência no tempo  $t$ . No método Holt–Winters o fator incluído  $S_t$  representa o índice sazonal no tempo  $t$ . Seja, ainda,  $s$  o fator sazonal, i.e., o número de períodos considerado no ciclo sazonal da série (e.g., durante o ano,  $s = 12$  para dados mensais e  $s = 4$  para dados trimestrais). Suponha, também, que no instante de tempo  $t$ , nós conhecemos os termos  $(N_1, T_1, S_1), \dots, (N_{t-1}, T_{t-1}, S_{t-1})$ . Então, após observar  $y_t$ , os termos  $N_t$ ,  $T_t$  e  $S_t$  são atualizados via método de alisamento exponencial.

Em dados clássicos, na presença de sazonalidade, as equações recorrentes para

atualização do nível,  $N_t$ , da tendência,  $T_t$ , e do índice sazonal,  $S_t$ , são, respectivamente,

$$\begin{aligned}\widehat{N}_t &= \alpha(y_t - \widehat{S}_{t-s}) + (1 - \alpha)(\widehat{N}_{t-1} + \widehat{T}_{t-1}), \\ \widehat{T}_t &= \beta(\widehat{N}_t - \widehat{N}_{t-1}) + (1 - \beta)\widehat{T}_{t-1}, \\ \widehat{S}_t &= \delta(y_t - \widehat{N}_t) + (1 - \delta)\widehat{S}_{t-s}.\end{aligned}$$

Os coeficientes  $\alpha$ ,  $\beta$  e  $\delta$  são denominados parâmetros de alisamento. A previsão um-passo-à-frente é dada por

$$\widehat{y}_{t+1} = \widehat{N}_t + \widehat{T}_t + \widehat{S}_{t-s+1}.$$

Como condições iniciais, podemos considerar

$$\begin{aligned}\widehat{N}_s &= y_s, \\ \widehat{T}_s &= (y_s - y_1)/s, \\ \widehat{S}_i &= y_i - (y_1 + \widehat{T}_s(i - 1)), \quad i = 1, \dots, d,\end{aligned}$$

definidas em Brockwell e Davis (2002, Capítulo 9) [14].

No método Holt–Winters cada uma das constantes de alisamento,  $\alpha$ ,  $\beta$  e  $\delta$ , está associada a uma componente da série e, como nos métodos de alisamento exponencial anteriores, essas constantes têm suporte entre 0 e 1.

Adotando a metodologia para extensão dos métodos clássicos, nós apresentamos o método Holt–Winters estendido para STI ( $\text{HW}^I$ ). Denotamos  $\mathbf{N}_t^I$ ,  $\mathbf{T}_t^I$  e  $\mathbf{S}_t^I$ , respectivamente, por nível intervalar, tendência intervalar e sazonalidade intervalar no tempo  $t$ .

Com base no método Holt–Winters clássico, as equações propostas no método  $\text{HW}^I$  para atualização do nível intervalar, da tendência intervalar e da sazonalidade intervalar, são, respectivamente,

$$\begin{aligned}\widehat{\mathbf{N}}_t^I &= \mathcal{A}(\mathbf{I}_t - \widehat{\mathbf{S}}_{t-s}^I) + (\mathbf{I} - \mathcal{A})(\widehat{\mathbf{N}}_{t-1}^I + \widehat{\mathbf{T}}_{t-1}^I), \\ \widehat{\mathbf{T}}_t^I &= \mathcal{B}(\widehat{\mathbf{N}}_t^I - \widehat{\mathbf{N}}_{t-1}^I) + (\mathbf{I} - \mathcal{B})\widehat{\mathbf{T}}_{t-1}^I, \\ \widehat{\mathbf{S}}_t^I &= \mathcal{D}(\mathbf{I}_t - \widehat{\mathbf{N}}_t^I) + (\mathbf{I} - \mathcal{D})\widehat{\mathbf{S}}_{t-s}^I,\end{aligned}$$

em que

$$\widehat{\mathbf{N}}_t^I = \begin{bmatrix} \widehat{N}_t^U \\ \widehat{N}_t^L \end{bmatrix}, \quad \widehat{\mathbf{T}}_t^I = \begin{bmatrix} \widehat{T}_t^U \\ \widehat{T}_t^L \end{bmatrix} \quad \text{e} \quad \widehat{\mathbf{S}}_{t-s}^I = \begin{bmatrix} \widehat{S}_{t-s}^U \\ \widehat{S}_{t-s}^L \end{bmatrix},$$

e as matrizes ( $2 \times 2$ ) de constantes de alisamento,  $\mathcal{A}$ ,  $\mathcal{B}$  e  $\mathcal{D}$ , são

$$\mathcal{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \quad \text{e} \quad \mathcal{D} = \begin{bmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{bmatrix},$$

com  $\mathbf{I}$  uma matriz identidade de ordem 2.

Após expansão matricial do método HW<sup>I</sup>, as equações de atualização do nível intervalar, da tendência intervalar e da sazonalidade intervalar podem, então, ser reescritas da forma

$$\begin{aligned} \widehat{\mathbf{N}}_t^{\mathbf{I}} &= \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} X_t^U - \widehat{S}_{t-s}^U \\ X_t^L - \widehat{S}_{t-s}^L \end{bmatrix} + \begin{bmatrix} 1 - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & 1 - \alpha_{22} \end{bmatrix} \begin{bmatrix} \widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U \\ \widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{11}(X_t^U - \widehat{S}_{t-s}^U) + \alpha_{12}(X_t^L - \widehat{S}_{t-s}^L) \\ \alpha_{22}(X_t^L - \widehat{S}_{t-s}^L) + \alpha_{21}(X_t^U - \widehat{S}_{t-s}^U) \end{bmatrix} \\ &\quad + \begin{bmatrix} (1 - \alpha_{11})(\widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U) - \alpha_{12}(\widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L) \\ (1 - \alpha_{22})(\widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L) - \alpha_{21}(\widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U) \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{11}(X_t^U - \widehat{S}_{t-s}^U) + (1 - \alpha_{11})(\widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U) + \alpha_{12}(X_t^L - \widehat{S}_{t-s}^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L) \\ \alpha_{22}(X_t^L - \widehat{S}_{t-s}^L) + (1 - \alpha_{22})(\widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L) + \alpha_{21}(X_t^U - \widehat{S}_{t-s}^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U) \end{bmatrix}, \end{aligned} \quad (4.18)$$

$$\begin{aligned} \widehat{\mathbf{T}}_t^{\mathbf{I}} &= \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{N}_t^U - \widehat{N}_{t-1}^U \\ \widehat{N}_t^L - \widehat{N}_{t-1}^L \end{bmatrix} + \begin{bmatrix} 1 - \beta_{11} & -\beta_{12} \\ -\beta_{21} & 1 - \beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{T}_{t-1}^U \\ \widehat{T}_{t-1}^L \end{bmatrix} \\ &= \begin{bmatrix} \beta_{11}(\widehat{N}_t^U - \widehat{N}_{t-1}^U) + \beta_{12}(\widehat{N}_t^L - \widehat{N}_{t-1}^L) \\ \beta_{22}(\widehat{N}_t^L - \widehat{N}_{t-1}^L) + \beta_{21}(\widehat{N}_t^U - \widehat{N}_{t-1}^U) \end{bmatrix} + \begin{bmatrix} (1 - \beta_{11})\widehat{T}_{t-1}^U - \beta_{12}\widehat{T}_{t-1}^L \\ (1 - \beta_{22})\widehat{T}_{t-1}^L - \beta_{21}\widehat{T}_{t-1}^U \end{bmatrix} \\ &= \begin{bmatrix} \beta_{11}(\widehat{N}_t^U - \widehat{N}_{t-1}^U) + (1 - \beta_{11})\widehat{T}_{t-1}^U + \beta_{12}(\widehat{N}_t^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L) \\ \beta_{22}(\widehat{N}_t^L - \widehat{N}_{t-1}^L) + (1 - \beta_{22})\widehat{T}_{t-1}^L + \beta_{21}(\widehat{N}_t^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U) \end{bmatrix}, \end{aligned} \quad (4.19)$$

$$\begin{aligned} \widehat{\mathbf{S}}_t^{\mathbf{I}} &= \begin{bmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{bmatrix} \begin{bmatrix} X_t^U - \widehat{N}_t^U \\ X_t^L - \widehat{N}_t^L \end{bmatrix} + \begin{bmatrix} 1 - \delta_{11} & -\delta_{12} \\ -\delta_{21} & 1 - \delta_{22} \end{bmatrix} \begin{bmatrix} \widehat{S}_{t-s}^U \\ \widehat{S}_{t-s}^L \end{bmatrix} \\ &= \begin{bmatrix} \delta_{11}(X_t^U - \widehat{N}_t^U) + \delta_{12}(X_t^L - \widehat{N}_t^L) \\ \delta_{22}(X_t^L - \widehat{N}_t^L) + \delta_{21}(X_t^U - \widehat{N}_t^U) \end{bmatrix} + \begin{bmatrix} (1 - \delta_{11})\widehat{S}_{t-s}^U - \delta_{12}\widehat{S}_{t-s}^L \\ (1 - \delta_{22})\widehat{S}_{t-s}^L - \delta_{21}\widehat{S}_{t-s}^U \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} \delta_{11}(X_t^U - \widehat{N}_t^U) + (1 - \delta_{11})\widehat{S}_{t-s}^U + \delta_{12}(X_t^L - \widehat{N}_t^L - \widehat{S}_{t-s}^L) \\ \delta_{22}(X_t^L - \widehat{N}_t^L) + (1 - \delta_{22})\widehat{S}_{t-s}^L + \delta_{21}(X_t^U - \widehat{N}_t^U - \widehat{S}_{t-s}^U) \end{bmatrix}. \quad (4.20)$$

■

No método HW<sup>I</sup>, a previsão um-passo-à-frente é obtida da forma

$$\widehat{\mathbf{I}}_{t+1} = \widehat{\mathbf{N}}_t^I + \widehat{\mathbf{T}}_t^I + \widehat{\mathbf{S}}_{t-s+1}^I,$$

em que  $s$  é o número de períodos no ciclo sazonal da STI.

Os parâmetros do modelo HW<sup>I</sup> podem ser estimados como no método de alisamento exponencial, minimizando a soma de quadrados dos erros de previsão um-passo-à-frente dada por

$$\begin{aligned} \mathcal{T}_2(\mathcal{A}, \mathcal{G}, \mathcal{D}) &= \sum_{t=s}^n (\mathbf{I}_{t+1} - \widehat{\mathbf{I}}_{t+1})^\top (\mathbf{I}_{t+1} - \widehat{\mathbf{I}}_{t+1}) \\ &= \sum_{t=s}^n \begin{bmatrix} X_{t+1}^U - \widehat{N}_t^U - \widehat{T}_t^U - \widehat{S}_{t-s+1}^U \\ X_{t+1}^L - \widehat{N}_t^L - \widehat{T}_t^L - \widehat{S}_{t-s+1}^L \end{bmatrix}^\top \begin{bmatrix} X_{t+1}^U - \widehat{N}_t^U - \widehat{T}_t^U - \widehat{S}_{t-s+1}^U \\ X_{t+1}^L - \widehat{N}_t^L - \widehat{T}_t^L - \widehat{S}_{t-s+1}^L \end{bmatrix} \\ &= \sum_{t=s}^n (X_{t+1}^U - \widehat{N}_t^U - \widehat{T}_t^U - \widehat{S}_{t-s+1}^U)^2 + \sum_{t=s}^n (X_{t+1}^L - \widehat{N}_t^L - \widehat{T}_t^L - \widehat{S}_{t-s+1}^L)^2. \end{aligned}$$

Conforme as equações (4.18), (4.19) e (4.20), a função objetivo do método HW<sup>I</sup> pode ainda ser expressa como

$$\begin{aligned} \mathcal{T}_2(\alpha_{ij}, \beta_{ij}, \delta_{ij}) &= \sum_{t=s}^n \left( X_{t+1}^U - \alpha_{11}(X_t^U - \widehat{S}_{t-s}^U) - (1 - \alpha_{11})(\widehat{N}_{t-1}^U + \widehat{T}_{t-1}^U) \right. \\ &\quad - \alpha_{12}(X_t^L - \widehat{S}_{t-s}^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L) - \beta_{11}(\widehat{N}_t^U - \widehat{N}_{t-1}^U) - (1 - \beta_{11})\widehat{T}_{t-1}^U \\ &\quad - \beta_{12}(\widehat{N}_t^L - \widehat{N}_{t-1}^L - \widehat{T}_{t-1}^L) - \delta_{11}(X_{t-s+1}^U - \widehat{N}_{t-s+1}^U) \\ &\quad \left. - (1 - \delta_{11})\widehat{S}_{t-2s+1}^U - \delta_{12}(X_{t-s+1}^L - \widehat{N}_{t-s+1}^L - \widehat{S}_{t-2s+1}^L) \right)^2 \\ &\quad + \sum_{t=s}^n \left( X_{t+1}^L - \alpha_{22}(X_t^L - \widehat{S}_{t-s}^L) - (1 - \alpha_{22})(\widehat{N}_{t-1}^L + \widehat{T}_{t-1}^L) \right. \\ &\quad - \alpha_{21}(X_t^U - \widehat{S}_{t-s}^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U) - \beta_{22}(\widehat{N}_t^L - \widehat{N}_{t-1}^L) - (1 - \beta_{22})\widehat{T}_{t-1}^L \\ &\quad - \beta_{21}(\widehat{N}_t^U - \widehat{N}_{t-1}^U - \widehat{T}_{t-1}^U) - \delta_{22}(X_{t-s+1}^L - \widehat{N}_{t-s+1}^L) \\ &\quad \left. - (1 - \delta_{22})\widehat{S}_{t-2s+1}^L - \delta_{21}(X_{t-s+1}^U - \widehat{N}_{t-s+1}^U - \widehat{S}_{t-2s+1}^U) \right)^2, \end{aligned}$$

para  $i = 1, 2$  e  $j = 1, 2$ . Assim, estimar as matrizes de parâmetros  $\mathcal{A}$ ,  $\mathcal{B}$  e  $\mathcal{D}$  do método HW<sup>I</sup>, equivale a resolver o problema de mínimos quadrados com restrições

$$\begin{aligned} \text{Minimizar} \quad & \mathcal{T}_2(\alpha_{ij}, \beta_{ij}, \delta_{ij}), \\ \text{sujeito a} \quad & 0 \leq \alpha_{ij}, \beta_{ij}, \delta_{ij} \leq 1. \end{aligned}$$

### Estimação das matrizes de constantes de alisamento

Muitos problemas de estimação de parâmetros na modelagem estatística de dados podem ser representados ou aproximados como problema de programação numérica com ou sem restrições.

O problema de estimação das matrizes de constantes de alisamento dos métodos de alisamento exponencial estendidos para STI – o AES<sup>I</sup>, o AED<sup>I</sup> e o HW<sup>I</sup> – é solucionado via técnicas de programação numérica. Especificamente, consideramos o problema em que as variáveis são limitadas, ou seja, o problema modelado por

$$\begin{aligned} \text{Minimizar} \quad & f(x), \\ \text{sujeito a} \quad & l \leq x \leq u, \end{aligned}$$

onde  $f$  é uma função de  $\mathbb{R}^m$  em  $\mathbb{R}$ ;  $l$  e  $u$  representam, respectivamente, os valores de limitação inferior e superior das variáveis (parâmetros) e  $m$  é o número de parâmetros. As restrições consideradas em nossos problemas são funções lineares.

Para minimizar as três funções objetivo  $\mathcal{S}_2(\alpha_{ij})$ ,  $\mathcal{R}_2(\alpha_{ij}, \beta_{ij})$  e  $\mathcal{T}_2(\alpha_{ij}, \beta_{ij}, \delta_{ij})$  dos métodos AES<sup>I</sup>, AED<sup>I</sup> e HW<sup>I</sup>, respectivamente, nós aplicamos o método BFGS *de memória limitada com restrições em caixa* (L-BFGS-B)<sup>5</sup> desenvolvido por Byrd et al. (1995) [18]. O algoritmo L-BFGS-B é uma extensão da versão BFGS *de memória limitada* (L-BFGS) usada para problemas sem restrições (Liu e Nocedal, 1989) [65] que, por sua vez, se baseia na técnica de projeção de gradiente através do usual método de otimização *quasi-Newton* conhecido como BFGS desenvolvido simultaneamente em 1970 em trabalhos individuais de Broyden [16], Fletcher [38], Goldfarb [45] e Shanno [83]. A ideia principal do algoritmo original BFGS é permitir a construção aproximada da matriz Hessiana da função a ser minimizada. Desta forma, a implementação do gradiente da função pode ser desprezada.

---

<sup>5</sup>Sigla em inglês que significa *limited memory BFGS with bound constraints*.

O método L-BFGS-B tem se mostrado eficaz na solução de problemas de otimização de grandes dimensões, com ou sem restrições. O termo “memória limitada” deve-se ao fato de que uma aproximação de baixo ranque da inversa da matriz Hessiana é armazenada em vez da inversa da matriz Hessiana completa. A aplicação deste método na estimação das matrizes de constantes de alisamento deve-se, principalmente, ao fato de que no L-BFGS-B são permitidas restrições em caixa, isto é, restrições em que os parâmetros são limitados inferior e/ou superiormente.

Basicamente, o método L-BFGS-B consiste na minimização de uma função custo  $f(x)$  de  $m$  variáveis restritas a uma limitação em caixa do  $\mathbb{R}^m$ . Como se baseia na técnica de projeção de gradiente, faz-se necessário determinar o valor da função  $f$  e de seu gradiente  $g$  avaliados em cada passo do algoritmo. Por esta razão, o L-BFGS-B torna-se útil para resolver grandes problemas em que o cálculo da matriz Hessiana e/ou sua inversa é de difícil ou impossível obtenção. O algoritmo L-BFGS-B está implementado na linguagem de programação R (Chambers, 2005) [20]. Para descrição mais detalhada sobre este algoritmo, ver, e.g., Nocedal e Wright (1999) [76].

A seguir apresentamos exemplos usando os métodos AES<sup>I</sup>, AED<sup>I</sup> e HW<sup>I</sup>, em que as matrizes de constantes de alisamento são estimadas via algoritmo L-BFGS-B.

**Exemplo 4.3** (Aplicação dos métodos AES<sup>I</sup> e AED<sup>I</sup>). Neste exemplo, nós ajustamos através dos métodos AES<sup>I</sup> e AED<sup>I</sup> a STI dos preços diários de negociações de ações da Petrobras descrita na Seção 4.2.1, Figura 4.6. Este exemplo tem por finalidade apenas ilustrar a aplicação destes métodos de alisamento exponencial, por isso, não avaliamos as precisões das previsões e não interpretamos as equações ajustadas.

A matriz de constantes de alisamento estimada no método AES<sup>I</sup> para a STI de preços de ações da Petrobras foi  $\widehat{\mathcal{A}} = \begin{bmatrix} 0,648 & 0,640 \\ 0,602 & 0,621 \end{bmatrix}$ , que resulta na seguinte equação de atualização do método AES<sup>I</sup> ajustada:

$$\widehat{\mathbf{I}}_{t+1} = \begin{bmatrix} 0,648X_t^U + 0,352\widehat{X}_t^U + 0,640e_t^L \\ 0,621X_t^L + 0,379\widehat{X}_t^L + 0,602e_t^U \end{bmatrix}.$$

A Figura 4.9 apresenta a STI com os preços reais (segmentos de reta pretos) e a STI ajustada pelo método AES<sup>I</sup> (segmentos de reta cinza).

Nós ajustamos novamente a STI de preços de ações da Petrobras através do método AED<sup>I</sup>. Para este método, as matrizes de constantes de alisamento estimadas

são  $\hat{\mathcal{A}} = \begin{bmatrix} 1,000 & 0,071 \\ 0,148 & 1,000 \end{bmatrix}$  e  $\hat{\mathcal{B}} = \begin{bmatrix} 0,203 & 0,000 \\ 0,039 & 0,000 \end{bmatrix}$ . Desta forma, o vetor de equações ajustado pelo método AED<sup>I</sup> para previsão um-passo-à-frente é

$$\hat{\mathbf{I}}_{t+1} = \begin{bmatrix} X_t^U + 0,071(X_t^L - \hat{N}_{t-1}^L - \hat{T}_{t-1}^L) \\ X_t^L + 0,148(X_t^U - \hat{N}_{t-1}^U - \hat{T}_{t-1}^U) \end{bmatrix} + \begin{bmatrix} 0,203(\hat{N}_t^U - \hat{N}_{t-1}^U) + 0,797\hat{T}_{t-1}^U \\ \hat{T}_{t-1}^L + 0,039(\hat{N}_t^U - \hat{N}_{t-1}^U - \hat{T}_{t-1}^U) \end{bmatrix}.$$

A Figura 4.10 mostra a STI dos preços reais (segmentos de reta pretos) de ações da Petrobras e a correspondente STI ajustada via AED<sup>I</sup> (segmentos de reta cinza).

**Exemplo 4.4** (Aplicação do método HW<sup>I</sup>). O método Holt–Winters, como descrito anteriormente, foi proposto com a finalidade de tratar variações sazonais nos dados. Por isso, nós ilustramos o método HW<sup>I</sup> no ajuste da STI de temperaturas da China, visto que esta série tem característica sazonal; ver Figura 4.5 na Seção 4.2.1. Como no exemplo anterior, não avaliamos as precisões das previsões e não interpretamos as equações ajustadas.

No exemplo da aplicação do método HW<sup>I</sup> no ajuste da STI de temperaturas medidas em uma estação meteorológica da China, as matrizes de constantes de alisamento estimadas são  $\hat{\mathcal{A}} = \begin{bmatrix} 0,000 & 0,016 \\ 0,000 & 0,348 \end{bmatrix}$ ,  $\hat{\mathcal{B}} = \begin{bmatrix} 0,000 & 0,000 \\ 0,941 & 0,023 \end{bmatrix}$  e  $\hat{\mathcal{D}} = \begin{bmatrix} 0,287 & 0,000 \\ 0,000 & 0,630 \end{bmatrix}$ . Desta forma, o vetor de equações ajustado pelo método HW<sup>I</sup> para previsão um-passo-à-frente é

$$\hat{\mathbf{I}}_{t+1} = \begin{bmatrix} (\hat{N}_{t-1}^U + \hat{T}_{t-1}^U) + 0,016(X_t^L - \hat{S}_{t-s}^L - \hat{N}_{t-1}^L - \hat{T}_{t-1}^L) \\ 0,348(X_t^L - \hat{S}_{t-s}^L) + 0,652(\hat{N}_{t-1}^L + \hat{T}_{t-1}^L) \end{bmatrix} + \begin{bmatrix} \hat{T}_{t-1}^U \\ 0,023(\hat{N}_t^L - \hat{N}_{t-1}^L) + 0,977\hat{T}_{t-1}^L + 0,941(\hat{N}_t^U - \hat{N}_{t-1}^U - \hat{T}_{t-1}^U) \end{bmatrix} + \begin{bmatrix} 0,287(X_t^U - \hat{N}_t^U) + 0,713\hat{S}_{t-s+1}^U \\ 0,630(X_t^L - \hat{N}_t^L) + 0,370\hat{S}_{t-s+1}^L \end{bmatrix}.$$

A Figura 4.11 mostra a STI de temperaturas reais (segmentos de reta pretos) medidas na China e a correspondente STI ajustada via HW<sup>I</sup> (segmentos de reta cinza).

Figura 4.9: Ajuste da STI de preços de ações da Petrobras via método AES<sup>I</sup>.

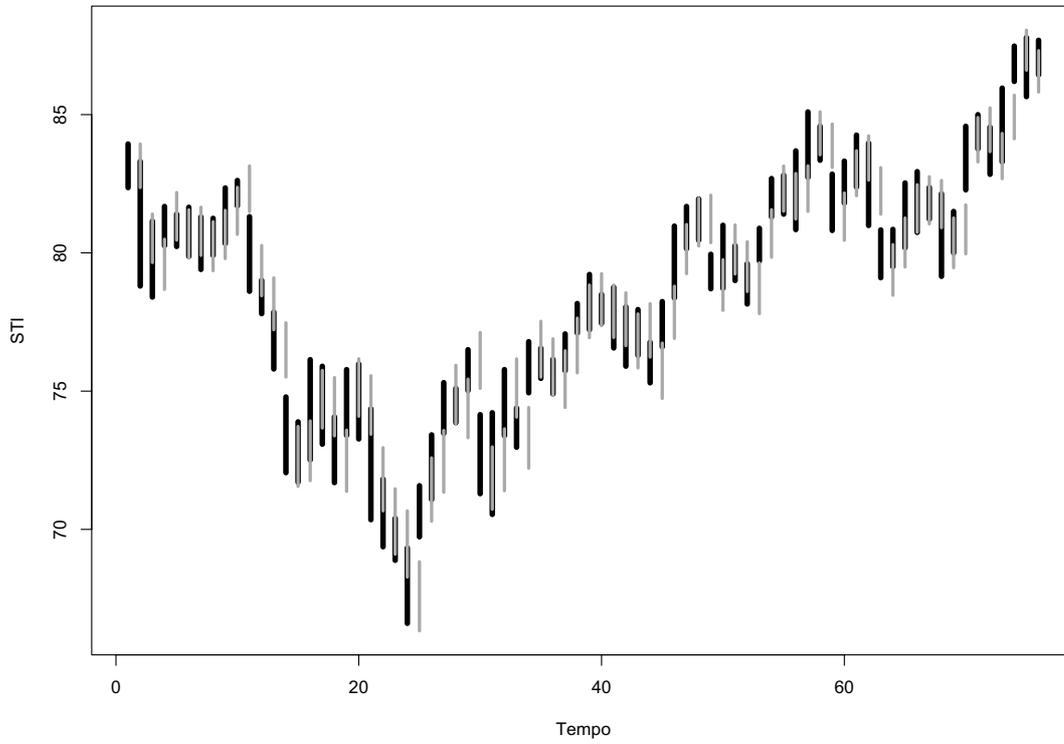


Figura 4.10: Ajuste da STI de preços de ações da Petrobras via método AED<sup>I</sup>.

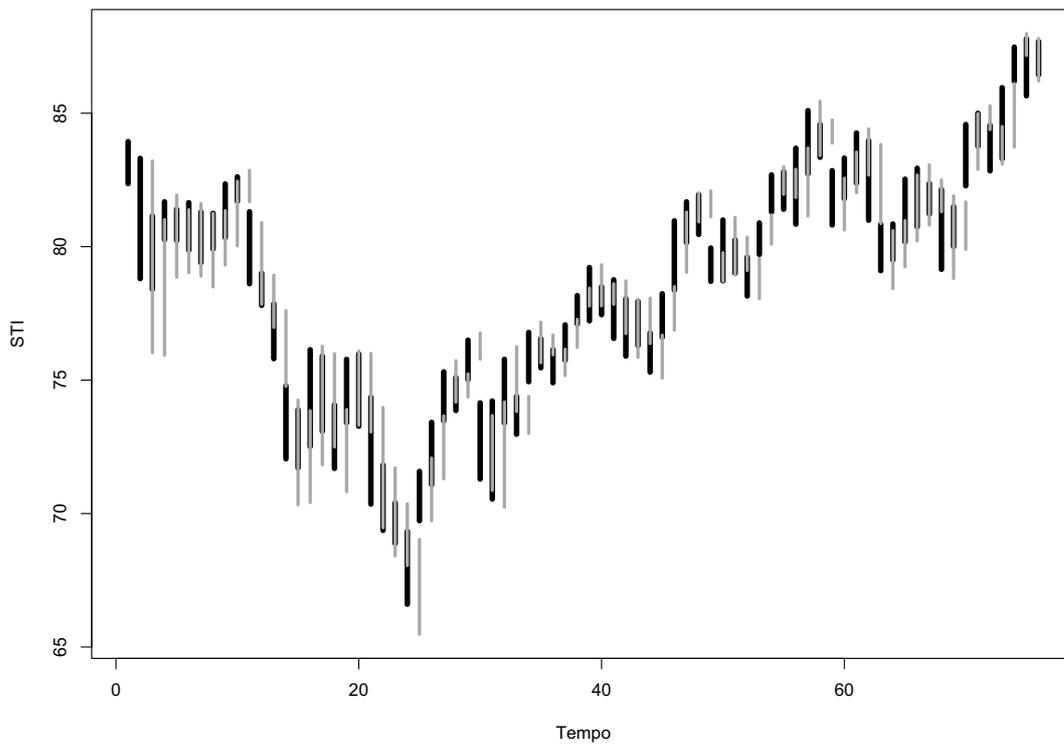
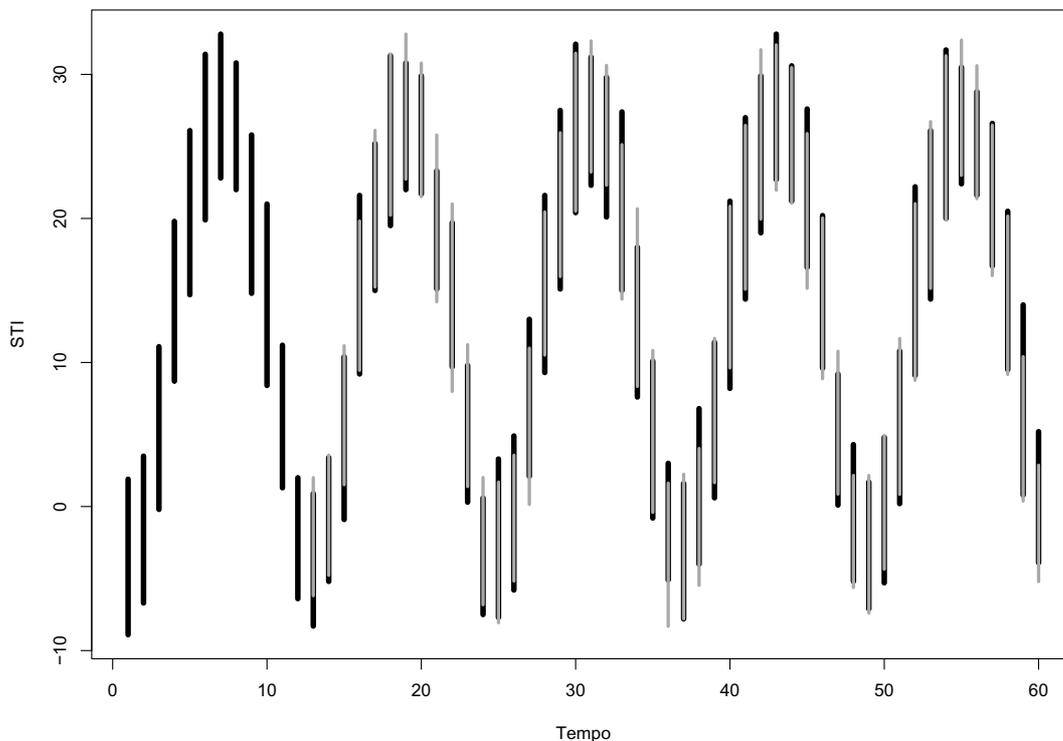


Figura 4.11: Ajuste da STI de temperaturas na China via método HW<sup>I</sup>.

Nesta seção foram apresentados três métodos para previsão de STI inspirados em métodos de alisamento exponencial usados para previsão de séries temporais univariadas. Denominamos os três métodos propostos por AES<sup>I</sup>, AED<sup>I</sup> e HW<sup>I</sup>. Nestes três métodos, as matrizes  $(2 \times 2)$  de constantes de alisamento são estimadas via técnicas de programação numérica empregadas para resolver problemas de mínimos quadrados lineares com restrições em caixa. As funções objetivo adotadas são a soma de quadrados dos erros de previsão intervalar de modo que os erros da previsão do limite superior e da previsão do limite inferior são conjuntamente computados.

## 4.5 Modelos para STI

Nesta seção estão apresentados, inicialmente, alguns processos estocásticos<sup>6</sup> que são utilizados com frequência na especificação de modelos para séries temporais univariadas e, em seguida, os correspondentes processos definidos para STI, incluindo alguns pressupostos considerados.

Basicamente, o nosso problema se restringe a estimar os parâmetros dos modelos estendidos para STI utilizando técnicas usuais, e.g., minimização de critério estabelecido na busca de um bom modelo para previsão de intervalos futuros da STI. Serão apresentados, então, modelos para STI inspirados em modelos convencionais conforme metodologia adotada e descrita na Seção 4.1.

### Processo intervalar aleatório

Em análise de séries temporais univariadas, um processo é dito ser *puramente aleatório* se consiste de uma sequência de variáveis aleatórias  $y_t$  independentes e identicamente distribuídas com média e variância constantes e covariância

$$c_k = \text{Cov}(y_t, y_{t+k}) = 0,$$

para  $k \neq 0$ .

Com base nesta ideia, dizemos que uma série temporal intervalar  $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n$  é um processo intervalar puramente aleatório se tem média e variância constantes dadas pelos vetores bidimensionais

$$\mathbf{E}(\mathbf{I}_t) = \boldsymbol{\mu}^I = \begin{bmatrix} \mathbf{E}(X_t^U) \\ \mathbf{E}(X_t^L) \end{bmatrix} = \begin{bmatrix} \mu^U \\ \mu^L \end{bmatrix} \quad \text{e} \quad \text{Var}(\mathbf{I}_t) = \boldsymbol{\sigma}^{I^2} = \begin{bmatrix} \text{Var}(X_t^U) \\ \text{Var}(X_t^L) \end{bmatrix} = \begin{bmatrix} \sigma^{U^2} \\ \sigma^{L^2} \end{bmatrix},$$

ou seja, se os limites superior e inferior são processos puramente aleatórios independentes e, ainda, a covariância intervalar, com  $k \neq 0$ , é dada pelo vetor

$$\begin{bmatrix} \text{Cov}(X_t^U, X_{t+k}^U) \\ \text{Cov}(X_t^L, X_{t+k}^L) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

---

<sup>6</sup>Processos estocásticos são processos cuja evolução no tempo é gerada por leis probabilísticas.

### Passeio aleatório intervalar

Seja  $\epsilon_t$  um processo puramente aleatório com média  $\mu_\epsilon$  e variância  $\sigma_\epsilon^2$ , um processo  $y_t$  é dito ser um *passeio aleatório* se

$$y_t = y_{t-1} + \epsilon_t.$$

Efetuada substituições sucessivas obtém-se

$$\begin{aligned} y_t &= y_{t-2} + \epsilon_{t-1} + \epsilon_t \\ &= y_{t-3} + \epsilon_{t-2} + \epsilon_{t-1} + \epsilon_t \\ &\vdots \\ &= y_0 + \sum_{j=1}^t \epsilon_j. \end{aligned} \tag{4.21}$$

Em STI, seja  $\mathbf{u}_t = (u_t^U, u_t^L)^\top$  um vetor bidimensional representando um processo puramente aleatório intervalar com vetor de médias  $\boldsymbol{\mu}_u^I$  e vetor de variâncias  $\boldsymbol{\sigma}_u^I$ , designamos  $\mathbf{I}_t$  um *passeio aleatório intervalar* se

$$\mathbf{I}_t = \mathbf{I}_{t-1} + \mathbf{u}_t. \tag{4.22}$$

Através de substituições sucessivas como em (4.21), chega-se a

$$\begin{aligned} \mathbf{I}_t &= \mathbf{I}_{t-2} + \mathbf{u}_{t-1} + \mathbf{u}_t \\ &= \mathbf{I}_{t-3} + \mathbf{u}_{t-2} + \mathbf{u}_{t-1} + \mathbf{u}_t \\ &\vdots \\ &= \mathbf{I}_0 + \sum_{j=1}^t \mathbf{u}_j \\ &= \begin{bmatrix} X_0^U \\ X_0^L \end{bmatrix} + \sum_{j=1}^t \begin{bmatrix} u_j^U \\ u_j^L \end{bmatrix} \\ &= \begin{bmatrix} X_0^U + \sum_{j=1}^t u_j^U \\ X_0^L + \sum_{j=1}^t u_j^L \end{bmatrix}, \end{aligned}$$

■

em que

$$\mathbf{u}_j = \begin{bmatrix} u_j^U \\ u_j^L \end{bmatrix} = \begin{bmatrix} X_j^U - X_{j-1}^U \\ X_j^L - X_{j-1}^L \end{bmatrix}.$$

No processo intervalar adaptado para STI e inspirado no processo passeio aleatório univalorado a equação do limite superior não considera qualquer informação do limite inferior e vice-versa.

Então, define-se o processo passeio aleatório intervalar como o processo intervalar em que a série do limite inferior  $\{X_t^L\}$  e a série do limite superior  $\{X_t^U\}$  são dois processos passeios aleatórios univalorados condicionados a  $X_t^U \geq X_t^L$  para todo  $t$ . Desta forma, nota-se que este processo definido para STI é uma generalização do processo univalorado.

### Processo intervalar estacionário

A estacionariedade é uma suposição crucial frequentemente feita sob séries temporais. Sob esta suposição, a série se desenvolve no tempo aleatoriamente em torno de uma média constante, refletindo alguma forma de equilíbrio estável (Morettin e Tolo, 2006) [74]. Intuitivamente, Morettin e Tolo (2006) [74] afirmam ainda que um processo é estacionário se ele se desenvolve no tempo de modo que a escolha de uma origem dos tempos não é importante, ou seja, as características de  $\{y_t\}$  são as mesmas de  $\{y_{t+\tau}\}$ , para todo  $\tau$ .

Recentemente, Han et al. (2008) [47] estenderam o conceito de estacionariedade para dados simbólicos do tipo intervalo e, então, propuseram uma definição formal de *estacionariedade intervalar*. Os autores partiram da seguinte definição de processos univalorados fracamente estacionários:

**Definição 4.5.1.** *Se a média  $\mu_t$  e a  $k$ -ésima autocovariância  $c_k$  não dependem do instante de tempo  $t$ , então o processo  $y_t$  é dito ser fracamente estacionário.*

A Definição 4.5.1 é a definição de estacionariedade fraca para séries temporais univaloradas usada no artigo de Han et al. (2008) [47]. Para definir estacionariedade para STI, eles partiram do princípio de que uma STI é limitada por duas séries univaloradas: a série do limite superior e a série do limite inferior. Sendo assim, eles consideram que a definição de estacionariedade intervalar pode ser derivada de forma direta do caso univalorado e afirmam que o processo univalorado

estacionário é um caso especial do processo intervalar estacionário. A definição de estacionariedade intervalar de Han et al. (2008) [47] é a seguinte:

**Definição 4.5.2** (Han et al., 2008). *Para um processo intervalar  $\{\mathbf{I}_t\} = \{[X_t^L; X_t^U] \mid X_t^L, X_t^U \in \mathbb{R}, X_t^U \geq X_t^L\}_{t=-\infty}^{\infty}$ , se as séries temporais univariadas  $\{X_t^L\}_{t=-\infty}^{\infty}$  e  $\{X_t^U\}_{t=-\infty}^{\infty}$  são ambas fracamente estacionárias, então este processo  $\{\mathbf{I}_t\}$  é dito ser intervalar estacionário.*

A definição dada por Han et al. (2008) [47] é de suma importância para a análise de STI. Através dele, por exemplo, testes de estacionariedade aplicados a séries univariadas se tornam válidos no exame da estacionariedade intervalar. Estes testes podem ser efetuados de modo independentes às séries dos limites superior e inferior dos intervalos e, pela Definição 4.5.2, constatada a estacionariedade de ambos os limites, constata-se a estacionariedade intervalar.

Apesar de não seguirmos rigorosamente as suposições que envolvem a teoria dos modelos tradicionais de séries temporais univariadas na proposição de modelos para STI, se necessário, nós adotaremos a Definição 4.5.2 de Han et al. (2008) [47] referindo-se a STI estacionárias.

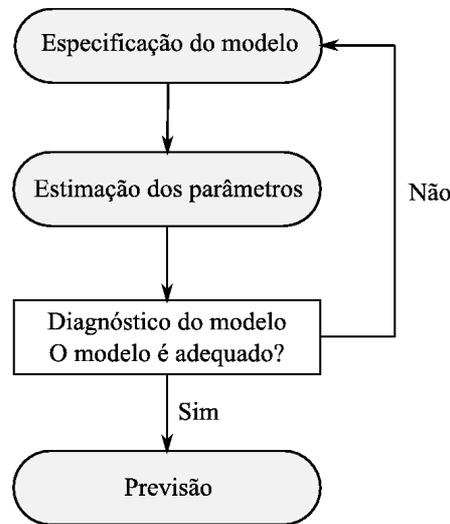
### 4.5.1 Modelos iARMA

A popularidade e a utilização bem sucedida da metodologia Box–Jenkins (Box e Jenkins, 1976) [12] para modelagem e previsão de séries temporais univariadas nos motivaram a considerar estratégias inspiradas nos modelos ARMA para o tratamento de STI. É razoável esperar que o uso de mais informações através da inclusão de variáveis correlacionadas no modelo aumente a precisão das previsões. No entanto, convimos que o processo de estimação de tal modelo é mais complexo que a estimação de modelos ARMA para dados univariados.

Os modelos ARMA estão fortemente atrelados aos princípios de especificação ou identificação do modelo, estimação dos parâmetros e previsão de séries temporais. Estes passos da metodologia Box–Jenkins estão apresentados na Figura 4.12.

Nós apresentamos, nesta seção, um novo paradigma para modelagem e previsão de STI usando modelos inspirados nos modelos ARMA univariados. Similarmente aos modelos ARMA univariados, em que os regressores são as séries defasadas da

Figura 4.12: Diagrama da metodologia Box–Jenkins.



própria série a ser prevista, nos modelos ARMA construídos para STI os regressores são as STI defasadas da própria STI a ser prevista. Conforme metodologia descrita na Seção 4.1, a principal ideia deste enfoque é assumir que a trajetória da série do limite superior da STI,  $\{X_t^U\}$ , pode ser afetada pela trajetória da série do limite inferior,  $\{X_t^L\}$ , e vice-versa. Os modelos resultantes são idênticos aos modelos ARMA bivariados que são casos particulares dos modelos vetores autorregressivos médias móveis (VARMA). No entanto, para ajuste de STI algumas considerações baseadas em restrições nos parâmetros são feitas para garantir a coerência matemática dos intervalos previstos.

Por convenção, intitulamos o modelo estendido aqui por *modelo intervalar baseado no modelo ARMA* ou, simplesmente, modelo iARMA. Também empregamos o termo *iARMA-based* em artigo que submetemos a um periódico em 2010.

Antes da apresentação dos modelos iARMA, nós apresentamos brevemente os processos denominados *médias móveis* (MA), *autorregressivos* (AR) e *autorregressivos médias móveis* (ARMA) para séries temporais univariadas. Seja  $\epsilon_t$  um processo puramente aleatório com média zero e variância constante  $\sigma_\epsilon^2$ , um processo  $y_t$  é dito ser um processo de médias móveis de ordem  $q$ , ou simplesmente MA( $q$ ), se

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}.$$

Por sua vez, um processo  $y_t$  é chamado de processo autorregressivo de ordem  $p$ , ou

AR( $p$ ), se

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \epsilon_t.$$

O processo AR( $p$ ) é similar ao modelo de regressão linear múltipla, em que os valores defasados (passados) de  $y_t$  assumem o papel das variáveis explicativas. Desta forma, processos AR( $p$ ) podem ser usados como modelos se for razoável assumir que o valor atual de uma série temporal depende do seu passado imediato mais um erro,  $\epsilon_t$ .

Através da combinação dos modelos AR( $p$ ) e MA( $q$ ) pode-se obter uma representação adequada com uma quantidade menor de parâmetros. Esta combinação é designada por modelos autorregressivos médias móveis de ordens  $p$  e  $q$ , ou simplesmente ARMA( $p, q$ ). A ordem  $p$  corresponde às componentes autorregressivas e a ordem  $q$  às componentes médias móveis. Em dados univariados, os modelos ARMA( $p, q$ ) são especialmente úteis por, geralmente, conduzirem a uma modelagem mais parcimoniosa que a obtida através dos modelos autorregressivos e dos modelos de médias móveis individualmente. O modelo ARMA( $p, q$ ) é dado por

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \\ &= \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}, \end{aligned} \quad (4.23)$$

em que  $\epsilon_t$  é um processo puramente aleatório com média zero e variância constante  $\sigma_\epsilon^2$ ;  $\phi_i$ , com  $i = 1, \dots, p$ , e  $\theta_j$ , com  $j = 1, \dots, q$ , são os parâmetros do modelo a serem estimados. Note que os modelos AR e MA podem ser obtidos como casos especiais do modelo ARMA, quando  $q = 0$  e  $p = 0$ , respectivamente.

Usando a metodologia descrita na Seção 4.1, i.e., com a substituição do escalar  $y_t$  na equação (4.23) pelo vetor intervalo  $\mathbf{I}_t$  e tomando, em vez de parâmetros escalares, matrizes ( $2 \times 2$ ) de parâmetros, promovemos a extensão dos modelos autorregressivos médias móveis de ordem  $p$  e  $q$  para tratar STI, designada modelos iARMA( $p, q$ ). Assim, o modelo iARMA( $p, q$ ) é dado por

$$\begin{aligned} \mathbf{I}_t &= \Phi_1 \mathbf{I}_{t-1} + \cdots + \Phi_p \mathbf{I}_{t-p} + \mathbf{u}_t + \Theta_1 \mathbf{u}_{t-1} + \cdots + \Theta_q \mathbf{u}_{t-q} \\ &= \sum_{i=1}^p \Phi_i \mathbf{I}_{t-i} + \mathbf{u}_t + \sum_{j=1}^q \Theta_j \mathbf{u}_{t-j}. \end{aligned} \quad (4.24)$$

em que  $\mathbf{u}_t = (u_t^U, u_t^L)^\top$  é um processo intervalar puramente aleatório com vetor de

médias nulo e vetor de variância constante, i.e.,

$$\mathbf{E}(\mathbf{u}_t) = \begin{bmatrix} \mathbf{E}(u_t^U) \\ \mathbf{E}(u_t^L) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{e} \quad \text{Var}(\mathbf{u}_t) = \begin{bmatrix} \text{Var}(u_t^U) \\ \text{Var}(u_t^L) \end{bmatrix} = \begin{bmatrix} (\sigma_u^U)^2 \\ (\sigma_u^L)^2 \end{bmatrix}.$$

As matrizes  $(2 \times 2)$  de parâmetros do modelo iARMA( $p, q$ ),  $\Phi_i$ , com  $i = 1, \dots, p$ , e  $\Theta_j$ , com  $j = 1, \dots, q$ , são dadas por

$$\Phi_i = \begin{bmatrix} \phi_{i1} & \phi_{i2} \\ \phi_{i3} & \phi_{i4} \end{bmatrix} \quad \text{e} \quad \Theta_j = \begin{bmatrix} \theta_{j1} & \theta_{j2} \\ \theta_{j3} & \theta_{j4} \end{bmatrix},$$

de modo que os elementos das matrizes de parâmetros  $\Phi_i$  e  $\Theta_j$  são os coeficientes do modelo.

Expandindo a expressão matricial (4.24), o modelo iARMA( $p, q$ ) pode ser reescrito como

$$\begin{aligned} \mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} &= \sum_{i=1}^p \begin{bmatrix} \phi_{i1} & \phi_{i2} \\ \phi_{i3} & \phi_{i4} \end{bmatrix} \begin{bmatrix} X_{t-i}^U \\ X_{t-i}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} + \sum_{j=1}^q \begin{bmatrix} \theta_{j1} & \theta_{j2} \\ \theta_{j3} & \theta_{j4} \end{bmatrix} \begin{bmatrix} u_{t-j}^U \\ u_{t-j}^L \end{bmatrix} \\ &= \sum_{i=1}^p \begin{bmatrix} \phi_{i1}X_{t-i}^U + \phi_{i2}X_{t-i}^L \\ \phi_{i3}X_{t-i}^U + \phi_{i4}X_{t-i}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} + \sum_{j=1}^q \begin{bmatrix} \theta_{j1}u_{t-j}^U + \theta_{j2}u_{t-j}^L \\ \theta_{j3}u_{t-j}^U + \theta_{j4}u_{t-j}^L \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^p (\phi_{i1}X_{t-i}^U + \phi_{i2}X_{t-i}^L) + \sum_{j=1}^q (\theta_{j1}u_{t-j}^U + \theta_{j2}u_{t-j}^L) + u_t^U \\ \sum_{i=1}^p (\phi_{i3}X_{t-i}^U + \phi_{i4}X_{t-i}^L) + \sum_{j=1}^q (\theta_{j3}u_{t-j}^U + \theta_{j4}u_{t-j}^L) + u_t^L \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^p \phi_{i1}X_{t-i}^U + u_t^U + \sum_{j=1}^q \theta_{j1}u_{t-j}^U + \sum_{i=1}^p \phi_{i2}X_{t-i}^L + \sum_{j=1}^q \theta_{j2}u_{t-j}^L \\ \sum_{i=1}^p \phi_{i4}X_{t-i}^L + u_t^L + \sum_{j=1}^q \theta_{j4}u_{t-j}^L + \sum_{i=1}^p \phi_{i3}X_{t-i}^U + \sum_{j=1}^q \theta_{j3}u_{t-j}^U \end{bmatrix}. \end{aligned} \tag{4.25}$$

■

Note que as duas equações do modelo têm os mesmos regressores – séries defasadas de  $X_t^L$ ,  $X_t^U$ ,  $u_t^L$  e  $u_t^U$  – e veja que os três primeiros termos de cada equação equivalem ao modelo ARMA para séries temporais univariadas. Assim, a previsão de um limite através do modelo iARMA( $p, q$ ) é dada pela previsão deste limite pelo modelo ARMA univariado mais dois termos referentes à informações do outro limite.

Veja também na expressão (4.25) que, se  $\Phi_i$  e  $\Theta_j$  são matrizes diagonais ( $\phi_{i2} = \phi_{i3} = 0$  e  $\theta_{j2} = \theta_{j3} = 0, \forall i, j$ ), o modelo iARMA tem duas equações independentes, i.e., a previsão de um dos limites independe de informações do outro limite. Isso faz com que a STI seja ajustada através de dois modelos ARMA univariados, em que as séries do limite superior e do limite inferior são tratadas como dados clássicos.

Por fim, o modelo iARMA mantém a propriedade de ser uma generalização do modelo univariado, ou seja, se  $X_t^L = X_t^U$  e  $u_t^L = u_t^U, \forall t$ , então o modelo iARMA se reduz ao modelo ARMA univariado.

Como nos modelos ARMA em dados univariados, aqui, os modelos autorregressivos para STI, iAR( $p$ ), e os modelos de médias móveis para STI, iMA( $p$ ), são obtidos como casos especiais dos modelos iARMA( $p, q$ ), quando  $q = 0$  e  $p = 0$ , respectivamente.

O modelo iAR( $p$ ) é dado por

$$\begin{aligned} \mathbf{I}_t &= \Phi_1 \mathbf{I}_{t-1} + \cdots + \Phi_p \mathbf{I}_{t-p} + \mathbf{u}_t \\ &= \sum_{i=1}^p \Phi_i \mathbf{I}_{t-i} + \mathbf{u}_t \end{aligned} \quad (4.26)$$

e expandido da forma

$$\begin{aligned} \mathbf{I}_t &= \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} = \sum_{i=1}^p \begin{bmatrix} \phi_{i1} & \phi_{i2} \\ \phi_{i3} & \phi_{i4} \end{bmatrix} \begin{bmatrix} X_{t-i}^U \\ X_{t-i}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} \\ &= \sum_{i=1}^p \begin{bmatrix} \phi_{i1} X_{t-i}^U + \phi_{i2} X_{t-i}^L \\ \phi_{i3} X_{t-i}^U + \phi_{i4} X_{t-i}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^p \phi_{i1} X_{t-i}^U + \sum_{i=1}^p \phi_{i2} X_{t-i}^L + u_t^U \\ \sum_{i=1}^p \phi_{i4} X_{t-i}^L + \sum_{i=1}^p \phi_{i3} X_{t-i}^U + u_t^L \end{bmatrix}. \end{aligned}$$

■

Enquanto que o modelo iMA( $p$ ) é da forma

$$\begin{aligned} \mathbf{I}_t &= \mathbf{u}_t + \Theta_1 \mathbf{u}_{t-1} + \cdots + \Theta_q \mathbf{u}_{t-q} \\ &= \mathbf{u}_t + \sum_{j=1}^q \Theta_j \mathbf{u}_{t-j}, \end{aligned}$$

e

$$\begin{aligned}
\mathbf{I}_t &= \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} = \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} + \sum_{j=1}^q \begin{bmatrix} \theta_{j1} & \theta_{j2} \\ \theta_{j3} & \theta_{j4} \end{bmatrix} \begin{bmatrix} u_{t-j}^U \\ u_{t-j}^L \end{bmatrix} \\
&= \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} + \sum_{j=1}^q \begin{bmatrix} \theta_{j1}u_{t-j}^U + \theta_{j2}u_{t-j}^L \\ \theta_{j3}u_{t-j}^U + \theta_{j4}u_{t-j}^L \end{bmatrix} \\
&= \begin{bmatrix} u_t^U + \sum_{j=1}^q \theta_{j1}u_{t-j}^U + \sum_{j=1}^q \theta_{j2}u_{t-j}^L \\ u_t^L + \sum_{j=1}^q \theta_{j4}u_{t-j}^L + \sum_{j=1}^q \theta_{j3}u_{t-j}^U \end{bmatrix},
\end{aligned}$$

■

em que  $\mathbf{u}_t = (u_t^U, u_t^L)^\top$  é um processo intervalar puramente aleatório com vetor de médias nulo e vetor de variância constante.

A seguir descreveremos um procedimento possível para estimação de parâmetros de modelos iARMA. O método aplicado é uma adaptação de métodos clássicos de estimação de modelos ARMA univariados.

### 4.5.2 Estimação irrestrita de modelos iARMA

Após formulação matemática do modelo iARMA( $p, q$ ), o próximo passo é proceder com a estimação das matrizes de parâmetros  $\Phi_i$  e  $\Theta_j$  condicionada à STI observada. O método de estimação dos parâmetros dos modelos iARMA( $p, q$ ) é derivado de métodos clássicos de estimação de modelos ARMA para séries temporais univariadas.

Visto que a distribuição do processo gerador de STI ainda não foi estudada e, então, é desconhecida, alguns métodos clássicos de estimação não são aplicáveis para estimação de modelos iARMA( $p, q$ ). Por exemplo, como não conhecemos a distribuição deste processo, não conhecemos a função verossimilhança que é requerida na estimação por máxima verossimilhança, o que nos impede de usar tal método. Uma alternativa que encontramos foi implementarmos técnicas de otimização numérica em que os parâmetros estimados são uma solução da minimização de critérios específicos descartando, assim, pressuposições sobre o processo  $\{\mathbf{I}_t\}$ .

Algumas considerações sobre os modelos iARMA merecem destaque antes de apresentarmos o procedimento de estimação adotado. Considere a seguinte forma expandida do modelo iARMA( $p, q$ ):

$$\mathbf{I}_t = \begin{bmatrix} \sum_{i=1}^p \phi_{i1} X_{t-i}^U + u_t^U + \sum_{j=1}^q \theta_{j1} u_{t-j}^U + \sum_{i=1}^p \phi_{i2} X_{t-i}^L + \sum_{j=1}^q \theta_{j2} u_{t-j}^L \\ \sum_{i=1}^p \phi_{i4} X_{t-i}^L + u_t^L + \sum_{j=1}^q \theta_{j4} u_{t-j}^L + \sum_{i=1}^p \phi_{i3} X_{t-i}^U + \sum_{j=1}^q \theta_{j3} u_{t-j}^U \end{bmatrix}.$$

Podemos ver que, apesar do modelo iARMA( $p, q$ ) ser escrito como um vetor bidimensional composto por dois modelos distintos – cada um consistindo da previsão de um limite –, a quantidade de termos autorregressivos,  $p$ , é a mesma para ambos os modelos, assim como é a mesma a quantidade de termos médias móveis,  $q$ . Com isso, o número de parâmetros a ser estimado para previsão do limite superior é o mesmo a ser estimado para previsão do limite inferior, i.e.,  $2(p + q)$  parâmetros em cada equação. Ou seja, o número de valores passados considerados no modelo do limite superior é igual ao número de valores passados considerados no modelo do limite inferior. Esta é uma característica importante que deve ser considerada nas equações simultâneas. Visto isso, o número total de parâmetros a serem estimados em um modelo iARMA( $p, q$ ) é  $4(p + q)$ .

Hannan e Rissanen (1982) [49] propuseram para dados univariados o método de estimação conhecido por *mínimos quadrados em dois estágios* (MQ2E). Na estimação de modelos ARMA univariados, a ideia do estimador de MQ2E é ajustar à série  $y_t$  um modelo autorregressivo puro de alta ordem em um primeiro estágio com a finalidade de estimar os resíduos. Os resíduos estimados,  $\hat{\epsilon}_t$ , por este modelo autorregressivo são estimativas dos verdadeiros resíduos,  $\epsilon_t$ , em (4.23). No segundo estágio, os resíduos  $\hat{\epsilon}_t$  são usados em substituição de  $\epsilon_t$  e uma nova estimação é efetuada, i.e., uma regressão de  $y_t$  em  $y_{t-i}$ , com  $i = 1, \dots, p$ , e  $\hat{\epsilon}_{t-j}$ , com  $j = 1, \dots, q$  para combinações de valores de  $p$  e  $q$ . O método de estimação de MQ2E é de fácil implementação computacional.

Para estimação de modelos univariados, diferentes regras para seleção do número de defasagens, i.e., da ordem do modelo AR( $l$ ) inicial foram propostas. Por exemplo, Hannan e Kavalieris (1984) [48] propuseram selecionar  $l$  através dos critérios *Akaike's Information Criterion* (AIC) ou *Bayesian Information Criterion* (BIC), enquanto Koreisha e Pukkila (1990) [58] propuseram escolher  $l = \sqrt{n}$  ou  $l = \frac{1}{2}\sqrt{n}$ , em

que  $l$  é o número de defasagens requerido e aumenta com o tamanho da amostra,  $n$ .

Para proceder com a estimação das matrizes de parâmetros dos modelos iARMA de ordem  $p, q > 0$ , nós consideramos o procedimento em dois estágios de Hannan e Rissanen (1982) [49], no entanto, vale salientar que algumas adaptações foram inseridas no processo de estimação do modelo iARMA.

O procedimento adotado é similar ao caso univariado e também de fácil implementação. A ideia é usar a representação finita iAR dada em (4.26) com a finalidade de estimar os resíduos  $\mathbf{u}_t$  no primeiro estágio. Por simplicidade, optamos pela sugestão em modelos univariados de Koreisha e Pukkila (1990) [58] na escolha de  $l = \frac{1}{2}\sqrt{n}$  no modelo iAR( $l$ ) do primeiro estágio. No segundo estágio do nosso procedimento de estimação, os resíduos do modelo iAR( $l$ ),  $\hat{\mathbf{u}}_t$ , são inseridos em (4.24) no lugar de  $\mathbf{u}_t$  e, então, a estimação da matriz ( $2 \times 2(p + q)$ ),

$$\mathbf{\Pi} = (\mathbf{\Phi}_1, \dots, \mathbf{\Phi}_p, \mathbf{\Theta}_1, \dots, \mathbf{\Theta}_q)$$

representando o conjunto de parâmetros desconhecidos do modelo iARMA( $p, q$ ), será executada. Como será detalhado a seguir, tanto para estimação do modelo iAR( $l$ ) no primeiro estágio quanto na estimação final da matriz  $\mathbf{\Pi}$  no segundo estágio, são utilizadas técnicas de otimização numérica.

Com o enfoque considerado nesta tese, nosso objetivo é encontrar o conjunto de  $4(p + q)$  valores de coeficientes do modelo,  $\mathbf{\Pi}$ , que minimizam o critério representado por uma medida de precisão de previsão. Assim, o método empregado na estimação de parâmetros de modelos iARMA( $p, q$ ) consiste da utilização de técnicas de otimização para minimizar o critério escolhido.

Nós assumimos como critério a soma de quadrados dos erros de previsão na forma intervalar que é uma função escalar dada por

$$\begin{aligned} G(\mathbf{\Pi}) &= \sum_{t=1}^n (\mathbf{I}_t - \hat{\mathbf{I}}_t)^\top (\mathbf{I}_t - \hat{\mathbf{I}}_t) \\ &= \sum_{t=1}^n (X_t^U - \hat{X}_t^U)^2 + \sum_{t=1}^n (X_t^L - \hat{X}_t^L)^2. \end{aligned} \quad (4.27)$$

Assim, a estimação das matrizes de parâmetros dos modelos iARMA pode ser tratada como um *problema de mínimos quadrados linear*. Reescrevendo a expressão

(4.27), temos que a função objetivo a ser minimizada para estimação dos parâmetros do modelo iARMA( $p, q$ ) é dada por

$$G(\phi_{ik}, \theta_{jk}) = \sum_{t=1}^n \left( X_t^U - \sum_{i=1}^p \phi_{i1} X_{t-i}^U - \sum_{j=1}^q \theta_{j1} u_{t-j}^U - \sum_{i=1}^p \phi_{i2} X_{t-i}^L - \sum_{j=1}^q \theta_{j2} u_{t-j}^L \right)^2 + \sum_{t=1}^n \left( X_t^L - \sum_{i=1}^p \phi_{i4} X_{t-i}^L - \sum_{j=1}^q \theta_{j4} u_{t-j}^L - \sum_{i=1}^p \phi_{i3} X_{t-i}^U - \sum_{j=1}^q \theta_{j3} u_{t-j}^U \right)^2, \quad (4.28)$$

com  $i = 1, \dots, p$ ,  $j = 1, \dots, q$  e  $k = 1, \dots, 4$ . No caso particular do modelo iAR( $p$ ), a função objetivo a ser minimizada é dada por

$$G(\phi_{ik}) = \sum_{t=1}^n \left( X_t^U - \sum_{i=1}^p \phi_{i1} X_{t-i}^U - \sum_{i=1}^p \phi_{i2} X_{t-i}^L \right)^2 + \sum_{t=1}^n \left( X_t^L - \sum_{i=1}^p \phi_{i4} X_{t-i}^L - \sum_{i=1}^p \phi_{i3} X_{t-i}^U \right)^2. \quad (4.29)$$

Baseado na função objetivo  $G(\phi_{ik}, \theta_{jk})$  em (4.28), nós escrevemos o processo de estimação das matrizes de parâmetros do modelo iARMA( $p, q$ ) como um problema de otimização, ou seja,

$$\min_{\phi, \theta} G(\phi_{ik}, \theta_{jk}).$$

O problema de minimização definido acima, em que  $G(\phi_{ik}, \theta_{jk})$  é formulada como (4.28), pode ser resolvido facilmente e de forma eficiente via métodos de programação numérica.

No ajuste de modelos iAR( $p$ ), as estimativas das matrizes de parâmetros  $\Phi_1, \dots, \Phi_p$  são obtidas em um único estágio, bastando minimizar a função objetivo  $G(\phi_{ik})$  dada em (4.29).

Se a finalidade é ajustar modelos iARMA( $p, q$ ) de ordem  $p, q \neq 0$ , o procedimento em dois estágios é requerido. Sendo assim, adotado o procedimento em dois estágios como em Hannan e Rissanen (1982) [49], a regra para seleção da ordem do modelo AR( $l$ ) inicial sugerida por Koreisha e Pukkila (1990) [58], i.e.,  $l = \frac{1}{2}\sqrt{n}$  e de posse da função objetivo  $G(\phi_{ik}, \theta_{jk})$  que nós definimos, o procedimento para estimação das matrizes de parâmetros do modelo iARMA( $p, q$ ) é como segue.

O primeiro estágio do algoritmo adaptado para STI consiste de uma autorregressão intervalar de alta ordem do tipo

$$\mathbf{I}_t = \sum_{i=1}^l \Phi_i^{(1)} \mathbf{I}_{t-i} + \mathbf{u}_t,$$

em que as matrizes de parâmetros do primeiro estágio  $\Phi_i^{(1)}$  são estimadas através da minimização do critério

$$\begin{aligned} G(\phi_{ik}) &= \sum_{t=1}^n \left( X_t^U - \sum_{i=1}^l \phi_{i1} X_{t-i}^U - \sum_{i=1}^l \phi_{i2} X_{t-i}^L \right)^2 \\ &+ \sum_{t=1}^n \left( X_t^L - \sum_{i=1}^l \phi_{i4} X_{t-i}^L - \sum_{i=1}^l \phi_{i3} X_{t-i}^U \right)^2, \end{aligned} \quad (4.30)$$

com  $l = \frac{1}{2}\sqrt{n}$ . Estimadas as matrizes  $\Phi_i^{(1)}$ , nós computamos os resíduos estimados,

$$\hat{\mathbf{u}}_t = \mathbf{I}_t - \sum_{i=1}^l \hat{\Phi}_i^{(1)} \mathbf{I}_{t-i},$$

em que  $\hat{\Phi}_i^{(1)}$  são as estimativas no primeiro estágio.

No segundo estágio, os resíduos estimados  $\hat{\mathbf{u}}_t$  são inseridos em (4.24) no lugar de  $\mathbf{u}_t$  e nós estimamos as matrizes de parâmetros  $\mathbf{\Pi} = (\Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_q)$  através da representação

$$\begin{aligned} \mathbf{I}_t &= \Phi_1 \mathbf{I}_{t-1} + \dots + \Phi_p \mathbf{I}_{t-p} + \Theta_1 \hat{\mathbf{u}}_{t-1} + \dots + \Theta_q \hat{\mathbf{u}}_{t-q} \\ &= \begin{bmatrix} \sum_{i=1}^p \phi_{i1} X_{t-i}^U + \sum_{j=1}^q \theta_{j1} \hat{u}_{t-j}^U + \sum_{i=1}^p \phi_{i2} X_{t-i}^L + \sum_{j=1}^q \theta_{j2} \hat{u}_{t-j}^L \\ \sum_{i=1}^p \phi_{i4} X_{t-i}^L + \sum_{j=1}^q \theta_{j4} \hat{u}_{t-j}^L + \sum_{i=1}^p \phi_{i3} X_{t-i}^U + \sum_{j=1}^q \theta_{j3} \hat{u}_{t-j}^U \end{bmatrix}, \end{aligned} \quad (4.31)$$

minimizando a função objetivo

$$\begin{aligned} G(\phi_{ik}, \theta_{jk}) &= \sum_{t=1}^n \left( X_t^U - \sum_{i=1}^p \phi_{i1} X_{t-i}^U - \sum_{j=1}^q \theta_{j1} \hat{u}_{t-j}^U - \sum_{i=1}^p \phi_{i2} X_{t-i}^L - \sum_{j=1}^q \theta_{j2} \hat{u}_{t-j}^L \right)^2 \\ &+ \sum_{t=1}^n \left( X_t^L - \sum_{i=1}^p \phi_{i4} X_{t-i}^L - \sum_{j=1}^q \theta_{j4} \hat{u}_{t-j}^L - \sum_{i=1}^p \phi_{i3} X_{t-i}^U - \sum_{j=1}^q \theta_{j3} \hat{u}_{t-j}^U \right)^2. \end{aligned} \quad (4.32)$$

Assim, as estimativas  $\hat{\mathbf{\Pi}} = (\hat{\Phi}_1, \dots, \hat{\Phi}_p, \hat{\Theta}_1, \dots, \hat{\Theta}_q)$  são obtidas.

**Exemplo 4.5** (Aplicação do modelo iARMA). Apresentamos um exemplo com o ajuste de modelos iARMA( $p, q$ ) à STI dos preços diários de negociações de ações da Petrobras citada na Seção 4.2.1. Nós ajustamos um modelo iAR(1) e um modelo iARMA(1,1). A finalidade deste exemplo não é avaliar as precisões das previsões e sim ilustrar o uso do modelo iARMA.

Para o modelo iAR(1), a matriz de parâmetros estimada para a STI de preços de ações da Petrobras foi  $\widehat{\Phi}_1 = \begin{bmatrix} 0,603 & 0,407 \\ 0,384 & 0,606 \end{bmatrix}$  e, conseqüentemente, o modelo ajustado foi

$$\widehat{\mathbf{I}}_t = \begin{bmatrix} 0,603X_{t-1}^U + 0,407X_{t-1}^L \\ 0,606X_{t-1}^L + 0,384X_{t-1}^U \end{bmatrix}.$$

A Figura 4.13 mostra os intervalos reais (segmentos de reta pretos) desta STI e os intervalos ajustados via modelo iAR(1) (segmentos de reta cinza).

Para o modelo iARMA(1,1), as matrizes de parâmetros estimadas para a STI de preços de ações da Petrobras foram  $\widehat{\Phi}_1 = \begin{bmatrix} 1,332 & -0,339 \\ 0,450 & 0,538 \end{bmatrix}$  e  $\widehat{\Theta}_1 = \begin{bmatrix} -0,658 & 0,838 \\ 0,193 & 0,079 \end{bmatrix}$  que implicam no modelo ajustado

$$\widehat{\mathbf{I}}_t = \begin{bmatrix} 1,332X_{t-1}^U - 0,658u_{t-1}^U - 0,339X_{t-1}^L + 0,838u_{t-1}^L \\ 0,538X_{t-1}^L + 0,079u_{t-1}^L + 0,450X_{t-1}^U + 0,193u_{t-1}^U \end{bmatrix}.$$

O modelo iAR no primeiro estágio do procedimento de estimação foi ajustado com  $l = 4$ . A Figura 4.14 mostra os intervalos reais (segmentos de reta pretos) desta STI e os intervalos ajustados via modelo iARMA(1,1) (segmentos de reta cinza).

Nesta seção, nós apresentamos a forma geral dos modelos iARMA que propomos para modelagem e previsão de STI e o método de estimação das matrizes de parâmetros. Na seção seguinte, nós propomos uma nova versão do procedimento de estimação destes modelos. Esta nova versão tem por finalidade considerar a coerência matemática inerente aos dados simbólicos intervalares em que  $X_t^U \geq X_t^L$ ,  $t = 1, \dots, n$ . O novo procedimento consiste em impor restrições durante a otimização com a finalidade de obter estimativas de parâmetros que garantam que  $\widehat{X}_t^U \geq \widehat{X}_t^L \forall t$ .

Figura 4.13: Ajuste da STI de preços de ações da Petrobras via modelo iAR(1).

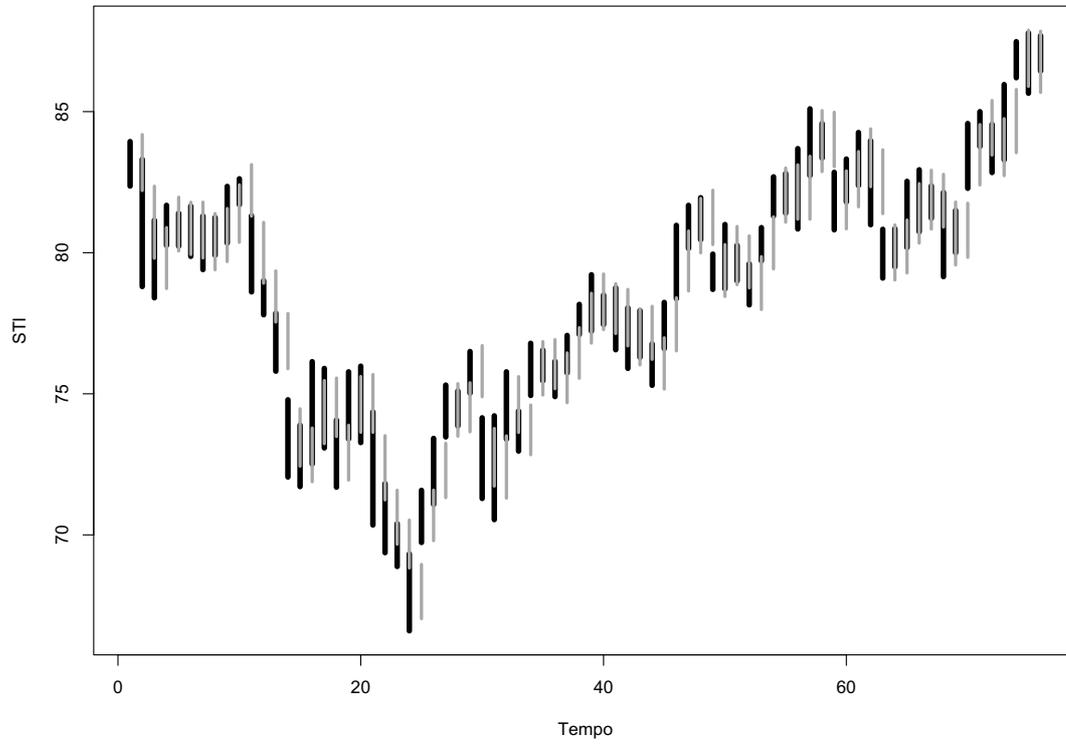
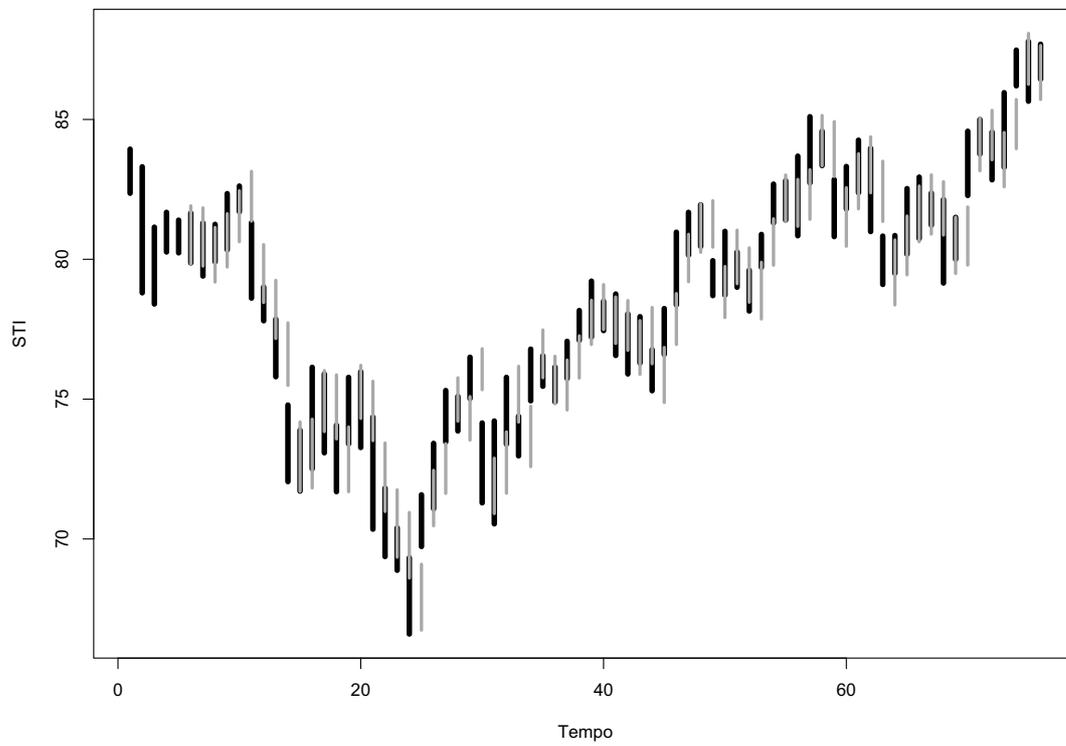


Figura 4.14: Ajuste da STI de preços de ações da Petrobras via modelo iARMA(1,1).



### 4.5.3 Estimação restrita de modelos iARMA

Apresentamos, nesta seção, uma versão modificada do procedimento de estimação das matrizes de parâmetros de modelos iARMA propostos na Seção 4.5.1. Esta nova versão surge da necessidade de considerar uma peculiaridade dos dados simbólicos do tipo intervalo. Especificamente, as modificações consistem em considerar a relação de ordem entre os valores de  $X_t^U$  e  $X_t^L$ , i.e., a coerência matemática inerente aos dados simbólicos intervalares, em que  $X_t^U \geq X_t^L$ , para todo  $t$ . Daí a necessidade da implementação de um mecanismo de estimação de parâmetros que forneça intervalos ajustados e previstos que também respeitem esta coerência matemática, i.e.,  $\hat{X}_t^U \geq \hat{X}_t^L$ .

As modificações aplicadas ao modelo iARMA consistem em incorporar à modelagem da STI métodos de estimação em que os parâmetros são sujeitos a algumas condições especiais. Com isso, o espaço paramétrico dos novos modelos passa a ser restrito. Por esta razão, chamamos a nova versão de *modelo iARMA com restrições nos parâmetros* ou, simplesmente, *modelo iARMA-restrito*. Para estimar as matrizes de parâmetros do modelo iARMA-restrito, nós empregamos técnicas de otimização com restrições lineares de desigualdade estabelecidas pelo usuário. O procedimento de estimação restrita será descrito em detalhes a seguir.

Antes de iniciarmos a apresentação do procedimento de estimação, faremos uma ressalva. As restrições apresentadas nesta seção são válidas sob a suposição de que  $\{\mathbf{I}_t\} = \{[X_t^L; X_t^U]\}$  é uma STI particular tal que a série univariada  $\{X_t^L\}_{t=-\infty}^{\infty}$  é uma série não-negativa, i.e.,  $X_t^L \geq 0, \forall t$ . No caso em que a série  $X_t^L$  não satisfaça esta condição para algum  $t$ , a operação matemática conhecida por *translação*<sup>7</sup> de função pode ser aplicada. Quando necessário, usaremos a translação para deslocar a STI  $\{\mathbf{I}_t\}$  no sentido vertical de modo que  $X_t^L \geq 0, \forall t$ . Em suma, se em algum instante de tempo  $t$ , observa-se  $X_t^L < 0$ , então, efetuamos a translação da STI realizando a seguinte operação em toda a STI:  $\{\mathbf{I}_t^+\} = \{\mathbf{I}_t + |\min\{X_t^L\}|\}$ . Esta operação, se aplicada, não altera a estrutura da STI e, então, os modelos podem ser aplicados normalmente à STI resultante  $\{\mathbf{I}_t^+\}$ . Os intervalos ajustados e previstos são obtidos

---

<sup>7</sup>Em geometria, define-se a operação de translação  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  como a aplicação  $T(x, y) = (x + h, y + k)$  onde  $h, k \in \mathbb{R}$ , sendo que  $h \neq 0$  e/ou  $k \neq 0$ . A ação da translação sobre uma curva faz com que a curva seja deslocada para outro local do plano  $\mathbb{R}^2$  tendo novos eixos principais paralelos aos eixos originais.

desfazendo a operação de translação, i.e.,  $\{\widehat{\mathbf{I}}_t\} = \{\widehat{\mathbf{I}}_t^+ - |\min\{X_t^L\}|\}$ . A razão para esta ressalva será melhor compreendida durante a apresentação das restrições ao longo desta seção.

Para facilitar a compreensão do procedimento de estimação restrita, iniciaremos com a apresentação da estimação de modelos iAR. Chamaremos o modelo iAR com estimação restrita de *modelo iAR-restrito*. De acordo com a versão irrestrita apresentada na Seção 4.5.1, a STI ajustada é representada por

$$\widehat{\mathbf{I}}_t = \begin{bmatrix} \sum_{i=1}^p \widehat{\phi}_{i1} X_{t-i}^U + \sum_{i=1}^p \widehat{\phi}_{i2} X_{t-i}^L \\ \sum_{i=1}^p \widehat{\phi}_{i4} X_{t-i}^L + \sum_{i=1}^p \widehat{\phi}_{i3} X_{t-i}^U \end{bmatrix}.$$

Sendo assim, incorporar restrições ao modelo iAR( $p$ ) com a finalidade de garantir  $\widehat{X}_t^U \geq \widehat{X}_t^L$  implica em impor a condição

$$\begin{aligned} \sum_{i=1}^p \widehat{\phi}_{i1} X_{t-i}^U + \sum_{i=1}^p \widehat{\phi}_{i2} X_{t-i}^L &\geq \sum_{i=1}^p \widehat{\phi}_{i4} X_{t-i}^L + \sum_{i=1}^p \widehat{\phi}_{i3} X_{t-i}^U \quad \Rightarrow \\ \widehat{\phi}_{11} X_{t-1}^U + \cdots + \widehat{\phi}_{p1} X_{t-p}^U + \widehat{\phi}_{12} X_{t-1}^L + \cdots + \widehat{\phi}_{p2} X_{t-p}^L &\geq \\ \widehat{\phi}_{13} X_{t-1}^U + \cdots + \widehat{\phi}_{p3} X_{t-p}^U + \widehat{\phi}_{14} X_{t-1}^L + \cdots + \widehat{\phi}_{p4} X_{t-p}^L. \end{aligned}$$

Podemos observar que ambos os lados da inequação acima apresentam termos comuns: defasagens da série do limite superior e da série do limite inferior. Os dois lados foram ordenados segundo os termos comuns e, assim, fica fácil observar que estes termos são multiplicados por coeficientes distintos. Observando ambos os lados da inequação acima, concluímos que uma forma de assegurar que os intervalos ajustados satisfaçam sempre a condição  $\widehat{X}_t^U \geq \widehat{X}_t^L$  é impor restrições aos parâmetros da forma

$$\begin{aligned} \widehat{\phi}_{11} &\geq \widehat{\phi}_{13}, \\ &\vdots \\ \widehat{\phi}_{p1} &\geq \widehat{\phi}_{p3}, \end{aligned}$$

(4.33)

$$\begin{aligned}\widehat{\phi}_{12} &\geq \widehat{\phi}_{14}, \\ &\vdots \\ \widehat{\phi}_{p2} &\geq \widehat{\phi}_{p4}.\end{aligned}$$

A autenticidade destas restrições pode ser verificada em um exemplo simples. Suponha que desejamos assegurar que  $\widehat{\phi}_{11}X_{t-1}^U \geq \widehat{\phi}_{13}X_{t-1}^U$ . É fácil ver que para isto acontecer basta que  $\widehat{\phi}_{11} \geq \widehat{\phi}_{13}$ .

As restrições impostas para estimação de modelos iAR podem, ainda, ser escritas matricialmente como

$$R\widehat{\gamma} = \begin{bmatrix} R_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & R_0 \end{bmatrix} \begin{bmatrix} \widehat{\phi}_{11} \\ \widehat{\phi}_{13} \\ \widehat{\phi}_{12} \\ \widehat{\phi}_{14} \\ \vdots \\ \widehat{\phi}_{p1} \\ \widehat{\phi}_{p3} \\ \widehat{\phi}_{p2} \\ \widehat{\phi}_{p4} \end{bmatrix} \geq 0, \quad (4.34)$$

em que  $R$  é uma matriz  $(2p \times 4p)$  conhecida e estabelecida pelo usuário,  $\widehat{\gamma}$  é um vetor  $(4p \times 1)$  irrestrito de parâmetros desconhecidos dado por

$$\widehat{\gamma} = \text{vec}(\widehat{\Phi}_1, \dots, \widehat{\Phi}_p),$$

em que “vec” denota o operador matemático que transforma uma matriz em um vetor através de concatenação das colunas da matriz.<sup>8</sup>  $\mathbf{0}$  é uma matriz  $(2 \times 4)$  nula e  $R_0$  é a matriz  $(2 \times 4)$  responsável pela inserção de restrições aos parâmetros dada por

$$R_0 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

A expressão  $R\widehat{\gamma} \geq 0$  resulta nas restrições (4.33) e é utilizada na implementação do algoritmo de otimização. A primeira linha da matriz  $R_0$  corresponde à restrição  $\widehat{\phi}_{i1} \geq \widehat{\phi}_{i3}$  e a segunda linha corresponde à restrição  $\widehat{\phi}_{i2} \geq \widehat{\phi}_{i4}$ , com  $i = 1, \dots, p$ .

---

<sup>8</sup>Seja  $M$  uma matriz  $(m \times k)$ ; o operador matemático vec transforma  $M$  em um vetor  $(mk \times 1)$  concatenando as colunas de  $M$ .

As estimativas restritas dos parâmetros do modelo iAR-restrito são obtidas via solução do problema de minimização da soma dos quadrados dos resíduos  $G(\phi_{ik})$  representada pela equação (4.29), sujeita a restrições lineares de desigualdade  $R\hat{\gamma} \geq 0$  como na expressão (4.34) impostas com a finalidade de assegurar que  $\hat{X}_t^U \geq \hat{X}_t^L$ . Formalmente, o procedimento de estimação restrita de modelos iAR( $p$ ) pode ser descrito como

$$\begin{aligned} &\text{Minimizar} && G(\phi_{ik}), \\ &\text{sujeito a} && R\hat{\gamma} \geq 0. \end{aligned}$$

Nenhuma modificação foi necessária na estrutura do modelo iAR para incorporar as restrições lineares de desigualdade  $R\hat{\gamma} \geq 0$  aos parâmetros do modelo. Isto faz com que o modelo iAR( $p$ )-restrito seja idêntico ao modelo irrestrito iAR( $p$ ) apresentado na equação (4.26) diferenciado apenas pelas restrições aos parâmetros impostas no procedimento de estimação.

Por sua vez, impor restrições ao procedimento de estimação de parâmetros de modelos iARMA( $p, q$ ), com  $p, q > 0$ , exige mais restrições lineares de desigualdade. Além das restrições nas matrizes de coeficientes dos termos autorregressivos,  $\hat{\Phi}_1, \dots, \hat{\Phi}_p$ , como no modelo iAR( $p$ )-restrito, devemos impor restrições às matrizes de coeficientes dos termos médias móveis,  $\hat{\Theta}_1, \dots, \hat{\Theta}_q$ . As restrições de desigualdade impostas para estimação de modelo iARMA( $p, q$ )-restrito são obtidas de maneira análoga às restrições do modelo iAR( $p$ )-restrito.

Consideremos o intervalo ajustado pelo modelo irrestrito iARMA( $p, q$ ) descrito em (4.31) como

$$\hat{\mathbf{I}}_t = \begin{bmatrix} \sum_{i=1}^p \hat{\phi}_{i1} X_{t-i}^U + \sum_{j=1}^q \hat{\theta}_{j1} \hat{u}_{t-j}^U + \sum_{i=1}^p \hat{\phi}_{i2} X_{t-i}^L + \sum_{j=1}^q \hat{\theta}_{j2} \hat{u}_{t-j}^L \\ \sum_{i=1}^p \hat{\phi}_{i4} X_{t-i}^L + \sum_{j=1}^q \hat{\theta}_{j4} \hat{u}_{t-j}^L + \sum_{i=1}^p \hat{\phi}_{i3} X_{t-i}^U + \sum_{j=1}^q \hat{\theta}_{j3} \hat{u}_{t-j}^U \end{bmatrix},$$

logo, a condição  $\hat{X}_t^U \geq \hat{X}_t^L$  pode ser garantida da seguinte maneira

$$\begin{aligned} &\sum_{i=1}^p \hat{\phi}_{i1} X_{t-i}^U + \sum_{j=1}^q \hat{\theta}_{j1} \hat{u}_{t-j}^U + \sum_{i=1}^p \hat{\phi}_{i2} X_{t-i}^L + \sum_{j=1}^q \hat{\theta}_{j2} \hat{u}_{t-j}^L && \geq \\ &\sum_{i=1}^p \hat{\phi}_{i4} X_{t-i}^L + \sum_{j=1}^q \hat{\theta}_{j4} \hat{u}_{t-j}^L + \sum_{i=1}^p \hat{\phi}_{i3} X_{t-i}^U + \sum_{j=1}^q \hat{\theta}_{j3} \hat{u}_{t-j}^U && \Rightarrow \end{aligned}$$

$$\begin{aligned}
& \widehat{\phi}_{11}X_{t-1}^U + \cdots + \widehat{\phi}_{p1}X_{t-p}^U + \widehat{\phi}_{12}X_{t-1}^L + \cdots + \widehat{\phi}_{p2}X_{t-p}^L \\
& \quad + \widehat{\theta}_{11}\widehat{u}_{t-1}^U + \cdots + \widehat{\theta}_{q1}\widehat{u}_{t-q}^U + \widehat{\theta}_{12}\widehat{u}_{t-1}^L + \cdots + \widehat{\theta}_{q2}\widehat{u}_{t-q}^L \geq \\
& \widehat{\phi}_{13}X_{t-1}^U + \cdots + \widehat{\phi}_{p3}X_{t-p}^U + \widehat{\phi}_{14}X_{t-1}^L + \cdots + \widehat{\phi}_{p4}X_{t-p}^L \\
& \quad + \widehat{\theta}_{13}\widehat{u}_{t-1}^U + \cdots + \widehat{\theta}_{q3}\widehat{u}_{t-q}^U + \widehat{\theta}_{14}\widehat{u}_{t-1}^L + \cdots + \widehat{\theta}_{q4}\widehat{u}_{t-q}^L.
\end{aligned}$$

Note que ambos os lados da inequação referente aos intervalos ajustados pelo modelo iARMA apresentam termos comuns: defasagens da série do limite superior e da série do limite inferior e defasagens dos resíduos do limite superior e dos resíduos do limite inferior superior. Estes resíduos são obtidos no primeiro estágio do algoritmo descrito na Seção 4.5.2. Ordenados ambos os lados da inequação acima segundo os termos comuns, notamos que eles são multiplicados por coeficientes distintos. Assim, uma forma de assegurar que no modelo iARMA os intervalos ajustados tenham sempre  $\widehat{X}_t^U \geq \widehat{X}_t^L$  é impor restrições aos parâmetros da forma

$$\begin{aligned}
\widehat{\phi}_{11} & \geq \widehat{\phi}_{13}, \\
& \vdots \\
\widehat{\phi}_{p1} & \geq \widehat{\phi}_{p3}, \\
\widehat{\phi}_{12} & \geq \widehat{\phi}_{14}, \\
& \vdots \\
\widehat{\phi}_{p2} & \geq \widehat{\phi}_{p4}, \\
\widehat{\theta}_{11} & \geq \widehat{\theta}_{13}, \\
& \vdots \\
\widehat{\theta}_{p1} & \geq \widehat{\theta}_{p3}, \\
\widehat{\theta}_{12} & \geq \widehat{\theta}_{14}, \\
& \vdots \\
\widehat{\theta}_{p2} & \geq \widehat{\theta}_{p4}.
\end{aligned} \tag{4.35}$$

As restrições acima podem ser reescritas na forma matricial como

$$C\hat{\boldsymbol{\pi}} = \begin{bmatrix} R_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & R_0 \end{bmatrix} \begin{bmatrix} \hat{\phi}_{11} \\ \hat{\phi}_{13} \\ \hat{\phi}_{12} \\ \hat{\phi}_{14} \\ \vdots \\ \hat{\phi}_{p1} \\ \hat{\phi}_{p3} \\ \hat{\phi}_{p2} \\ \hat{\phi}_{p4} \\ \hat{\theta}_{11} \\ \hat{\theta}_{13} \\ \hat{\theta}_{12} \\ \hat{\theta}_{14} \\ \vdots \\ \hat{\theta}_{q1} \\ \hat{\theta}_{q3} \\ \hat{\theta}_{q2} \\ \hat{\theta}_{q4} \end{bmatrix} \geq 0, \quad (4.36)$$

em que  $C$  é uma matriz  $(2(p+q) \times 4(p+q))$  conhecida e estabelecida pelo usuário,  $\hat{\boldsymbol{\pi}}$  é um vetor  $(4(p+q) \times 1)$  irrestrito de parâmetros desconhecidos dado por

$$\hat{\boldsymbol{\pi}} = \text{vec}(\hat{\boldsymbol{\Pi}}) = \text{vec}(\hat{\boldsymbol{\Phi}}_1, \dots, \hat{\boldsymbol{\Phi}}_p, \hat{\boldsymbol{\Theta}}_1, \dots, \hat{\boldsymbol{\Theta}}_q),$$

em que “vec” é o operador matemático que transforma uma matriz em um vetor,  $\mathbf{0}$  é uma matriz  $(2 \times 4)$  nula e  $R_0$  é a matriz  $(2 \times 4)$  responsável pela inserção de restrições aos parâmetros dada por

$$R_0 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Observe que a matriz  $C$  e a matriz  $R$  usada nos modelos iAR-restrito são construídas com a mesma matriz  $R_0$  e diferem-se apenas quanto à ordem;  $R$  tem ordem  $(2p \times 4p)$ , enquanto  $C$  tem ordem  $(2(p+q) \times 4(p+q))$ . Desta forma,  $C\hat{\boldsymbol{\pi}} \geq 0$  é uma generalização

das restrições impostas aos modelos iAR-restrito, pois se  $q = 0$ , as restrições  $C\hat{\boldsymbol{\pi}} \geq 0$  em (4.36) se reduzem a  $R\hat{\boldsymbol{\gamma}} \geq 0$  dadas em (4.34).

A estimação restrita de modelos iARMA para assegurar  $\hat{X}_t^U \geq \hat{X}_t^L$  também se dá em dois estágios inspirados no método de Hannan e Rissanen (1982) [49]. O primeiro estágio para estimação do modelo iARMA-restrito é idêntico ao primeiro estágio do modelo irrestrito, em que se estima um modelo irrestrito iAR( $l$ ), como descrito na Seção 4.5.2. O segundo estágio é análogo ao da estimação irrestrita, diferenciando no fato de que a função objetivo dada em (4.32) é agora minimizada sob imposição das restrições lineares de desigualdade  $C\hat{\boldsymbol{\pi}} \geq 0$ . O algoritmo adotado para estimação das matrizes de parâmetros do modelo iARMA-restrito está descrito a seguir.

No primeiro estágio da estimação do modelo iARMA-restrito, estima-se um modelo irrestrito iAR( $l$ ), com  $l = \frac{1}{2}\sqrt{n}$ , minimizando-se o critério

$$G(\phi_{ik}) = \sum_{t=1}^n \left( X_t^U - \sum_{i=1}^l \phi_{i1} X_{t-i}^U - \sum_{i=1}^l \phi_{i2} X_{t-i}^L \right)^2 + \sum_{t=1}^n \left( X_t^L - \sum_{i=1}^l \phi_{i4} X_{t-i}^L - \sum_{i=1}^l \phi_{i3} X_{t-i}^U \right)^2.$$

Então, computamos os resíduos estimados pelo modelo iAR( $l$ ),

$$\hat{\mathbf{u}}_t = \mathbf{I}_t - \sum_{i=1}^l \hat{\boldsymbol{\Phi}}_i^{(1)} \mathbf{I}_{t-i}.$$

Até então, nenhuma restrição é imposta.

No segundo estágio da estimação do modelo iARMA-restrito, os resíduos estimados  $\hat{\mathbf{u}}_t$  são inseridos no lugar de  $\mathbf{u}_t$  e estima-se o modelo iARMA( $p, q$ )-restrito através da minimização do critério

$$G(\phi_{ik}, \theta_{jk}) = \sum_{t=1}^n \left( X_t^U - \sum_{i=1}^p \phi_{i1} X_{t-i}^U - \sum_{j=1}^q \theta_{j1} \hat{u}_{t-j}^U - \sum_{i=1}^p \phi_{i2} X_{t-i}^L - \sum_{j=1}^q \theta_{j2} \hat{u}_{t-j}^L \right)^2 + \sum_{t=1}^n \left( X_t^L - \sum_{i=1}^p \phi_{i4} X_{t-i}^L - \sum_{j=1}^q \theta_{j4} \hat{u}_{t-j}^L - \sum_{i=1}^p \phi_{i3} X_{t-i}^U - \sum_{j=1}^q \theta_{j3} \hat{u}_{t-j}^U \right)^2,$$

sujeito às restrições lineares de desigualdade  $C\hat{\boldsymbol{\pi}} \geq 0$ . Formalmente, o segundo estágio do procedimento de estimação do modelo iARMA( $p, q$ )-restrito pode ser

descrito como

$$\begin{aligned} \text{Minimizar} \quad & G(\phi_{ik}, \theta_{jk}), \\ \text{sujeito a} \quad & C\hat{\boldsymbol{\pi}} \geq 0. \end{aligned}$$

Assim, as estimativas restritas das matrizes de parâmetros do modelo iARMA( $p, q$ )-restrito são obtidas.

O procedimento adotado para estimação restrita do modelo iARMA( $p, q$ ) sempre fornece estimativas das matrizes de coeficientes no espaço paramétrico tal que  $\hat{X}_t^U \geq \hat{X}_t^L$ , i.e., o intervalo previsto  $\hat{\mathbf{I}}_t$  respeita a condição  $\hat{X}_t^U \geq \hat{X}_t^L, \forall t$ . Chamamos este espaço paramétrico restrito de região factível (em inglês, *feasible region*). Esta estratégia reduz o espaço de soluções dos parâmetros em comparação ao método de estimação irrestrita.

**Exemplo 4.6** (Aplicação do modelo iARMA-restrito). Para ilustrar a modelagem de STI via modelos iARMA( $p, q$ )-restritos, voltamos ao exemplo dos preços diários de negociações de ações da Petrobras citada na Seção 4.2.1. Como no exemplo dos modelos irrestritos, nós ajustamos um modelo iAR(1) e um modelo iARMA(1,1), porém, neste exemplo, a estimação é condicionada a restrições lineares de desigualdade.

Para o modelo iAR(1)-restrito, a matriz de parâmetros estimada para a STI de preços de ações da Petrobras foi  $\hat{\boldsymbol{\Phi}}_1 = \begin{bmatrix} 0,522 & 0,491 \\ 0,497 & 0,490 \end{bmatrix}$ , resultando no modelo ajustado

$$\hat{\mathbf{I}}_t = \begin{bmatrix} 0,522X_{t-1}^U + 0,491X_{t-1}^L \\ 0,490X_{t-1}^L + 0,497X_{t-1}^U \end{bmatrix}.$$

A Figura 4.13 mostra os intervalos reais (segmentos de reta pretos) desta STI e os intervalos ajustados via modelo iAR(1)-restrito (segmentos de reta cinza).

As matrizes de parâmetros estimadas no modelo iARMA(1,1)-restrito para a STI de preços de ações da Petrobras foram  $\hat{\boldsymbol{\Phi}}_1 = \begin{bmatrix} 0,631 & 0,379 \\ 0,607 & 0,378 \end{bmatrix}$  e  $\hat{\boldsymbol{\Theta}}_1 = \begin{bmatrix} -0,033 & 0,207 \\ -0,033 & 0,207 \end{bmatrix}$  que implicam no modelo ajustado

$$\hat{\mathbf{I}}_t = \begin{bmatrix} 0,631X_{t-1}^U - 0,033u_{t-1}^U + 0,379X_{t-1}^L + 0,207u_{t-1}^L \\ 0,378X_{t-1}^L + 0,207u_{t-1}^L + 0,607X_{t-1}^U - 0,033u_{t-1}^U \end{bmatrix}.$$

No primeiro estágio da estimação foi ajustado um modelo iAR(4) para estimar os resíduos do modelo iARMA. A Figura 4.14 mostra os intervalos reais (segmentos de reta pretos) desta STI e os intervalos ajustados via modelo iARMA(1,1)-restrito (segmentos de reta cinza).

#### 4.5.4 Modelos iARIMA

Em análise de séries temporais univariadas, i.e., séries temporais de dados clássicos, os modelos ARMA são apropriados para séries temporais estacionárias. Assim, o ajuste de modelos ARMA pode ser mais adequado em séries onde não são observadas fontes de variações não-estacionárias. Algumas transformações nos dados podem ser aplicadas com a finalidade de eliminar tais fontes de variações e, então, gerar nova série estacionária. Por exemplo, se a série observada for não-estacionária na média, pode-se tentar remover a tendência tomando-se uma ou mais diferenças. Muito frequentemente, quando a série temporal é não-estacionária, a série resultante após algumas (uma ou duas) diferenças se torna estacionária. Tomar diferenças da série nos leva a considerar a classe de modelos *autorregressivos integrado médias móveis* (ARIMA).

Um modelo ARMA no qual a série original  $y_t$  é substituída pela sua  $d$ -ésima diferença  $\nabla^d y_t$  é capaz de descrever alguns tipos de séries não-estacionárias. A série diferenciada (integrada)  $d$  vezes é dada por

$$w_t = \nabla^d y_t = (1 - B)^d y_t, \quad (4.37)$$

em que  $B$  é o operador de retardo (*backward shift operator*) e definido como

$$B^d y_t = y_{t-d},$$

para todo  $d \geq 1$ . Especificamente, a diferença de ordem  $d = 1$  é dada por

$$\nabla y_t = y_t - y_{t-1} = (1 - B)y_t.$$

Polinômios em  $B$  e  $\nabla$  são manipulados da mesma forma como funções polinomiais de variáveis reais. Por exemplo,

$$\nabla^2 y_t = \nabla(\nabla y_t) = (1 - B)(1 - B)y_t = (1 - 2B + B^2)y_t = y_t - 2y_{t-1} + y_{t-2}.$$

Figura 4.15: Ajuste da STI de preços de ações da Petrobras via modelo iAR(1)-restrito.

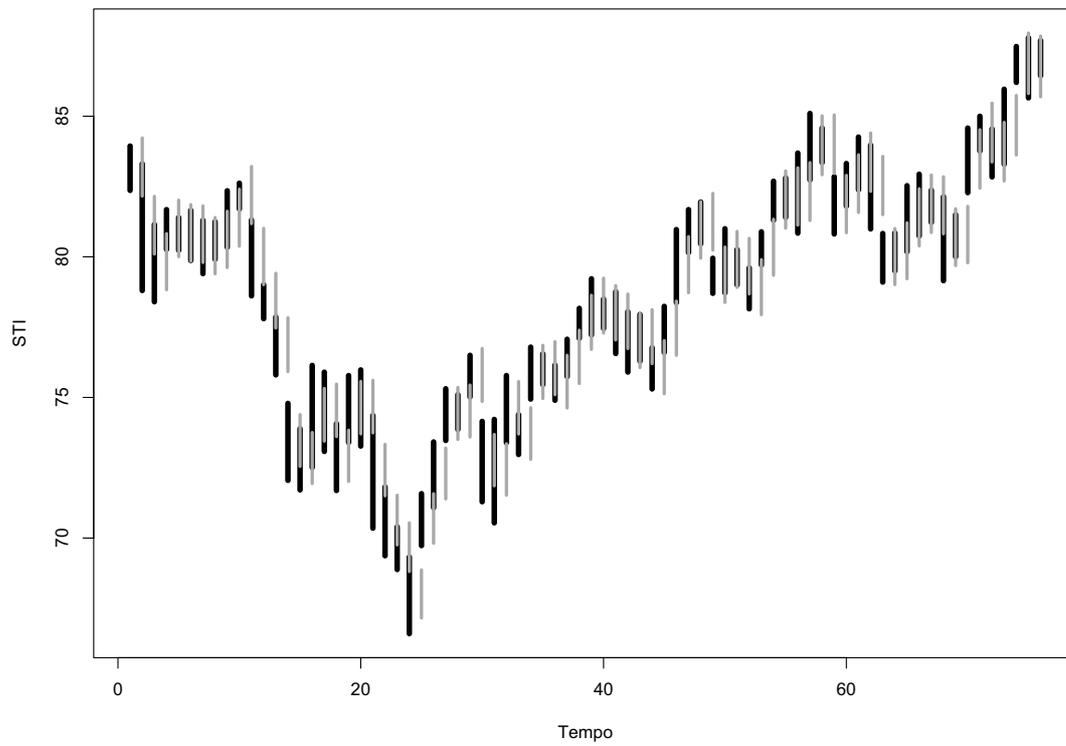
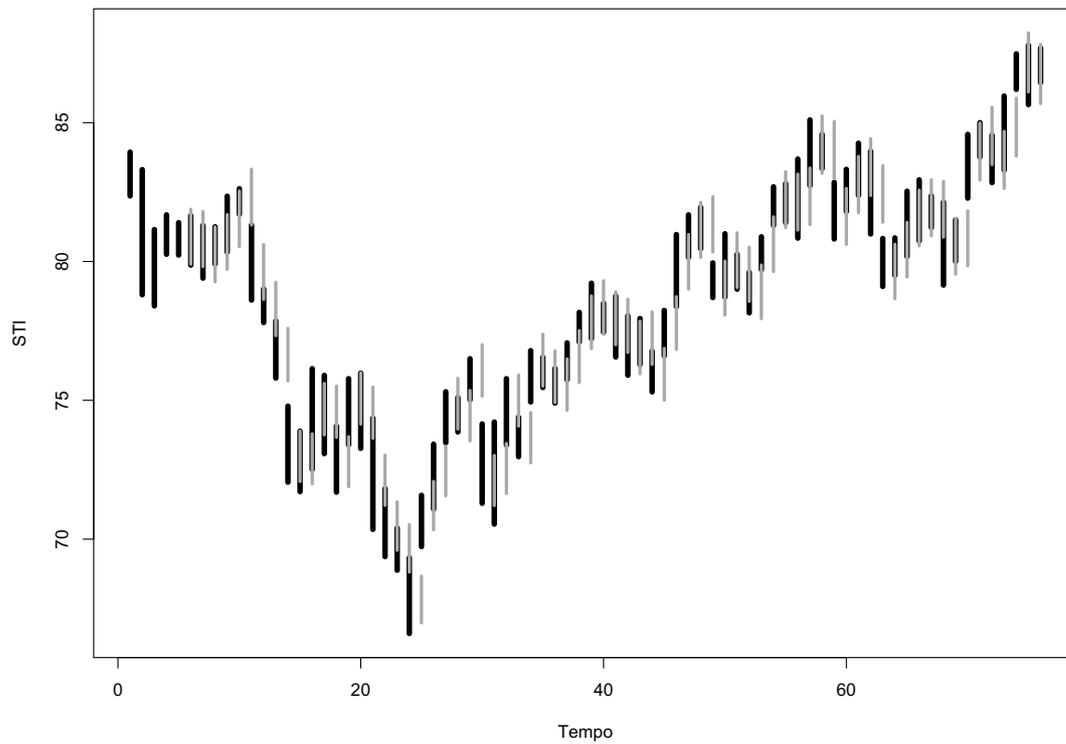


Figura 4.16: Ajuste da STI de preços de ações da Petrobras via modelo iARMA(1,1)-restrito.



Denotando  $w_t$  como em (4.37), o processo autorregressivo integrado médias móveis denotado  $\text{ARIMA}(p, d, q)$  é dado por

$$\begin{aligned} w_t &= \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \\ &= \sum_{i=1}^p \phi_i w_{t-i} + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}. \end{aligned}$$

O modelo  $\text{ARIMA}$  é similar ao modelo  $\text{ARMA}$ , no entanto, no primeiro modela-se a série diferenciada  $w_t$  tendo como regressores as defasagens desta série,  $w_{t-1}, \dots, w_{t-p}$ , além dos termos médias móveis.

Como definido em Brockwell e Davis (2002) [14] no contexto de séries temporais univariadas, se  $d$  é um inteiro não-negativo, então  $y_t$  é um processo  $\text{ARIMA}(p, d, q)$  se  $w_t = \nabla^d y_t$  é um processo  $\text{ARMA}(p, q)$ , i.e.,

$$\nabla^d y_t \text{ é um processo } \text{ARMA}(p, q) \implies y_t \text{ é um processo } \text{ARIMA}(p, d, q).$$

Vale salientar que os modelos  $\text{ARIMA}$  são uma generalização dos modelos  $\text{ARMA}$  e, em análise de dados univariados, são mais indicados para modelagem de séries temporais não-estacionárias, em que a operação de diferença as tornam estacionárias.

A extensão dos modelos  $\text{ARIMA}$  para modelar  $\text{STI}$  é obtida da mesma forma como os modelos  $\text{iARMA}(p, q)$ . No entanto, para construir modelos inspirados em modelos  $\text{ARIMA}$  univariados e capazes de modelar  $\text{STI}$ , faz-se necessário definir a *série intervalar diferenciada ou integrada*. Para definir  $\text{STI}$  diferenciada que denotamos por  $\mathbf{J}_t$ , nós adotamos a metodologia empregada nas extensões já apresentadas, em que o escalar  $y_t$  é substituído pelo intervalo  $\mathbf{I}_t$ . Então, definimos  $\mathbf{J}_t$  por

$$\mathbf{J}_t = \nabla^d \mathbf{I}_t = (1 - B)^d \mathbf{I}_t = \begin{bmatrix} (1 - B)^d X_t^U \\ (1 - B)^d X_t^L \end{bmatrix} = \begin{bmatrix} \nabla^d X_t^U \\ \nabla^d X_t^L \end{bmatrix} = \begin{bmatrix} w_t^U \\ w_t^L \end{bmatrix}$$

e o operador de retardo para  $\text{STI}$  por

$$B^d \mathbf{I}_t = \mathbf{I}_{t-d} = \begin{bmatrix} X_{t-d}^U \\ X_{t-d}^L \end{bmatrix}.$$

Assim, uma  $\text{STI}$  diferenciada é representada pelo vetor bidimensional com entradas dadas pelas séries diferenciadas dos limites superior e inferior, respectivamente,  $w_t^U$  e  $w_t^L$ . Em particular, denotamos a primeira diferença intervalar como

$$\nabla \mathbf{I}_t = \mathbf{I}_t - \mathbf{I}_{t-1} = \begin{bmatrix} X_t^U - X_{t-1}^U \\ X_t^L - X_{t-1}^L \end{bmatrix} = \begin{bmatrix} (1 - B)X_t^U \\ (1 - B)X_t^L \end{bmatrix}.$$

Se  $\mathbf{J}_t$  é uma diferença intervalar de  $\mathbf{I}_t$ , então,  $\mathbf{I}_t$  é denominado uma *integral* de  $\mathbf{J}_t$ , e dizemos que  $\mathbf{I}_t$  segue um processo *autorregressivo integrado médias móveis intervalar* (iARIMA). De acordo com as definições anteriores, o modelo ARIMA para modelagem de STI, ou modelo iARIMA( $p, d, q$ ), é dado por

$$\begin{aligned} \mathbf{J}_t &= \Phi_1 \mathbf{J}_{t-1} + \cdots + \Phi_q \mathbf{J}_{t-q} + \mathbf{u}_t + \Theta_1 \mathbf{u}_{t-1} + \cdots + \Theta_q \mathbf{u}_{t-q} \\ &= \sum_{i=1}^p \Phi_i \mathbf{J}_{t-i} + \mathbf{u}_t + \sum_{j=1}^q \Theta_j \mathbf{u}_{t-j}, \end{aligned} \quad (4.38)$$

em que  $\mathbf{u}_t = (u_t^U, u_t^L)^\top$  é um processo intervalar puramente aleatório com vetor de médias nulo e vetor de variância constante. Assim como as matrizes de parâmetros dos modelos iARMA( $p, q$ ), nos modelos iARIMA( $p, d, q$ ) as matrizes  $\Phi_i$  e  $\Theta_j$  são dadas por

$$\Phi_i = \begin{bmatrix} \phi_{i1} & \phi_{i2} \\ \phi_{i3} & \phi_{i4} \end{bmatrix} \quad \text{e} \quad \Theta_j = \begin{bmatrix} \theta_{j1} & \theta_{j2} \\ \theta_{j3} & \theta_{j4} \end{bmatrix}.$$

Expandindo a equação (4.38), temos

$$\begin{aligned} \mathbf{J}_t = \begin{bmatrix} w_t^U \\ w_t^L \end{bmatrix} &= \sum_{i=1}^p \begin{bmatrix} \phi_{i1} & \phi_{i2} \\ \phi_{i3} & \phi_{i4} \end{bmatrix} \begin{bmatrix} w_{t-i}^U \\ w_{t-i}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} + \sum_{j=1}^q \begin{bmatrix} \theta_{j1} & \theta_{j2} \\ \theta_{j3} & \theta_{j4} \end{bmatrix} \begin{bmatrix} u_{t-j}^U \\ u_{t-j}^L \end{bmatrix} \\ &= \sum_{i=1}^p \begin{bmatrix} \phi_{i1} w_{t-i}^U + \phi_{i2} w_{t-i}^L \\ \phi_{i3} w_{t-i}^U + \phi_{i4} w_{t-i}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix} + \sum_{j=1}^q \begin{bmatrix} \theta_{j1} u_{t-j}^U + \theta_{j2} u_{t-j}^L \\ \theta_{j3} u_{t-j}^U + \theta_{j4} u_{t-j}^L \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^p (\phi_{i1} w_{t-i}^U + \phi_{i2} w_{t-i}^L) + \sum_{j=1}^q (\theta_{j1} u_{t-j}^U + \theta_{j2} u_{t-j}^L) + u_t^U \\ \sum_{i=1}^p (\phi_{i3} w_{t-i}^U + \phi_{i4} w_{t-i}^L) + \sum_{j=1}^q (\theta_{j3} u_{t-j}^U + \theta_{j4} u_{t-j}^L) + u_t^L \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^p \phi_{i1} w_{t-i}^U + u_t^U + \sum_{j=1}^q \theta_{j1} u_{t-j}^U + \sum_{i=1}^p \phi_{i2} w_{t-i}^L + \sum_{j=1}^q \theta_{j2} u_{t-j}^L \\ \sum_{i=1}^p \phi_{i3} w_{t-i}^U + u_t^L + \sum_{j=1}^q \theta_{j3} u_{t-j}^U + \sum_{i=1}^p \phi_{i4} w_{t-i}^L + \sum_{j=1}^q \theta_{j4} u_{t-j}^L \end{bmatrix}. \end{aligned} \quad (4.39)$$

■

Portanto, o modelo (4.39) supõe que a  $d$ -ésima diferença da série intervalar

$\mathbf{I}_t$  pode ser representada por um modelo iARMA. Assim, se  $d$  é um inteiro não-negativo, então  $\mathbf{I}_t$  é um processo intervalar iARIMA( $p, d, q$ ) se  $\mathbf{J}_t = \nabla^d \mathbf{I}_t$  é um processo iARMA( $p, q$ ), i.e.,

$$\nabla^d \mathbf{I}_t \text{ é um processo iARMA}(p, q) \implies \mathbf{I}_t \text{ é um processo iARIMA}(p, d, q).$$

Neste ponto devemos atentar para um detalhe dos modelos iARIMA. A ordem de relação dos limites do intervalo garante que  $X_t^U \geq X_t^L$ . No entanto, nos modelos iARIMA modela-se a série das diferenças  $\mathbf{J}_t = (w_t^U, w_t^L)^\top$  que não mantém esta coerência matemática dos intervalos, ou seja,  $w_t^U$  não necessariamente é maior que  $w_t^L$ , para todo  $t$ . Por exemplo, dados dois intervalos  $\mathbf{I}_1 = (18, 10)^\top$  e  $\mathbf{I}_2 = (30, 25)^\top$ , a primeira diferença é  $\mathbf{J}_2 = \mathbf{I}_2 - \mathbf{I}_1 = (12, 15)^\top$ . Neste caso,  $w_2^U < w_2^L$ . Isso tem uma implicação importante para o processo de estimação de modelos iARIMA.

A condição de inexistência da relação de ordem na série das diferenças  $\mathbf{J}_t = (w_t^U, w_t^L)^\top$  implica que o método empregado para estimação das matrizes de parâmetros dos modelos iARMA sem levar em consideração as restrições pode ser aplicado diretamente para estimação das matrizes de parâmetros dos modelos iARIMA. Por esta razão, não dedicamos uma nova seção para apresentar o método de estimação de modelos iARIMA, pois pode ser consultado na Seção 4.5.2.

Como pode ser visto, os modelos estendidos nesta seção para previsão de STI são inspirados nos modelos ARIMA para dados univariados. Resumidamente, aplicar os modelos iARIMA( $p, d, q$ ) a STI consiste em aplicar modelos iARMA( $p, q$ ) a esta STI transformada, i.e., após tomada a  $d$ -ésima diferença da série intervalar  $\mathbf{I}_t$ . Esta característica torna os modelos iARIMA( $p, d, q$ ) distintos dos modelos iARMA( $p, q$ ), porém, se  $d = 0$ , o modelo iARIMA( $p, d, q$ ) se reduz ao modelo iARMA( $p, q$ ).

Nas duas próximas seções, construímos métodos alternativos para previsão de STI baseados em redes neurais artificiais. O propósito da adaptação destes métodos para previsão de STI é proporcionar uma alternativa eficiente e capaz de modelar séries temporais com características complexas como, por exemplo, relação funcional não-linear entre o valor atual e os valores passados da série.

## 4.6 Métodos baseados em redes neurais para STI

Um tópico popular em análise de dados moderna é redes neurais artificiais (RNA) e muito já se sabe sobre sua capacidade na solução de problemas de aproximação de funções escalares. Elas têm sido frequentemente aplicadas e investigadas em problemas práticos de vários domínios, incluindo previsão de séries temporais. Se tratando de previsão de séries temporais, RNA podem ser úteis para modelagem de processos não-lineares que apresentam uma relação funcional desconhecida e, conseqüentemente, são difíceis de se ajustar (Darbellay e Slama, 2000) [24]. A ideia geral da modelagem de dados através de RNA é que fatores de entrada, ou variáveis dependentes, são filtrados através de uma ou mais camadas escondidas compostas de unidades de processamento, ou nodos, antes de chegar à variável de saída.

O uso de RNA, seja para previsão ou para qualquer outro tipo de aplicação, envolve um conjunto de etapas, tais como: análise exploratória dos dados, o pré-processamento dos dados, a escolha do modelo de redes neurais, o treinamento da rede e a sua validação. Com relação à escolha do modelo, existem vários tipos de RNA adequados para previsão de séries temporais como, por exemplo, redes função de base radial (RBF) ou redes recorrentes, embora a maioria das aplicações envolvam as redes neurais do tipo *perceptron* de múltiplas camadas (MLP).

Nesta tese nos detemos aos modelos MLP não apenas pelo fato desta rede ser a mais utilizada em problemas de previsão de séries temporais, mas também devido à simplicidade matemática e computacional destes modelos. Muitos programas, dentre os quais podemos citar o R (R Development Core Team, 2006) [78], já apresentam pacotes implementados para modelagem de dados via redes neurais MLP e, através de simples manipulações nestes pacotes, podemos adaptar as redes neurais e torná-las capazes de modelar e prever STI. Além disso, a metodologia empregada para adaptar as redes MLP mostra que qualquer outra RNA pode ser adaptada para previsão de STI.

Com base na teoria das RNA, nós apresentamos duas metodologias para previsão de STI em que as redes MLP são utilizadas. O primeiro método a ser apresentado consiste de uma MLP, método comumente utilizado em previsão de séries temporais univariadas. O segundo método é uma adaptação da metodologia híbrida proposta por Zhang (2003) [94] para séries temporais univariadas. Nesta adaptação, nós

combinamos o modelo iARMA com a rede neural MLP construída para previsão de STI.

Os métodos para previsão de STI que serão apresentados a seguir diferem dos métodos apresentados na Seção 3.2 devido à metodologia adotada conforme descrita na Seção 4.1.

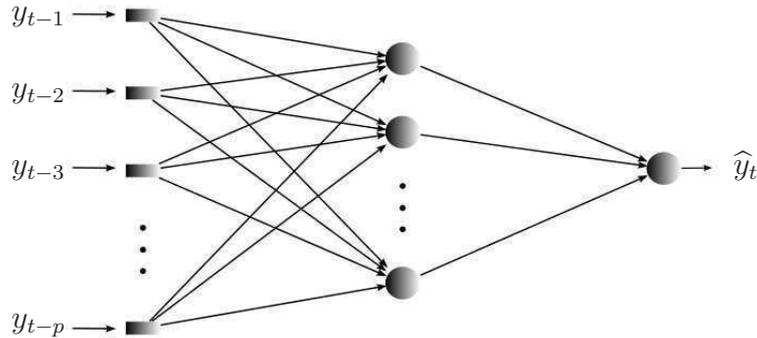
#### 4.6.1 Redes neurais MLP para STI

Redes neurais artificiais com arquitetura do tipo *perceptron* de múltiplas camadas, ou *multilayer perceptron*, são frequentemente usadas para processamento de séries temporais de forma similar aos modelos autorregressivos (AR) descritos na Seção 4.5.1. Em especial, redes MLP com duas camadas – uma camada escondida e uma camada de saída – são frequentemente usadas para previsão de séries temporais. Redes MLP com uma única camada escondida são capazes de representar qualquer função contínua (ver, e.g., Russell e Norvig, 1995) [81]. Em geral, as redes MLP têm camadas completamente conectadas, i.e., um nodo em qualquer camada da rede está conectado a todos os nodos da camada anterior (Haykin, 1999) [51].

A estrutura de uma rede neural MLP frequentemente utilizada para previsão de séries temporais univariadas está mostrada na Figura 4.17. Esta MLP consiste de duas camadas compostas por unidades de processamento, também denominadas neurônios ou nodos. Os valores de entrada (dados de entrada) alimentam a rede neural na parte esquerda da Figura 4.17, onde são passados para diante aos nodos da camada escondida. É importante destacar que a entrada da rede, na verdade, não é formada por neurônios reais, pois eles não realizam nenhum processamento; simplesmente distribuem os valores das entradas da rede para os neurônios da camada escondida. As setas que tocam a camada escondida na figura indicam conexões entre as entradas e os nodos escondidos e entre estes nodos e o nodo de saída. Ao longo das conexões os valores das entradas são passados para os nodos escondidos e, por sua vez, os valores obtidos como entradas para os nodos escondidos são processados novamente dentro deles e passados para a camada de saída. Finalmente, no nodo de saída, as informações provenientes da camada escondida são combinadas e a previsão é fornecida. Também estão presentes nas redes neurais MLP unidades especiais conhecidas por *bias* que estão ligadas a todos os nodos da

camada escondida e da saída da rede. O peso das ligações das conexões *bias* tem finalidade similar ao intercepto em modelos de regressão. Os *bias* foram omitidos da Figura 4.17.

Figura 4.17: Estrutura de uma rede neural MLP com  $p$  entradas, uma camada escondida com  $q$  nodos e uma saída para previsão de séries temporais univariadas.



Cada conexão da rede MLP tem um parâmetro associado denominado peso que indica a intensidade desta conexão. Através de atualização dos pesos usando técnicas específicas, a rede neural é capaz de aprender com os padrões apresentados nas entradas e com base nos valores da variável resposta. A descrição do procedimento no qual esta adaptação de pesos é realizada é chamada de aprendizagem ou algoritmo de treinamento.

Na rede neural MLP com uma única camada escondida para previsão de séries temporais, temos que o  $j$ -ésimo nodo na camada escondida é definido como

$$h_j = g \left( \beta_{0j} + \sum_{i=1}^p \beta_{ij} y_{t-i} \right),$$

em que  $y_{t-i}$  é o valor da  $i$ -ésima entrada,  $g$  é uma função de ativação tipicamente não-linear,  $\beta_{0j}$  é o *bias* e  $\beta_{ij}$  representa o peso da conexão entre a  $i$ -ésima entrada e o  $j$ -ésimo nodo na camada escondida.

O nodo da camada de saída com função de ativação linear é definido como

$$o = \alpha_0 + \sum_{j=1}^q \alpha_j h_j,$$

e que  $\alpha_0$  é o *bias* ligado ao nodo de saída e  $\alpha_j$  é o peso da conexão entre o  $j$ -ésimo nodo na camada escondida e o nodo de saída da rede.

Combinando as entradas, a camada escondida e o nodo de saída da rede MLP, a relação entre o valor atual da série  $y_t$ , e os valores passados da série  $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ , é dada como

$$\hat{y}_t = \alpha_0 + \sum_{j=1}^q \alpha_j \cdot g \left( \beta_{0j} + \sum_{i=1}^p \beta_{ij} y_{t-i} \right),$$

em que  $\alpha$ 's e  $\beta$ 's são os pesos (parâmetros do modelo);  $p$  é o número de entrada da rede;  $q$  é o número de nodos na camada escondida; e  $g$  é a função de ativação nos nodos da camada escondida. Existem vários tipos de funções de ativação, mas a considerada nesta MLP é a função logística,

$$g(u) = \frac{1}{1 + \exp(-u)}. \quad (4.40)$$

Note que  $g$  é uma função não-linear. A maior vantagem desta função é sua derivada de fácil obtenção. A derivada da função de ativação será necessária no processo de treinamento da rede neural (Kaastra e Boyd, 1996) [56].

Observe que a rede neural MLP com uma única camada escondida tem estrutura similar ao modelo clássico AR, porém pode ser mais poderosa devido ao fato de realizar um mapeamento não-linear de valores passados da série para os valores futuros. Assim, podemos considerar uma rede neural deste tipo como um modelo AR não-linear.

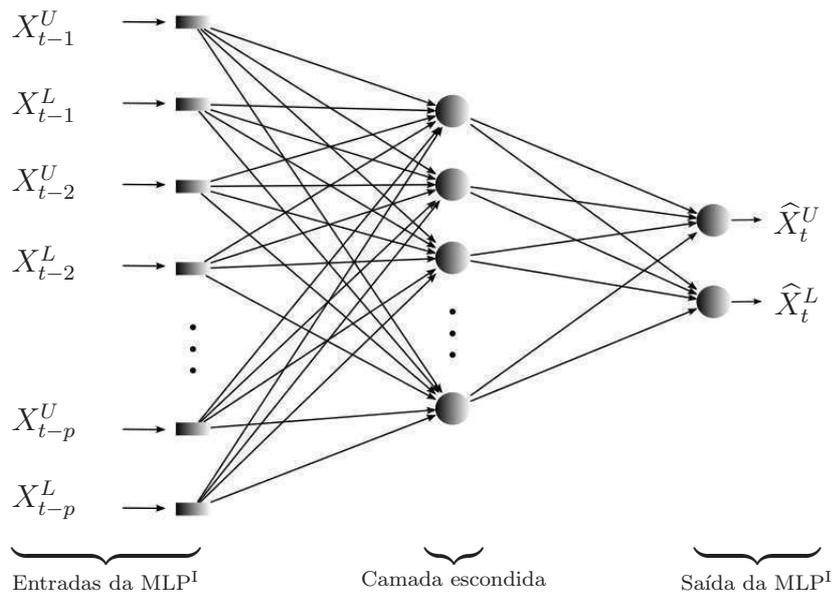
Uma grande vantagem dos modelos de redes neurais sobre os modelos lineares é que eles são capazes de aproximar várias formas de não-linearidade presentes nos dados e, ainda, não requer qualquer pressuposição sobre a forma do modelo. Esta característica de não-linearidade é muito importante visto que muitos problemas reais são de natureza não-linear. De modo geral, as redes neurais são aproximadores universais de funções com alto grau de precisão (Cybenko, 1989) [23].

Para o processamento e previsão de STI, usaremos uma rede MLP semelhante à apresentada anteriormente. Designamos a rede neural por *perceptron* de múltiplas camadas para previsão de séries intervalares, ou simplesmente MLP<sup>I</sup>. A principal diferença com relação à rede MLP para séries univariadas apresentada antes está no uso de dois nodos de saída na MLP<sup>I</sup>, cada um responsável pela previsão de um limite da série intervalar,  $X_t^L$  e  $X_t^U$ .

Baseado nas redes MLP tipicamente utilizadas para previsão de séries temporais univariadas, sem perda de generalização, configuramos as redes MLP<sup>I</sup> com apenas

uma camada escondida. A Figura 4.18 apresenta a arquitetura da MLP<sup>I</sup> utilizada para previsão de STI. Como pode ser visto na figura, na rede neural MLP<sup>I</sup>, o número de nodos da camada de entrada é  $2p$ , indicado pelo número de intervalos defasados, i.e.,  $\{\mathbf{I}_{t-1}, \mathbf{I}_{t-2}, \dots, \mathbf{I}_{t-p}\} = \{(X_{t-1}^U, X_{t-1}^L)^\top, (X_{t-2}^U, X_{t-2}^L)^\top, \dots, (X_{t-p}^U, X_{t-p}^L)^\top\}$ . Observe que são  $2p$  nodos de entrada, pois, para cada um dos  $p$  intervalos passados correspondem dois valores de entrada: o valor do limite superior e o valor do limite inferior. Como função de ativação na camada escondida, usamos a função logística dada em (4.40). As setas nesta figura indicam a direção de processamento dos dados. As conexões *bias* foram omitidas na Figura 4.18. Como nos modelos apresentados antes, na rede MLP<sup>I</sup> a previsão de um dos limites leva em consideração informações do outro limite.

Figura 4.18: Estrutura da rede neural MLP<sup>I</sup> com  $2p$  entradas, uma camada escondida com  $q$  nodos e duas saídas para previsão de STI.



A regra de treinamento frequentemente utilizada nas redes neurais MLP é a de retropropagação de erros (*error backpropagation*). A ideia deste algoritmo é atualizar os pesos utilizando as derivadas dos erros em relação aos pesos. Para detalhes sobre o algoritmo de retropropagação de erros, ver Rumelhart et al. (1986) [80].

Na rede MLP<sup>I</sup> com uma camada escondida com  $q$  nodos e dois nodos na camada

de saída, sejam  $\varphi_{ij}$  e  $\omega_j$  os vetores bidimensionais de pesos

$$\varphi_{ij} = \begin{bmatrix} \varphi_{ij}^U \\ \varphi_{ij}^L \end{bmatrix} \quad \text{e} \quad \omega_j = \begin{bmatrix} \omega_j^U \\ \omega_j^L \end{bmatrix},$$

a relação entre os dois nodos de saída da rede MLP<sup>I</sup> e as entradas,  $\mathbf{I}_{t-1}$ ,  $\mathbf{I}_{t-2}$ , ...,  $\mathbf{I}_{t-p}$ , é expressa como

$$\begin{aligned} \hat{\mathbf{I}}_t = \begin{bmatrix} \hat{X}_t^U \\ \hat{X}_t^L \end{bmatrix} &= \omega_0 + \sum_{j=1}^q \omega_j \odot g \left( \begin{bmatrix} \varphi_{0j}^U + \sum_{i=1}^p (\varphi_{ij})^\top \mathbf{I}_{t-i} \\ \varphi_{0j}^L + \sum_{i=1}^p (\varphi_{ij})^\top \mathbf{I}_{t-i} \end{bmatrix} \right) \\ &= \begin{bmatrix} \omega_0^U + \sum_{j=1}^q \omega_j^U \cdot g \left( \varphi_{0j}^U + \sum_{i=1}^p (\varphi_{ij})^\top \mathbf{I}_{t-i} \right) \\ \omega_0^L + \sum_{j=1}^q \omega_j^L \cdot g \left( \varphi_{0j}^L + \sum_{i=1}^p (\varphi_{ij})^\top \mathbf{I}_{t-i} \right) \end{bmatrix} \\ &= \begin{bmatrix} \omega_0^U + \sum_{j=1}^q \omega_j^U \cdot g \left( \varphi_{0j}^U + \sum_{i=1}^p (\varphi_{ij}^U X_{t-i}^U + \varphi_{ij}^L X_{t-i}^L) \right) \\ \omega_0^L + \sum_{j=1}^q \omega_j^L \cdot g \left( \varphi_{0j}^L + \sum_{i=1}^p (\varphi_{ij}^U X_{t-i}^U + \varphi_{ij}^L X_{t-i}^L) \right) \end{bmatrix}, \end{aligned}$$

em que os parâmetros  $\varphi^U$ 's e  $\varphi^L$ 's são os pesos das conexões entre o vetor de entrada (intervalos passados) e os nodos da camada escondida, os  $\omega^U$ 's são os pesos das conexões da camada escondida para a camada de saída correspondente à previsão do limite superior e os  $\omega^L$ 's são os pesos das conexões da camada escondida para a saída correspondente à previsão do limite inferior. O símbolo  $\odot$  na equação acima denota o *produto de Hardamard*.<sup>9</sup>

Os mecanismos utilizados para seleção do número de entradas e do número de nodos escondidos nas redes MLP para dados univariados podem ser utilizados para configurar as redes MLP<sup>I</sup>. Podemos usar técnicas de otimização, e.g., minimização de critério como nos modelos anteriores. A tarefa de encontrar o número de entradas utilizado para prever intervalos futuros é mais fácil, pois os próprios dados podem indicar este número. Por exemplo, o número de nodos de entrada que define o

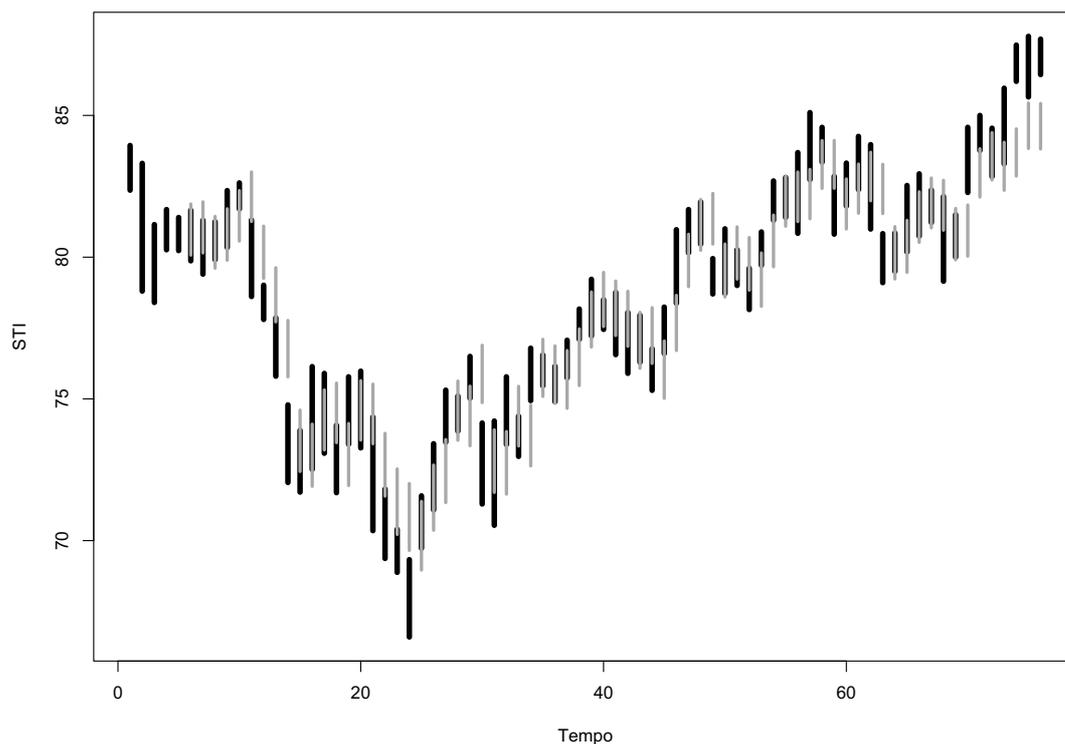
---

<sup>9</sup>O *produto de Hardamard* consiste da multiplicação elemento-por-elemento de dois vetores ou duas matrizes de mesma dimensão; e.g.,  $\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \odot \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} \\ a_{21}b_{21} \end{bmatrix}$ .

número de intervalos passados pode ser definido através do exame da correlação serial intervalar a partir da função de autocorrelação amostral  $ACF^I$  apresentada na Seção 4.2.

**Exemplo 4.7** (Aplicação da rede neural  $MLP^I$ ). Para ilustrar a aplicação do modelo  $MLP^I$ , ajustamos a STI de preços de ações da Petrobras. A série intervalar real (segmentos de reta pretos) e os intervalos ajustados através da rede neural  $MLP^I$  (segmentos de reta cinza) estão apresentadas na Figura 4.19. A rede  $MLP^I$  foi treinada com dez entradas correspondentes à cinco cinco defasagens intervalares:  $I_{t-1}$ ,  $I_{t-2}$ ,  $I_{t-3}$ ,  $I_{t-4}$  e  $I_{t-5}$ . Na camada escondida foram usados oito nodos todos com função de ativação logística e na camada de saída dois nodos com função linear.

Figura 4.19: Ajuste da STI de preços de ações da Petrobras via rede neural  $MLP^I$ .



Na teoria, redes neurais artificiais são capazes de aproximar qualquer função contínua. Tal flexibilidade fornece uma ferramenta potencialmente poderosa para previsão de séries temporais, enquanto ao mesmo tempo, o número elevado de parâmetros a ser ajustado torna o processo de seleção da rede uma tarefa não

muito simples. Na prática, a tarefa de seleção do melhor formato da rede envolve métodos de tentativa e erro (Kaastra e Boyd, 1996) [56].

### 4.6.2 Modelo híbrido para STI

A metodologia híbrida consiste basicamente em combinar um modelo intervalar linear com uma rede neural MLP. O modelo híbrido apresentado trata-se da combinação do método linear AED<sup>I</sup> para STI com a rede neural MLP<sup>I</sup>. No entanto, qualquer um dos métodos lineares para STI apresentados – AES<sup>I</sup>, HW<sup>I</sup>, iAR, iARMA ou as versões restritas dos modelos iAR e iARMA – pode ser utilizado. A metodologia será descrita a seguir.

Algumas questões nos motivam a considerar a metodologia híbrida. Primeiro, na prática é muito difícil determinar quando uma série temporal é gerada por um processo linear ou não-linear ou quando um método particular é mais eficiente que um outro para prever a série. Geralmente, muitos modelos são ajustados e o que fornece resultados mais precisos no conjunto de treinamento é selecionado para prever valores futuros da série. Entretanto, o modelo selecionado não é necessariamente o melhor modelo para uso futuro devido a influência de outros fatores, por exemplo, variações amostrais, incerteza do modelo e mudança na estrutura da série. A segunda razão para uso de um modelo híbrido é que séries temporais reais raramente são processos lineares ou não-lineares puros, frequentemente, contêm ambos os padrões. Terceiro, na literatura de previsão de séries temporais concorda-se que não existe um método melhor em todas as situações. Através de combinações de modelos lineares com modelos de RNA, estruturas de autocorrelações complexas nos dados podem ser modeladas mais precisamente.

Muitos estudos empíricos sugerem que através de combinação de diferentes modelos o desempenho das previsões pode ser melhor que o desempenho de cada um dos modelos individualmente. Neste contexto, a ideia básica da combinação de modelos para previsão de séries temporais é que cada modelo individual seja capaz de capturar diferentes padrões nos dados.

A literatura sobre metodologia híbrida envolvendo combinação de modelos lineares e redes neurais para previsão de séries temporais é ampla. Por exemplo, Su et al. (1997) [86] usaram modelo híbrido para prever séries com tendência crescente.

Seus resultados mostraram que o modelo híbrido fornece melhores previsões que os modelos ARMA ou que as redes neurais individualmente. Wedding e Cios (1996) [91] exploraram um sistema híbrido envolvendo redes do tipo função de base radial (RBF) e modelos tradicionais ARMA e conseguiram melhorar a precisão de previsões de algumas séries temporais. Hansen e Nelson (1997) [50] obtiveram sucesso na combinação de redes neurais com modelos convencionais de séries temporais para previsão de renda para o estado de Utah, EUA. Em vez de usar diretamente redes neurais para previsão, Tian et al. (1997) [88] usaram uma rede neural recorrente para estimar os parâmetros de modelos AR. Outros estudos podem ser vistos em Tseng et al. (2002) [89] e Taskaya et al. (2005) [87].

Inicialmente, introduzimos a metodologia proposta por Zhang (2003) [94] para previsão de séries temporais univariadas que consiste do hibridismo dos modelos ARIMA e RNA. Em seguida, apresentamos uma metodologia híbrida inspirada na proposta de Zhang (2003) [94], porém usando os modelos apresentados para previsão de séries temporais do tipo intervalo.

Conforme já descrito no Capítulo 3, para justificar a metodologia híbrida que combina modelo ARIMA com redes neurais MLP, Zhang (2003) [94] considerou que uma série temporal univariada é composta de uma estrutura de autocorrelação linear e uma componente não-linear,

$$y_t = \ell_t + \eta_t,$$

em que  $\ell_t$  e  $\eta_t$  denotam, respectivamente, a componente linear e a componente não-linear. Então, Zhang (2003) [94] desenvolveu um modelo híbrido que consiste de duas etapas. Para modelar  $y_t$ , primeiramente, usa-se o modelo ARIMA para modelar a componente linear. Com isso, os resíduos do modelo ARIMA deverão conter informações sobre a não-linearidade da série

$$e_t = y_t - \hat{\ell}_t.$$

Após ajuste do modelo ARIMA, os resíduos são modelados através da rede neural MLP com a finalidade de capturar a relação não-linear da série usando  $k$  valores de entrada,

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-k}) + \epsilon_t.$$

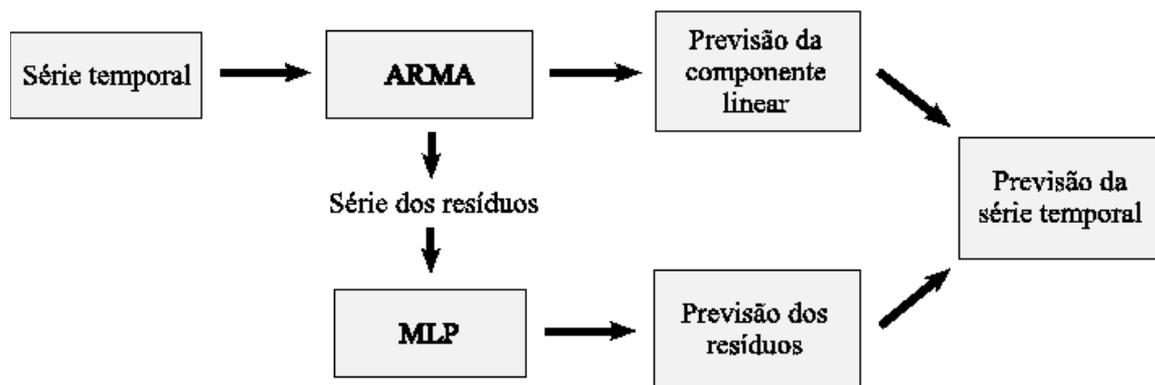
Denota-se a previsão do resíduo,  $e_t$ , através da rede MLP como  $\hat{\eta}_t$ . Desta forma, a previsão combinada fornecida pelo modelo híbrido será dada por

$$\hat{y}_t = \hat{\ell}_t + \hat{\eta}_t.$$

Note que esta metodologia não requer quaisquer pressuposições sobre a estrutura de correlação da série temporal.

A Figura 4.20 mostra o diagrama do modelo híbrido adotado em Zhang (2003) [94], em que o modelo ARIMA é usado na primeira etapa para ajuste da componente linear e a rede neural MLP é usada para ajustar os resíduos. Neste diagrama, podemos ver claramente as duas etapas do modelo.

Figura 4.20: Diagrama do modelo híbrido de Zhang (2003) para previsão de séries temporais univariadas.



Segundo Zhang (2003) [94], em problemas complexos, em que a série temporal apresenta estrutura de correlação linear e não-linear, o método híbrido proposto pode ser um método eficiente para melhorar a precisão das previsões. Em seus experimentos, ele concluiu que o modelo híbrido apresenta desempenho superior aos modelos usados individualmente.

Baseado na ideia de Zhang (2003) [94], nós estendemos a aplicação da metodologia híbrida para previsão de STI. Diferente da adaptação do modelo híbrido apresentada no Capítulo 3, o modelo que apresentamos aqui emprega a metodologia de substituição do escalar  $y_t$  pelo vetor intervalo  $\mathbf{I}_t = [X_t^L; X_t^U]$ . Com isso, nosso modelo híbrido corrente está inserido na classe de métodos para os limites superior e inferior dos intervalos.

Para estender o modelo híbrido e torná-lo útil na previsão de STI, consideramos que o vetor  $\mathbf{I}_t$  é composto de componente linear dada pelo vetor  $\boldsymbol{\ell}_t$  e de componente não-linear dada pelo vetor  $\boldsymbol{\eta}_t$ , i.e.,

$$\mathbf{I}_t = \boldsymbol{\ell}_t + \boldsymbol{\eta}_t = \begin{bmatrix} \ell_t^U \\ \ell_t^L \end{bmatrix} + \begin{bmatrix} \eta_t^U \\ \eta_t^L \end{bmatrix}.$$

Agora, os modelos que serão considerados para formar o modelo híbrido intervalar são os modelos propostos para STI. Qualquer modelo linear que nós apresentamos para previsão de STI que fazem parte dos métodos para os limites superior e inferior dos intervalos pode ser considerado para modelar a componente linear da STI,  $\boldsymbol{\ell}_t$ , e a rede neural MLP<sup>I</sup> para modelar as séries dos resíduos no limite inferior e no limite superior,

$$\widehat{\boldsymbol{\eta}}_t = \mathbf{e}_t = \mathbf{I}_t - \widehat{\boldsymbol{\ell}}_t = \begin{bmatrix} X_t^U - \widehat{\ell}_t^U \\ X_t^L - \widehat{\ell}_t^L \end{bmatrix},$$

em que  $\widehat{\ell}_t^U$  e  $\widehat{\ell}_t^L$  são as previsões obtidas pelo modelo linear, respectivamente, para os limites superior e inferior. Assim, as séries resultantes dos resíduos no limite superior e no limite inferior contêm informações sobre a não-linearidade da STI. Após ajuste do modelo linear, a série bidimensional de resíduos  $\mathbf{e}_t$  é modelada através da rede neural MLP<sup>I</sup>,

$$\mathbf{e}_t = f(\mathbf{e}_{t-1}, \mathbf{e}_{t-2}, \dots, \mathbf{e}_{t-k}) + \boldsymbol{\epsilon}_t,$$

com a finalidade de capturar a relação não-linear da STI usando  $2k$  entradas. Aqui,  $f$  é uma função não-linear determinada pela rede neural MLP<sup>I</sup>.

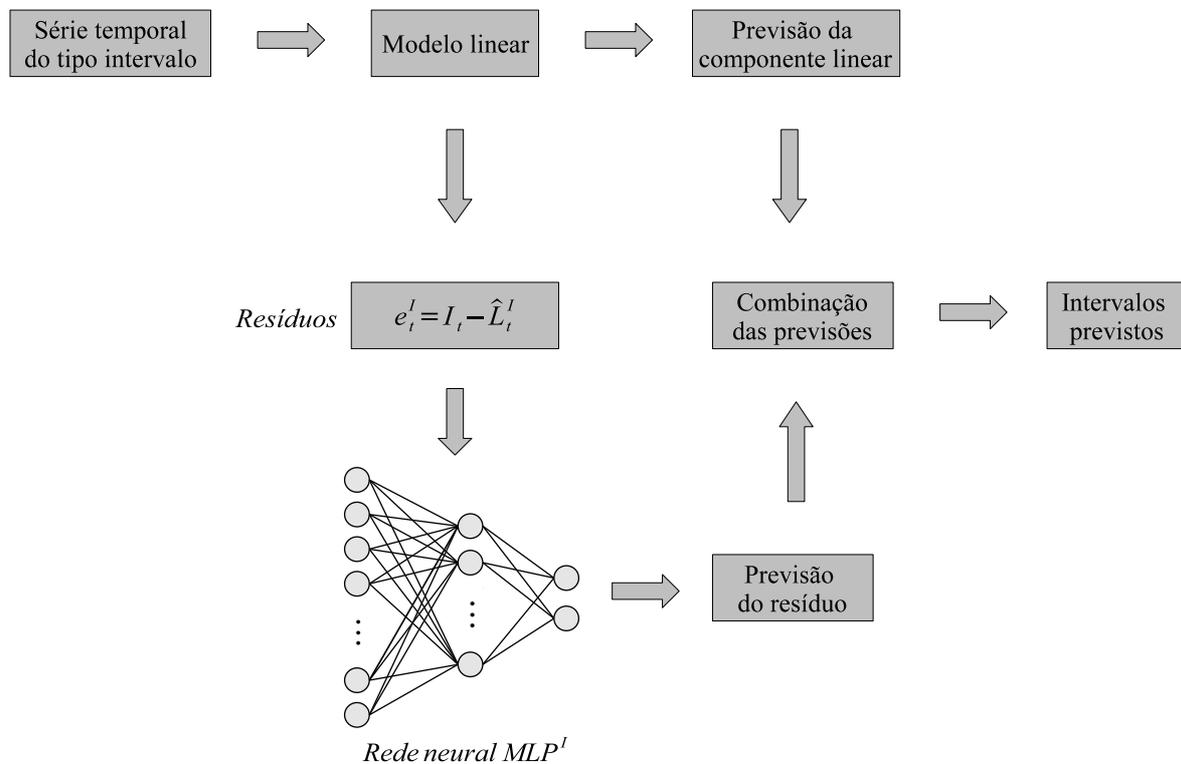
O intervalo previsto através do modelo híbrido é, então, dado por

$$\widehat{\mathbf{I}}_t = \widehat{\boldsymbol{\ell}}_t + \widehat{\boldsymbol{\eta}}_t = \begin{bmatrix} \widehat{\ell}_t^U + \widehat{\eta}_t^U \\ \widehat{\ell}_t^L + \widehat{\eta}_t^L \end{bmatrix}.$$

Como dito antes, os modelos AES<sup>I</sup>, AED<sup>I</sup>, HW<sup>I</sup>, iAR, iARMA e as versões restritas dos dois últimos podem ser adotados como modelo linear na primeira etapa do modelo híbrido intervalar. A Figura 4.21 mostra o diagrama do modelo híbrido considerando os modelos propostos para previsão de STI.

A metodologia híbrida apresentada não requer quaisquer pressuposições sobre a estrutura de correlação da série temporal. Além disso, pelo fato de combinar dois

Figura 4.21: Diagrama do modelo híbrido para previsão de STI.



modelos com relações funcionais distintas, ela pode ser vantajosa sobre os modelos lineares ou não-lineares ajustados separadamente.

## 4.7 Medidas de precisão de previsões para STI

A tarefa de avaliação da precisão das previsões fornecidas pelos modelos é fundamental na análise de séries temporais. Esta etapa consiste em avaliar o desempenho preditivo de modelos através da comparação dos valores ajustados e previstos pelo modelo com os dados reais da série. Na prática, esta comparação indica, dentre os métodos utilizados, qual o método é mais preciso na previsão de uma mesma série temporal.

Nesta seção, apresentamos algumas medidas baseadas nos resíduos dos modelos comumente utilizadas na avaliação de métodos de previsão, tais como, o *erro quadrático médio* (MSE), a *estatística U de Theil* (U) e a *variância relativa média* (ARV) e as representações estendidas para avaliação de modelos de previsão de séries temporais do tipo intervalo.

Em dados univariados, o MSE, a estatística U de Theil e a ARV são definidos, respectivamente, através das medidas

$$\begin{aligned} \text{MSE} &= \frac{1}{m} \sum_{j=1}^m (y_j - \hat{y}_j)^2, \\ U &= \sqrt{\frac{\sum_{j=1}^m (y_{j+1} - \hat{y}_{j+1})^2}{\sum_{j=1}^m (y_{j+1} - y_j)^2}}, \\ \text{ARV} &= \frac{\sum_{j=1}^m (y_{j+1} - \hat{y}_{j+1})^2}{\sum_{j=1}^m (y_{j+1} - \bar{y})^2}, \end{aligned}$$

em que  $m$  é o número de valores ajustados e/ou previstos pelo modelo.

Para desenvolver medidas de precisão de previsões capazes de avaliar modelos intervalares, nós consideramos o resíduo intervalar definido como

$$\mathbf{e}_t = \mathbf{I}_t - \hat{\mathbf{I}}_t.$$

Assim, para avaliação de modelos de previsão de STI, nós apresentamos as medidas que denominamos de *erro quadrático intervalar médio* ( $\text{MSE}^I$ ), a *estatística U de Theil intervalar* ( $U^I$ ) e a *variância relativa intervalar média* ( $\text{ARV}^I$ ) que mensuram o erro de ajuste de intervalos. As medidas são dadas, respectivamente, por

$$\begin{aligned} \text{MSE}^I &= \frac{1}{2m} \sum_{j=1}^m (\mathbf{I}_t - \hat{\mathbf{I}}_t)^\top (\mathbf{I}_t - \hat{\mathbf{I}}_t) \\ &= \frac{1}{2m} \sum_{j=1}^m \left[ (X_j^U - \hat{X}_j^U)^2 + (X_j^L - \hat{X}_j^L)^2 \right], \\ U^I &= \sqrt{\frac{\sum_{j=1}^m (\mathbf{I}_{j+1} - \hat{\mathbf{I}}_{j+1})^\top (\mathbf{I}_{j+1} - \hat{\mathbf{I}}_{j+1})}{\sum_{j=1}^m (\mathbf{I}_{j+1} - \mathbf{I}_j)^\top (\mathbf{I}_{j+1} - \mathbf{I}_j)}} \end{aligned}$$

$$\begin{aligned}
&= \sqrt{\frac{\sum_{j=1}^m (X_{j+1}^U - \widehat{X}_{j+1}^U)^2 + \sum_{j=1}^m (X_{j+1}^L - \widehat{X}_{j+1}^L)^2}{\sum_{j=1}^m (X_{j+1}^U - X_j^U)^2 + \sum_{j=1}^m (X_{j+1}^L - X_j^L)^2}}, \\
\text{ARV}^I &= \frac{\sum_{j=1}^m (\mathbf{I}_{j+1} - \widehat{\mathbf{I}}_{j+1})^\top (\mathbf{I}_{j+1} - \widehat{\mathbf{I}}_{j+1})}{\sum_{j=1}^m (\mathbf{I}_{j+1} - \bar{\mathbf{I}})^\top (\mathbf{I}_{j+1} - \bar{\mathbf{I}})} \\
&= \frac{\sum_{j=1}^m (X_{j+1}^U - \widehat{X}_{j+1}^U)^2 + \sum_{j=1}^m (X_{j+1}^L - \widehat{X}_{j+1}^L)^2}{\sum_{j=1}^m (X_{j+1}^U - \bar{X}^U)^2 + \sum_{j=1}^m (X_{j+1}^L - \bar{X}^L)^2},
\end{aligned}$$

em que  $m$  é o número de intervalos ajustados e/ou previstos pelo modelo,  $\widehat{\mathbf{I}}_t = (\widehat{X}_t^U, \widehat{X}_t^L)^\top$  é o  $t$ -ésimo intervalo ajustado,  $\bar{\mathbf{I}} = (\bar{X}^U, \bar{X}^L)^\top$  é o intervalo médio amostral e  $\bar{X}^U$  e  $\bar{X}^L$  são o limite superior e o limite inferior médios, respectivamente. O denominador  $2m$  no  $\text{MSE}^I$  deve-se ao fato de estarmos somando dois vetores de tamanho  $m$ .

Uma importante característica nas estatísticas apresentadas para avaliação de modelos na previsão de STI é que no cálculo do  $\text{MSE}^I$ , da  $U^I$  e da  $\text{ARV}^I$  são levados em consideração, simultaneamente, o erro no limite superior e o erro no limite inferior. Além disso, nas duas últimas medidas acima, os erros de previsão são avaliados independente da escala dos dados.

A estatística  $U$  de Theil intervalar,  $U^I$ , é usada para comparar os erros de predição fornecidos pelo modelo ajustado contra o modelo ingênuo passeio aleatório intervalar, em que o intervalo mais recente é considerado a previsão do intervalo futuro da STI, i.e.,  $\widehat{\mathbf{I}}_{t+1} = \mathbf{I}_t$ .  $U^I$  igual a 1 indica que o modelo preditor tem o mesmo desempenho de um passeio aleatório.  $U^I$  maior que 1 indica que o modelo ajustado é pior que o passeio aleatório, enquanto que  $U^I$  menor que 1 indica que o desempenho do modelo ajustado é superior ao modelo passeio aleatório.  $U^I$  tende a zero quando o modelo é perfeito (Mills, 2003) [72].

Por sua vez, a variância relativa intervalar média,  $\text{ARV}^I$ , compara o desempenho

do modelo ajustado com a predição dada pelo intervalo médio da STI. Quanto menor é o valor de  $ARV^I$ , melhor é o desempenho do modelo considerado. Além disso,  $ARV^I$  igual a 0 indica previsão perfeita, enquanto  $ARV^I$  igual a 1 indica que o modelo apenas fornece previsões iguais ao intervalo médio da STI.

As medidas apresentadas nesta seção serão usadas na avaliação da qualidade do ajuste dos modelos obtidos para previsão de STI.

# Capítulo 5

## Resultados experimentais

A melhor maneira de validar as técnicas e modelos para STI obtidos neste trabalho é através de aplicação prática destes. Assim, relatamos alguns resultados obtidos através das técnicas desenvolvidas na tese no contexto de análise de séries temporais do tipo intervalo.

As estimações e cálculos foram implementados no programa estatístico livre R (R Development Core Team, 2006) [78]. O R é um ambiente de programação para realização de análises estatísticas de dados e de análises gráficas. Trata-se de uma linguagem orientada a objetos que corresponde a uma versão ampliada e aprimorada da linguagem S. O R é um programa bastante flexível, gratuito e de código aberto e encontra-se disponível em <http://www.r-project.org>. Revisão detalhada sobre este ambiente de programação pode ser vista no livro de Venables e Ripley (2002) [90].

Além dos métodos propostos com o objetivo específico de modelar STI, nós ajustamos as STI através de modelos univalorados de séries temporais. Os modelos univalorados foram ajustados às séries dos limites superiores e dos limites inferiores separadamente e independentemente. Ou seja, para ajustar os modelos univalorados, as STI foram ajustadas através de dois modelos independentes: um modelo ajustado à série univalorada dos limites superiores dos intervalos,  $\{X_t^U\}$ , e outro modelo ajustado à série univalorada dos limites inferiores dos intervalos,  $\{X_t^L\}$ . O intervalo previsto é dado pela previsão do limite superior e pela previsão do limite inferior obtidas de modo independente. A finalidade é comparar e avaliar se os modelos propostos especificamente para STI apresentam ganho de precisão quando comparados aos modelos clássicos ajustados de modo independente. Vale salientar

que estes modelos ajustados de modo independente são casos especiais dos modelos de ajuste dependente, em que as matrizes  $(2 \times 2)$  de parâmetros têm elementos da diagonal secundária nulos.

Os métodos e modelos que são comparados neste capítulo estão descritos resumidamente na Tabela 5.1. Na primeira coluna da tabela estão os nomes dos modelos como serão apresentados nos resultados ao longo da seção e na segunda coluna estão as descrições propriamente ditas. Na coluna referente aos nomes dos modelos, o índice  $c,r$  designa que são ajustados dois modelos univariados (clássicos) independentes, um à série dos centros e outro à série das amplitudes dos intervalos. O modelo com a letra I sobrescrita consistem dos modelos em que o intervalo é previsto de modo dependente, i.e., a previsão do limite superior leva em consideração informações relativas ao limite inferior e vice-versa. O índice  $s,i$ , por sua vez, designa que são ajustados dois modelos também univariados independentes, um à série dos limites superiores e outro à série dos limites inferiores dos intervalos. Os três primeiros modelos da tabela foram detalhados no Capítulo 3 e os nove seguintes foram detalhados no Capítulo 4.

O procedimento adotado para encontrar os valores ótimos das constantes de alisamento dos métodos univariados,  $AES_{s,i}$ ,  $AED_{s,i}$  e  $HW_{s,i}$ , aplicados às séries dos limites superiores e dos limites inferiores separadamente, é a minimização da soma de quadrados dos erros de previsão um-passo-à-frente.

Para os métodos de alisamento exponencial estendidos para previsão de séries intervalares,  $AES^I$ ,  $AED^I$  e  $HW^I$ , as matrizes de constantes de alisamento foram estimadas através da minimização do critério da soma de quadrados dos erros de previsão um-passo-à-frente, em que as três funções objetivo são, respectivamente,  $\mathcal{S}_2(\alpha_{ij})$ ,  $\mathcal{R}_2(\alpha_{ij}, \beta_{ij})$  e  $\mathcal{T}_2(\alpha_{ij}, \beta_{ij}, \delta_{ij})$  apresentadas na Seção 4.4. O algoritmo utilizado para minimizar estas funções é o L-BFGS-B que admite restrições em caixa conforme descrito na Seção 4.4.3.

Os parâmetros do modelo  $ARMA_{c,r}$  ajustado à série dos centros e à série das amplitudes e os parâmetros dos modelos  $ARMA_{s,i}$  e  $ARIMA_{s,i}$  ajustados à série dos limites superiores e à série dos limites inferiores foram estimados via método de máxima verossimilhança. Os valores referentes à ordem  $p$  e  $q$  dos modelos  $ARMA_{c,r}$  e  $ARMA_{s,i}$  e referentes à ordem  $p$ ,  $d$  e  $q$  do modelo  $ARIMA_{s,i}$  foram obtidos pela minimização do critério de informação de Akaike (AIC) (Akaike, 1974) [1]. O uso

Tabela 5.1: Nome e descrição dos métodos e modelos utilizados na previsão de STI.

Modelo	Descrição
$ARMA_{c,r}$	Um modelo ARMA para a série dos centros, $X_t^c$ , e um modelo ARMA para a série das amplitudes, $X_t^r$ , ajustados de modo independente.
$MLP_{c,r}$	Uma rede neural MLP com uma saída para a série dos centros, $X_t^c$ , e uma rede MLP com uma saída para a série das amplitudes, $X_t^r$ , dos intervalos ajustadas de modo independente.
Híbrido $_{c,r}$	Modelo híbrido que ajusta o modelo $ARMA_{c,r}$ na primeira etapa e a rede $MLP_{c,r}$ na segunda etapa. Um modelo híbrido para a série dos centros, $X_t^c$ , e um modelo híbrido para a série das amplitudes, $X_t^r$ , ajustados de modo independente.
AES <sup>I</sup>	Método de AES para STI, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
AED <sup>I</sup>	Método de AED para STI, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
HW <sup>I</sup>	Método Holt–Winters para STI, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
iARMA-irrestrito	Modelo iARMA para STI com estimação irrestrita, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
iARMA-restrito	Modelo iARMA para STI com estimação restrita, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
iARMA-combinado	Modelo que combina os modelos iARMA-irrestrito e iARMA-restrito, em que os intervalos previstos invertidos pelo modelo irrestrito são substituídos pelos intervalos previstos pelo modelo restrito.
iARIMA	Modelo iARIMA para STI, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
MLP <sup>I</sup>	Rede neural MLP com duas saídas para a STI, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
Híbrido <sup>I</sup>	Modelo híbrido para STI que ajusta o modelo iARMA-restrito na primeira etapa e a rede neural MLP <sup>I</sup> na segunda etapa, com séries $X_t^U$ e $X_t^L$ ajustadas de modo dependente.
AES $_{s,i}$	Um AES para a série dos limites superiores, $X_t^U$ , e um AES para a série dos limites inferiores, $X_t^L$ , ajustados de modo independente.
AED $_{s,i}$	Um AED para a série dos limites superiores, $X_t^U$ , e um AED para a série dos limites inferiores, $X_t^L$ , ajustados de modo independente.
HW $_{s,i}$	Um HW para a série dos limites superiores, $X_t^U$ , e um HW para a série dos limites inferiores, $X_t^L$ , ajustados de modo independente.
ARMA $_{s,i}$	Um ARMA para a série dos limites superiores, $X_t^U$ , e um ARMA para a série dos limites inferiores, $X_t^L$ , ajustados de modo independente.
ARIMA $_{s,i}$	Um ARIMA para a série dos limites superiores, $X_t^U$ , e um ARIMA para a série dos limites inferiores, $X_t^L$ , ajustados de modo independente.
MLP $_{s,i}$	Uma rede MLP com uma saída para a série dos limites superiores, $X_t^U$ , e uma rede MLP com uma saída para a série dos limites inferiores, $X_t^L$ , ajustados de modo independente.
Híbrido $_{s,i}$	Modelo híbrido que ajusta o modelo ARMA $_{s,i}$ na primeira etapa e a rede MLP $_{s,i}$ na segunda etapa. Um modelo híbrido para a série dos limites inferiores, $X_t^L$ , ajustados de modo independente.

do método de máxima verossimilhança foi possível pelo fato destes modelos serem univalorados. Nos experimentos, nós variamos os valores para busca da ordem destes modelos da seguinte forma:  $p = 1, 2, \dots, 6$ ;  $q = 1, 2, \dots, 6$ ; e  $d = 1, 2, 3$ . Assim, a ordem que fornece menor valor para o AIC é selecionada para previsão da série temporal.

Para proceder com a estimação das matrizes de parâmetros dos modelos iARMA-irrestrito e iARMA-restrito, nós adaptamos o procedimento de mínimos quadrados em dois estágios de Hannan e Rissanen (1982) [49] com descrito na Seção 4.5.2. No primeiro estágio, ajustamos o modelo iAR( $l$ ) com  $l = \frac{1}{2}\sqrt{n_T}$ , em que  $n_T$  denota o tamanho da série no conjunto de treinamento. A estimação de parâmetros do modelo iARMA-restrito que consiste em considerar a relação de ordem entre as séries  $\{X_t^U\}$  e  $\{X_t^L\}$  foi detalhada na Seção 4.5.3. Nos modelos iARMA-irrestrito e iARMA-restrito, estimar os parâmetros consiste em minimizar a soma de quadrados dos erros de previsão em STI.

Como o custo computacional da estimação dos modelos iARMA-irrestrito e iARMA-restrito não é tão grande, a determinação da ordem  $p, q$  consiste em estimar modelos com  $p$  variando de 1 a 6 e  $q$  variando de 0 a 5 e calcular o erro quadrático médio para cada combinação de  $p$  e  $q$ . Então, determinar a ordem  $p, q$  para a qual o modelo apresentou melhor desempenho no conjunto de treinamento.

O ajuste do modelo iARIMA consiste de tomar diferenças da STI e, então, estimar os parâmetros como os modelos iARMA. A seleção da ordem  $p, d, q$  é realizada como nos modelos iARMA e com  $d = 1, 2, 3$ .

Redes neurais MLP com uma camada escondida e uma única saída, como mostrada na Figura 4.17 no Capítulo 4, foram utilizadas para modelar as séries univaloradas dos centros e das amplitudes na MLP <sub>$c,r$</sub>  e para modelar as séries univaloradas dos limites superiores e dos limites inferiores na MLP <sub>$s,i$</sub> . Este mesmo tipo de rede MLP foi utilizada na segunda etapa dos modelos híbridos Híbrido <sub>$c,r$</sub>  e Híbrido <sub>$s,i$</sub>  para modelar as séries univaloradas dos resíduos dos modelos lineares ajustados às séries univaloradas dos centros, das amplitudes, dos limites superiores e dos limites inferiores. A rede neural MLP<sup>I</sup> tem uma camada escondida e duas saídas, como mostrada na Figura 4.18 no Capítulo 4.

Para treinamento das redes MLP para as séries univaloradas e das redes MLP<sup>I</sup>,

nós usamos o método do gradiente conjugado. O interesse não é investigar as propriedades dos algoritmos de treinamento das redes neurais, no entanto, encontramos uma boa razão para uso do método do gradiente conjugado que é o baixo custo computacional, evitando as exigências associadas com o cálculo, armazenamento e inversão da matriz hessiana. Maior velocidade no treinamento das redes neurais é uma qualidade desejável, visto que elas são ajustadas diversas vezes durante os experimentos realizados. O método do gradiente conjugado é bastante adequado para o treinamento de redes *perceptrons* de múltiplas camadas, com aplicações típicas que incluem aproximação de funções, controle e análise de séries temporais (Haykin, 2001, Capítulo 4) [52].

A seleção do número de entradas e do número ideal de nodos na camada escondida nas redes  $MLP_{c,r}$ ,  $MLP_{s,i}$  e  $MLP^I$  envolve experimentação. Nos experimentos aqui apresentados, um grupo de redes neurais com diferentes números de entradas variando de 1 a 10 e diferentes números de nodos escondidos variando no conjunto  $\{2,5,8,12,15,20\}$  foram treinadas e cada uma delas foi avaliada no conjunto de treinamento. Nos experimentos, utilizamos 50 iterações do algoritmo do gradiente conjugado, em que os pesos iniciais foram gerados aleatoriamente para reduzir o problema de mínimos locais (Pretchelt, 1994) [77]. Para determinar a melhor configuração – número de entradas e número de nodos escondidos –, nós computamos o erro quadrático médio das 50 iterações e a rede com melhor desempenho no conjunto de treinamento foi selecionada para efetuar as previsões da série temporal. Este mesmo procedimento foi adotado nas redes neurais usadas na segunda etapa dos modelos híbridos  $Híbrido_{c,r}$ ,  $Híbrido_{s,i}$  e  $Híbrido^I$ .

Depois de o modelo ter sido ajustado a uma dada STI, nós avaliamos se este modelo fornece uma descrição adequada dos dados. A ideia é verificar o comportamento dos resíduos representados pela diferença entre o intervalo real e o intervalo previsto, i.e.,

$$\mathbf{I}_t - \widehat{\mathbf{I}}_t = \begin{bmatrix} X_t^U - \widehat{X}_t^U \\ X_t^L - \widehat{X}_t^L \end{bmatrix}.$$

Baseados nos resíduos acima, os resultados da avaliação da qualidade do ajuste dos métodos e modelos para STI são computados mediante duas medidas de precisão de previsões para STI: a *estatística U de Theil intervalar* ( $U^I$ ) e a *variância relativa*

*intervalar média* (ARV<sup>I</sup>). Estas medidas são representadas, respectivamente, por

$$U^I = \sqrt{\frac{\sum_{j=1}^m (X_{j+1}^U - \widehat{X}_{j+1}^U)^2 + \sum_{j=1}^m (X_{j+1}^L - \widehat{X}_{j+1}^L)^2}{\sum_{j=1}^m (X_{j+1}^U - X_j^U)^2 + \sum_{j=1}^m (X_{j+1}^L - X_j^L)^2}},$$

$$ARV^I = \frac{\sum_{j=1}^m (X_{j+1}^U - \widehat{X}_{j+1}^U)^2 + \sum_{j=1}^m (X_{j+1}^L - \widehat{X}_{j+1}^L)^2}{\sum_{j=1}^m (X_{j+1}^U - \bar{X}^U)^2 + \sum_{j=1}^m (X_{j+1}^L - \bar{X}^L)^2},$$

apresentadas em detalhes na Seção 4.7.

Para um modelo adequado, espera-se que os resíduos sejam serialmente (ao longo do tempo) não correlacionados. Assim, a autocorrelação residual é útil para verificar se o modelo ajustado é suficiente para filtrar a autocorrelação da série. Evidência de correlação serial nos resíduos é uma indicação de que uma ou mais características da série não foi adequadamente capturada pelo modelo.

Consequentemente, além de observar as medidas de erro de previsão, nós examinamos a correlação serial dos resíduos fornecidos por cada um dos modelos ajustados. Para isso, nós executamos o teste Ljung–Box da hipótese nula de não-correlação da série residual. Ele consiste em testar se um grupo de autocorrelações é significativamente diferente de zero, ou seja, consiste em testar a hipótese nula de que as autocorrelações residuais são todas nulas, contra a hipótese alternativa de que ao menos uma delas é diferente de zero, i.e.,

$$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0,$$

$$H_1: \text{pelo menos uma autocorrelação não é nula.}$$

O teste Ljung–Box é considerado nos experimentos e foi realizado para as séries dos resíduos no limite superior e no limite inferior separadamente. Para isso, computamos as autocorrelações amostrais residuais referentes às 30 primeiras defasagens,  $r_1, r_2, \dots, r_{30}$ , para cada um dos modelos ajustados. A rejeição de  $H_0$  pode sugerir que o modelo não foi capaz de filtrar toda a correlação serial da STI.

Além do exame das medidas de precisão e da correlação serial dos resíduos fornecidos pelos modelos, nós estamos interessados em avaliar os desempenhos destes

modelos em relação à ordem entre  $X_t^U$  e  $X_t^L$ . Ou seja, queremos observar se os modelos fornecem intervalos ajustados que não respeitam a coerência matemática  $\widehat{X}_t^U \geq \widehat{X}_t^L$ . Para isso, nós computamos o número de vezes em que um determinado modelo forneceu previsões com erro do tipo  $\widehat{X}_t^U < \widehat{X}_t^L$ . Nós denominamos estes pontos de previsões com erro por *previsões invertidas*. Obviamente, estas previsões invertidas não são observadas no modelo iARMA-restrito, pois as restrições impostas não permitem. Desta forma, os resultados foram computados considerando o ajuste de todos os intervalos (incluindo as previsões invertidas) e considerando os ajustes após eliminação dos intervalos em que pelo menos um dos modelos produziu previsões invertidas.

Devido à possibilidade de produção de previsões invertidas, nós avaliamos ainda outro método que combina as previsões dos modelos iARMA-irrestrito e iARMA-restrito para evitar previsões invertidas. O mecanismo de combinação é bastante simples. Nela, os intervalos previstos invertidos pelo modelo irrestrito são substituídos por intervalos previstos pelo modelo restrito. Por isso, o modelo combinado também é livre de previsões invertidas. Esta combinação dos modelos iARMA-irrestrito e iARMA-restrito foi chamada de iARMA-combinado e pode ser útil no caso em que o modelo irrestrito apresente desempenho superior ao modelo restrito, mas produza previsões invertidas.

Inicialmente, avaliamos os métodos estendidos para STI em séries sintéticas simuladas com diferentes características. Em seguida, apresentamos resultados práticos em STI reais.

## 5.1 Experimentos com STI sintéticas

O objetivo principal dos experimentos com STI sintéticas realizados é mostrar a efetividade dos métodos propostos para análise de STI, procurando responder algumas indagações surgidas no desenvolvimento do trabalho. Examinamos a efetividade de métodos propostos e efetuamos comparações entre os modelos de previsão de STI propostos aqui com o objetivo de avaliar os seus desempenhos em situações diversas e, se possível, identificar qual o modelo de previsão é recomendado em cada situação.

Antes de apresentar os resultados, nós apresentamos cada processo gerador dos

dados (PGD) utilizado para simular as STI sintéticas com diferentes características. A seguir encontram-se descritos os experimentos realizados com estas séries.

Iniciamos com um experimento simples que consiste de uma breve avaliação da função de autocorrelação amostral para STI.

### 5.1.1 Avaliação da ACF<sup>I</sup> para STI sintéticas

O primeiro experimento realizado com dados sintéticos consiste em avaliar se as autocorrelações amostrais apresentadas para STI são úteis na identificação de padrões em STI de modo semelhante ao que ocorre em séries temporais univariadas. Neste experimento, são geradas STI com estruturas diversas e são fornecidas suas representações gráficas. Para avaliar a correlação serial em STI, são computadas as autocorrelações amostrais e o correlograma para STI é construído. Vale lembrar que as autocorrelações para STI podem ser interpretadas da mesma forma que a correlação linear em dados univariados, i.e.,  $-1 \leq r_k^I \leq 1$  e quanto mais próximo dos extremos, maior é a correlação linear entre duas variáveis. Com relação à interpretação do correlograma em séries univariadas, uma breve revisão pode ser vista em Ehlers (2005) [37].

Neste experimento, para cada processo gerador da STI considerado, nós simulamos 60 intervalos ( $n = 60$ ). Para avaliar os resultados, representações gráficas das séries simuladas e os correspondentes correlogramas para STI são apresentados.

O primeiro PGD considerado consiste de uma STI do tipo processo intervalar puramente aleatório representada por

$$\mathbf{I}_t = \mathbf{u}_t = \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix}, \quad (5.1)$$

em que  $u_t^U$  e  $u_t^L$ , com  $u_t^U \geq u_t^L$ , são processos puramente aleatórios com médias, respectivamente,  $\mu^U$  e  $\mu^L$  e variâncias constantes  $\sigma^{U^2}$  e  $\sigma^{L^2}$ . A Figura 5.1 apresenta vários gráficos referentes a este PGD temporal. Nesta figura, os gráficos à esquerda são as representações gráficas das STI simuladas com 60 intervalos e os gráficos à direita representam os correlogramas para STI correspondentes nos quais computamos as 30 primeiras autocorrelações para STI. Em séries temporais univariadas, espera-se que as autocorrelações do processo puramente aleatório sejam nulas para

$k \neq 0$ . Podemos notar na Figura 5.1 que, para as séries simuladas, nenhuma das 30 primeiras autocorrelações para STI apresenta valor alto. Com isso, há indícios de que as realizações intervalares são, de fato, independentes.

O segundo PGD analisado por intermédio das autocorrelações amostrais para STI é o que se chama série com correlação de curto-prazo. Neste processo, o valor atual da série depende substancialmente do valor mais recente dela mesma. Em séries temporais univariadas, o correlograma deste processo deve exibir um valor relativamente grande na autocorrelação para  $k = 1$ , i.e.,  $r_1$ , seguido de valores que tendem a decair tornando-se aproximadamente 0 (zero). A Figura 5.2 apresenta os gráficos, em que simulamos STI através do processo

$$\mathbf{I}_t = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{13} & \phi_{14} \end{bmatrix} \begin{bmatrix} X_{t-1}^U \\ X_{t-1}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix}, \quad (5.2)$$

para  $0 < \phi_{1j} < 1$ ,  $j = 1, 2, 3, 4$ . Os  $u_t^U$  e  $u_t^L$  são independentes e identicamente distribuídos (i.i.d.) com médias zero e variâncias constantes  $\sigma_u^2$ . Nas STI simuladas para este PGD, Figura 5.2, nota-se nos correlogramas para STI que a segunda autocorrelação amostral, equivalente à primeira defasagem,  $k = 1$ , é a que apresenta maior valor, desconsiderando a  $r_0^I = 1$ .

Simulamos também STI em que o intervalo atual depende dos dois intervalos anteriores, como

$$\mathbf{I}_t = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{13} & \phi_{14} \end{bmatrix} \begin{bmatrix} X_{t-1}^U \\ X_{t-1}^L \end{bmatrix} + \begin{bmatrix} \phi_{21} & \phi_{22} \\ \phi_{23} & \phi_{24} \end{bmatrix} \begin{bmatrix} X_{t-2}^U \\ X_{t-2}^L \end{bmatrix} + \begin{bmatrix} u_t^U \\ u_t^L \end{bmatrix},$$

para  $0 < \phi_{1j}, \phi_{2j} < 1$ ,  $j = 1, 2, 3, 4$ . Os  $u_t^U$  e  $u_t^L$  são i.i.d. com médias zero e variâncias constantes  $\sigma_u^2$ . Os correlogramas resultantes encontram-se na Figura 5.3. Nesta figura podemos ver que as maiores autocorrelações amostrais para STI são observadas para as defasagens  $k = 1$  e  $k = 2$ .

O próximo processo avaliado através da função de autocorrelação amostral para STI é a série temporal com correlação negativa. Esta característica também é de fácil percepção no correlograma. Em dados clássicos, uma série com correlação negativa apresenta valores alternados acima e abaixo de um valor médio e as autocorrelações também tendem a se alternar. A STI com correlação negativa foi gerada

a partir de

$$\mathbf{I}_t = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{13} & \phi_{14} \end{bmatrix} \begin{bmatrix} X_{t-1}^U \\ X_{t-1}^L \end{bmatrix} + \begin{bmatrix} \epsilon_t^U \\ \epsilon_t^L \end{bmatrix},$$

para coeficientes negativos,  $-1 < \phi_{1j} < 0$ ,  $j = 1, 2, 3, 4$ . Os correlogramas referentes a estes processos intervalares simulados encontram-se na Figura 5.4. Vemos que a função de autocorrelação amostral proposta é capaz de detectar o comportamento oscilatório em STI com correlação negativa. Vale notar nos correlogramas para as STI com correlação negativa que a autocorrelação  $r_1^I$  é a mais alta, desconsiderando a  $r_0^I = 1$ .

Uma importante característica frequentemente presente em séries temporais é a tendência. Em séries temporais univariadas, na presença de tendência os valores da autocorrelação decaem lentamente para zero. Na prática, qualquer tendência deve ser removida antes do cálculo da função de autocorrelação amostral, pois este comportamento pode ocultar outras características que podem ser verificadas no correlograma. Agora, avaliamos o comportamento da função de autocorrelação amostral em STI com diferentes formas supostas de tendência. A Figura 5.5 apresenta as STI simuladas e os correlogramas correspondentes. Nas STI simuladas com formas distintas de tendência, o correlograma apresenta comportamento análogo ao comportamento em séries univariadas com tendência, ou seja, as autocorrelações decaem lentamente.

Outro aspecto importante a ser identificado em séries temporais é a sazonalidade. Se a função de autocorrelação amostral exibe sazonalidade, há indícios de que a série temporal tem comportamento sazonal. Com isso, em geral um padrão sazonal é facilmente identificado através do correlograma. Na prática, se o padrão sazonal da série temporal já for evidente no gráfico da série original o correlograma trará pouca ou nenhuma informação adicional (Ehlers, 2005) [37]. As STI simuladas com supostas formas de sazonalidade e os correlogramas correspondentes estão dispostos na Figura 5.6. Como acontece em séries univariadas, em STI com sazonalidade o correlograma indica o comportamento sazonal da série.

De modo geral, a função de autocorrelação amostral proposta para STI,  $ACF^I$ , é capaz de identificar características presentes nos dados de forma semelhante à função de autocorrelação amostral em séries univariadas. Sendo assim, o estudo detalhado do comportamento da  $ACF^I$  em STI pode ser útil na sugestão do verdadeiro processo

Figura 5.1: Correlogramas em processos intervalares puramente aleatórios.

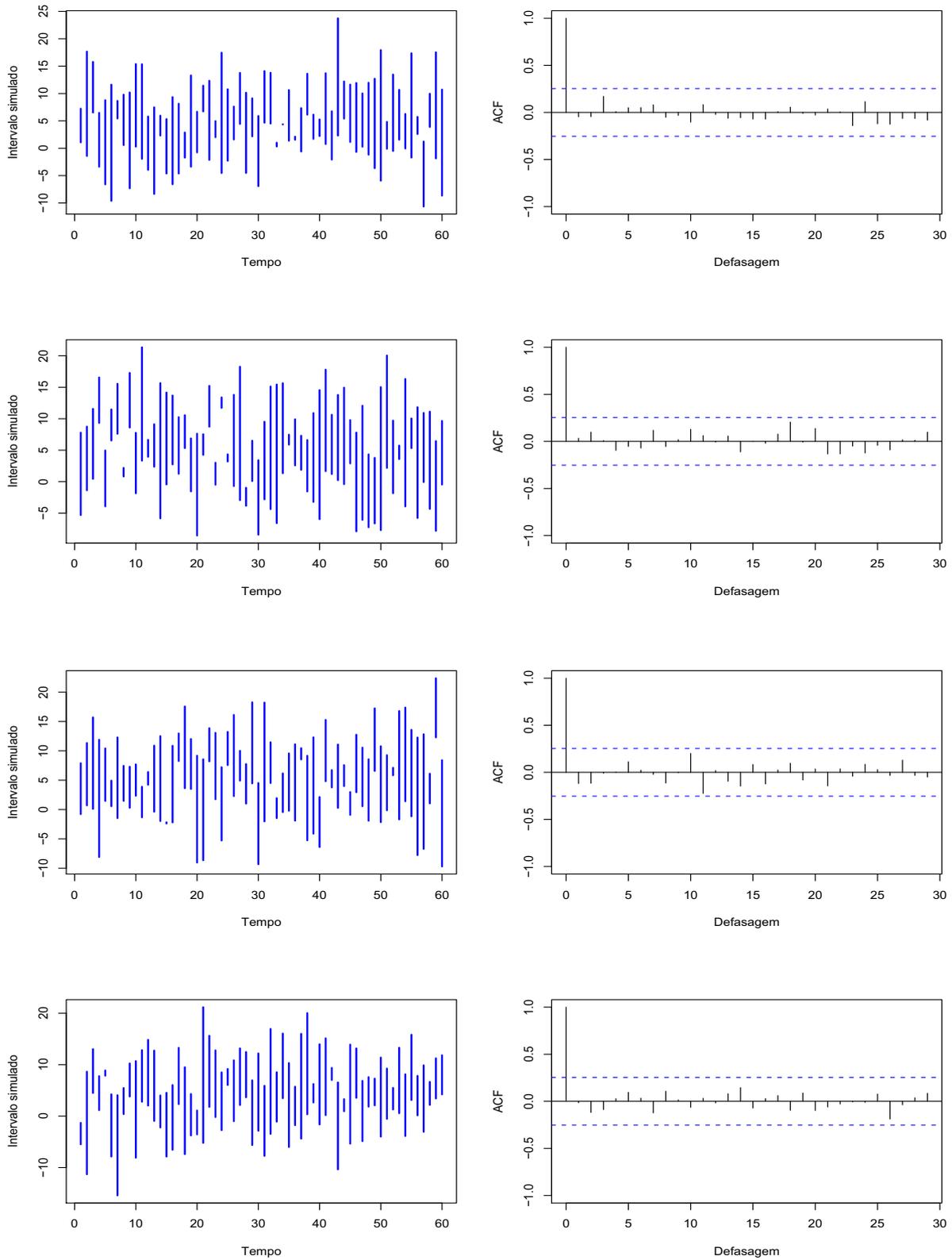


Figura 5.2: Correlogramas em processos intervalares com correlação de curto-prazo.

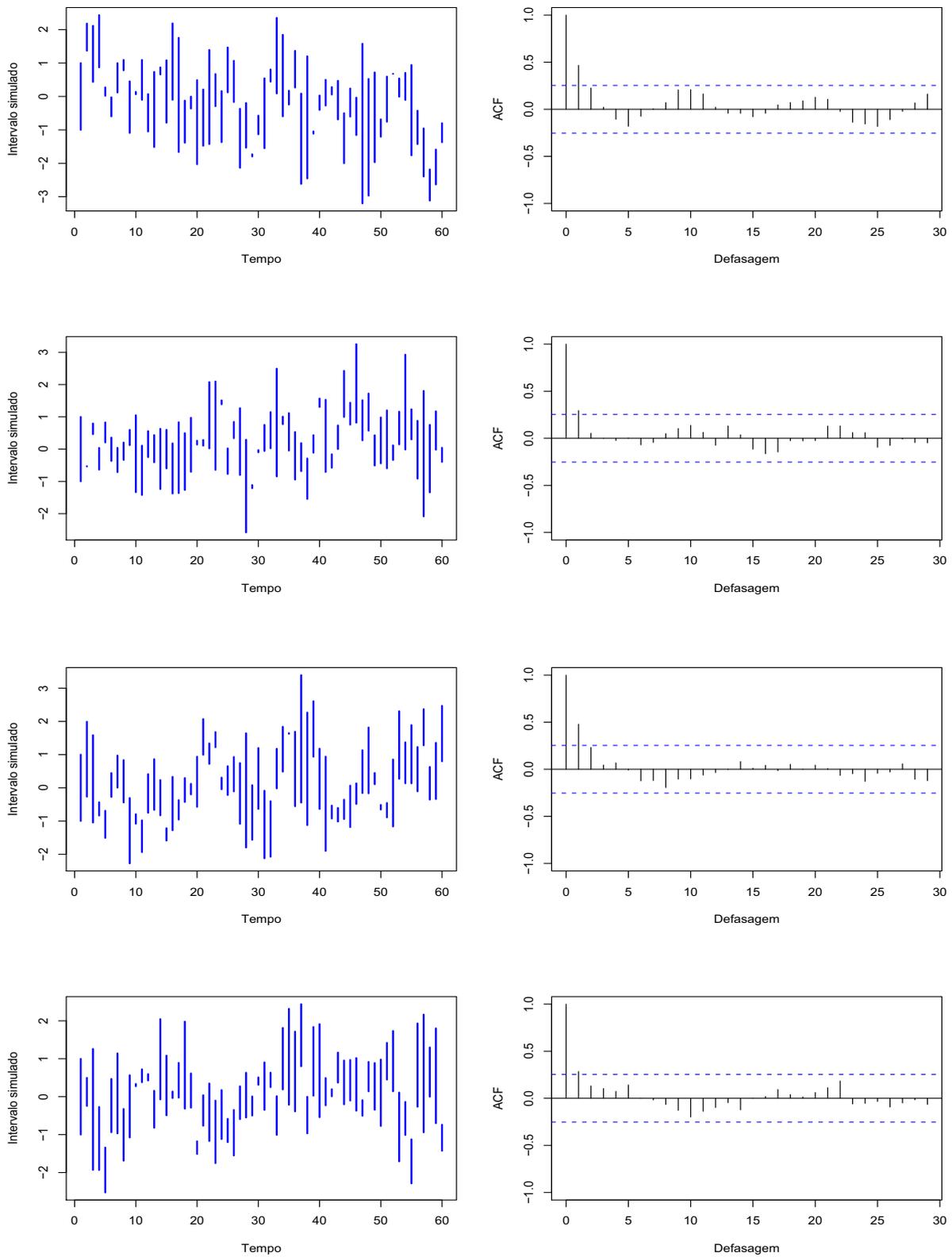


Figura 5.3: Correlogramas em processos intervalares iAR(2).

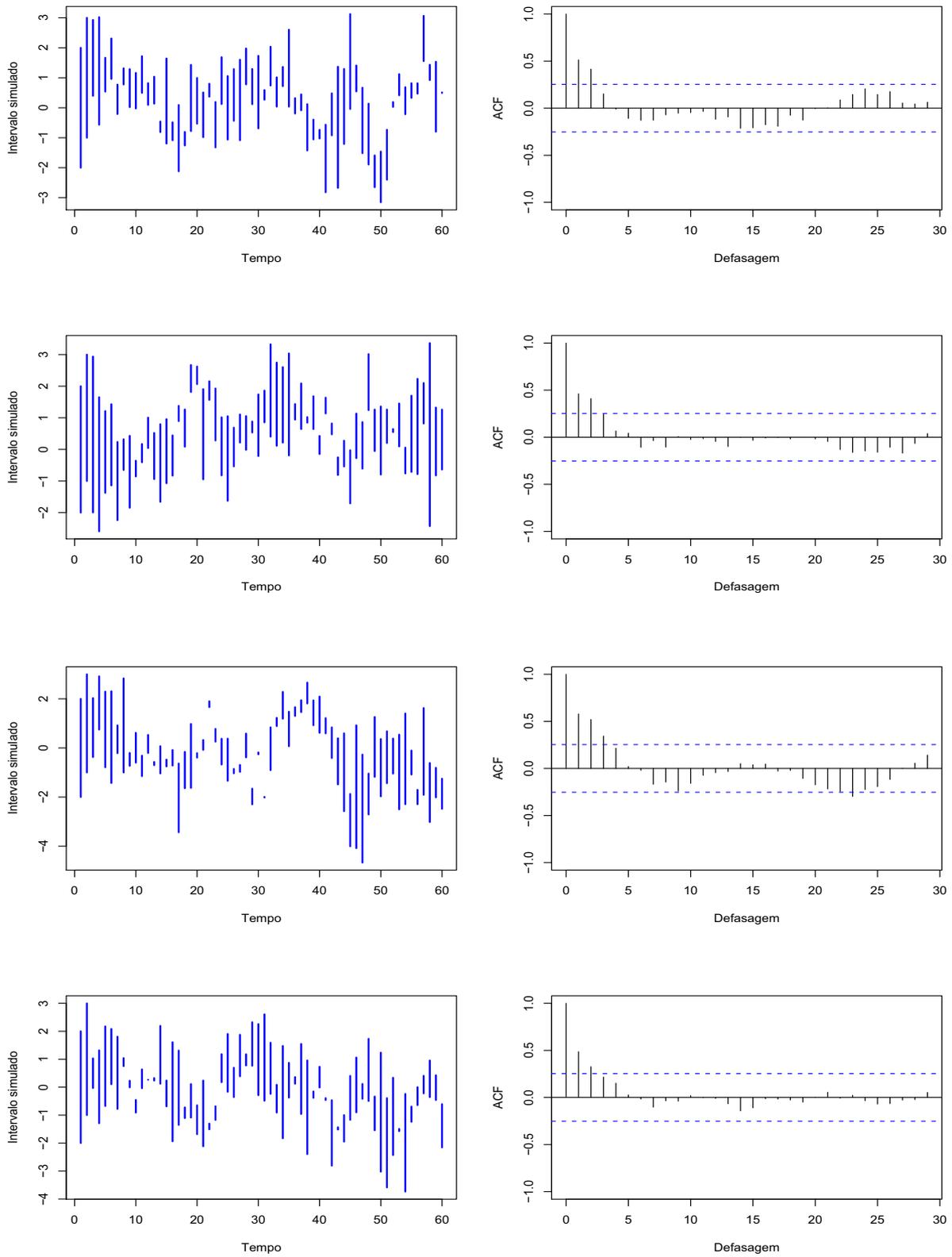


Figura 5.4: Correlogramas em processos intervalares com correlação negativa.

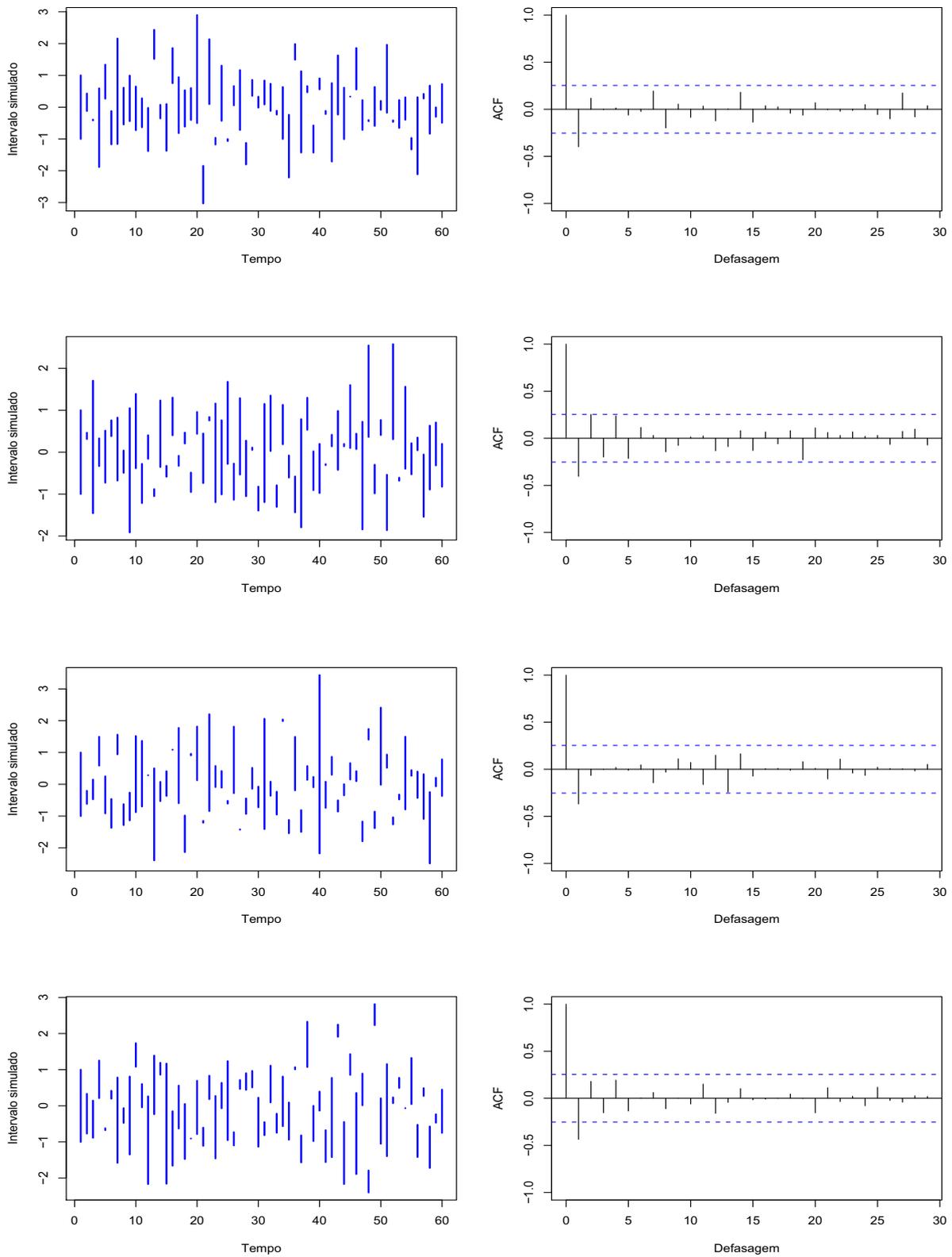


Figura 5.5: Correlogramas em processos intervalares com tendência.

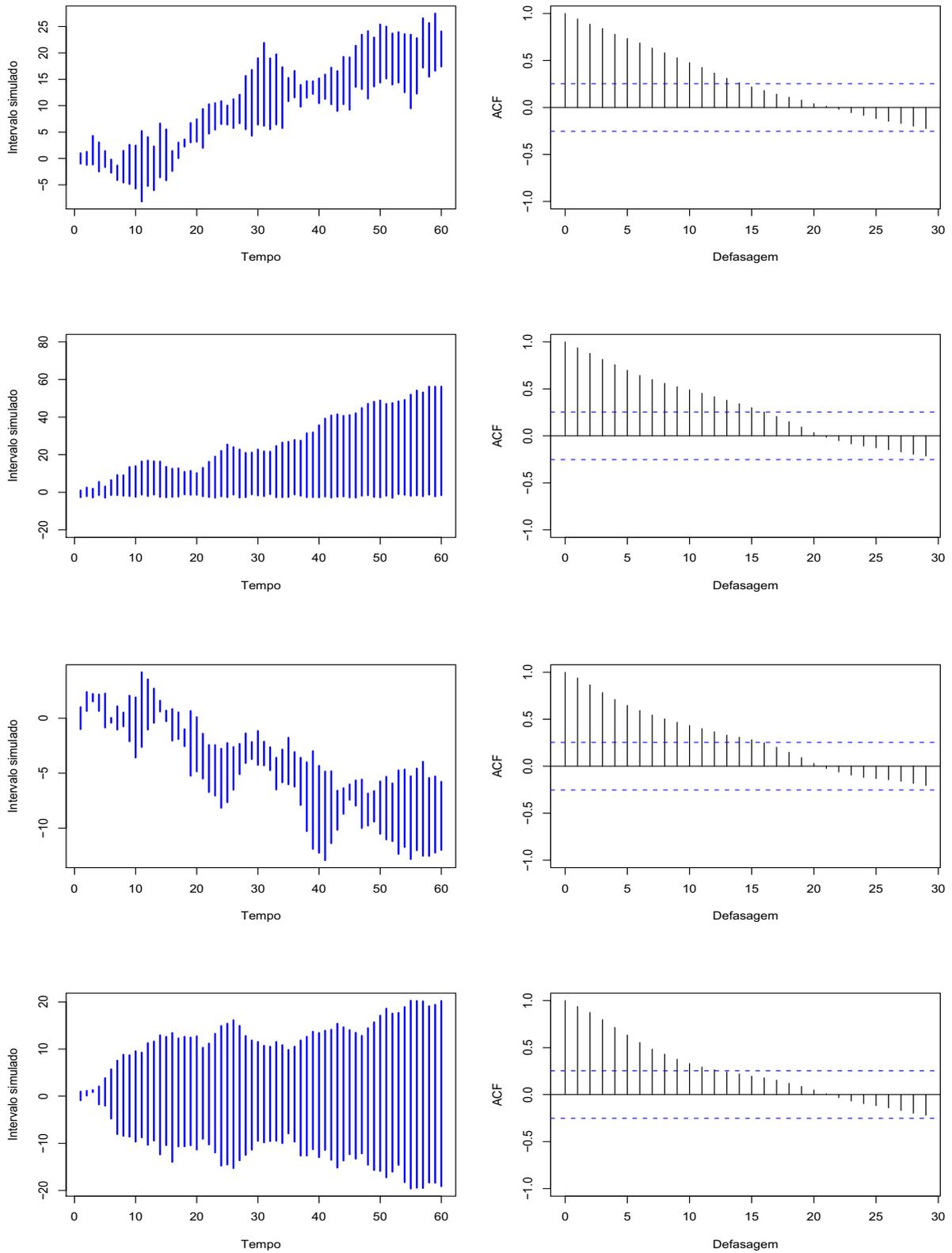
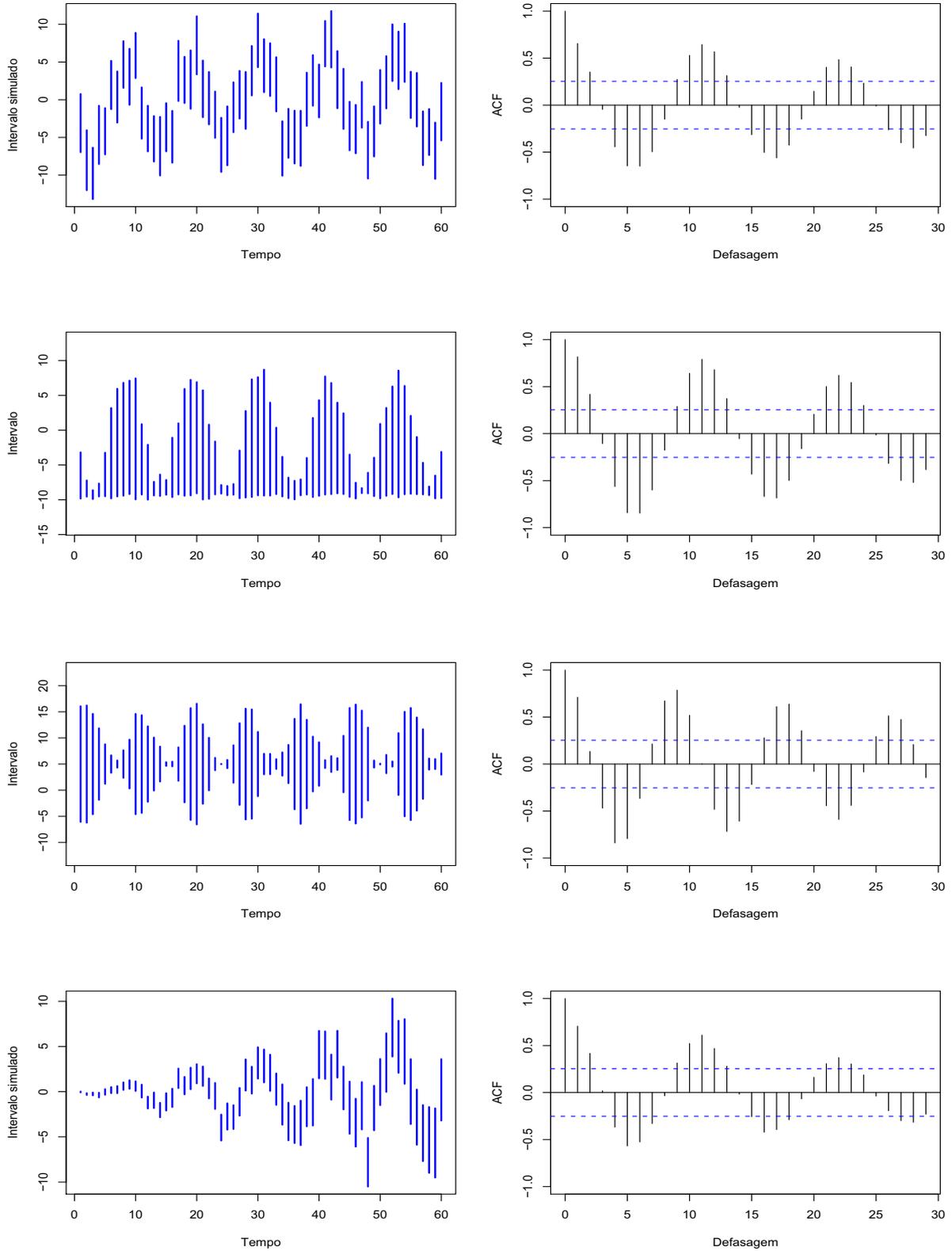


Figura 5.6: Correlogramas em processos intervalares com sazonalidade.



gerador da STI.

Durante o desenvolvimento e exame da  $ACF^I$  em STI surgiu o interesse em comparar os seus resultados com os valores das autocorrelações obtidas para a série dos limites superiores, para a série dos limites inferiores e para a série dos centros dos intervalos. Para isso, realizamos um único experimento simples, em que calculamos as autocorrelações para a série dos limites superiores,  $X_t^U$ , para a série dos limites inferiores,  $X_t^L$ , para a série dos centros,  $X_t^c = (X_t^L + X_t^U)/2$ , e para a STI  $\mathbf{I}_t$ . As autocorrelações para as séries univariadas  $X_t^L$ ,  $X_t^U$  e  $X_t^c$  foram calculadas através da forma usual para dados clássicos, como na Equação (4.2). Para a série intervalar  $\mathbf{I}_t$ , as autocorrelações amostrais foram calculadas através da  $ACF^I$  dada pela Equação (4.5), em que as séries dos limites superior e inferior são consideradas simultaneamente. O objetivo é mostrar a diferença existente entre os valores das autocorrelações fornecidos pela  $ACF^I$  com os fornecidos para as séries univariadas via forma clássica.

Para comparar os valores das autocorrelações, consideramos a STI dada pelo processo intervalar puramente aleatório dado em (5.1). A Tabela 5.2 apresenta os valores da função de autocorrelação amostral calculados para as primeiras 11 defasagens das séries univariadas  $X_t^L$ ,  $X_t^U$ ,  $X_t^c$  e da série intervalar  $\mathbf{I}_t$ . Notamos as diferenças nos valores fornecidos na Tabela 5.2, e, então, observamos que as autocorrelações fornecidas pela  $ACF^I$  não são iguais às obtidas para as séries univariadas  $X_t^L$ ,  $X_t^U$  e  $X_t^c$ .

Tabela 5.2: Função de autocorrelação amostral para as séries  $X_t^U$ ,  $X_t^L$ ,  $X_t^c$  e  $\mathbf{I}_t$ .

Série	Defasagem										
	0	1	2	3	4	5	6	7	8	9	10
$X_t^U$	1,000	-0,148	-0,215	0,099	-0,056	0,104	0,049	0,006	-0,048	-0,065	0,025
$X_t^L$	1,000	-0,014	0,053	-0,104	0,065	-0,006	-0,081	-0,113	0,051	-0,012	-0,028
$X_t^c$	1,000	-0,111	-0,096	-0,039	-0,086	0,041	-0,010	-0,023	-0,053	-0,044	0,098
$\mathbf{I}_t$	1,000	-0,085	-0,088	0,002	0,001	0,051	-0,012	-0,050	-0,001	-0,040	0,000

Na prática, a análise da ACF é frequentemente utilizada para avaliar a correlação serial dos resíduos do modelo de previsão. Com isso, podemos avaliar a estrutura dos resíduos e, então, identificar alguma característica que não foi captada pelo modelo adotado. Por exemplo, se os resíduos apresentam correlação serial, em que o correlograma aponta para presença de tendência, pode ser sinal de que a tendência deve ser removida antes da modelagem da série temporal. O mesmo deve

ser realizado com a sazonalidade se o correlograma dos resíduos indica a presença de tal característica.

### 5.1.2 Comparação dos modelos no ajuste de STI sintéticas

Seguindo com os experimentos de avaliação de técnicas estendidas para análise de séries temporais do tipo intervalo, a seguir, focamos nos métodos e modelos propostos para previsão de STI. O objetivo da comparação é verificar a qualidade destes métodos e modelos no ajuste de STI.

Nós efetuamos a comparação de todos os métodos e modelos descritos anteriormente na Tabela 5.1 seguindo os procedimentos de estimação e identificação explicados antes.

Antes de mostrar os resultados com relação aos desempenhos dos modelos, faz-se necessária a apresentação dos processos geradores dos dados das STI sintéticas consideradas. A seguir apresentamos cada um deles.

**Primeiro processo gerador dos dados (PGD-I):** No primeiro PGD das STI, são considerados dois processos temporais independentes, um representando a série univariada dos centros dos intervalos,  $X_t^c$ , e outro representando a série univariadas das amplitudes dos intervalos,  $X_t^r$ . De acordo com o procedimento adotado neste primeiro PGD, a série do limite superior não leva em consideração informações da série do limite inferior e vice-versa. Assim, no PGD-I a STI é gerada da forma

$$\mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} = \begin{bmatrix} X_t^c + \frac{X_t^r}{2} \\ X_t^c - \frac{X_t^r}{2} \end{bmatrix},$$

em que as séries dos centros e das amplitudes dos intervalos foram geradas, respectivamente, através dos processos independentes

$$\begin{aligned} X_t^c &= 15,0 + 0,9X_{t-1}^c - 0,5X_{t-2}^c - 0,3u_{t-1}^c + 0,2u_{t-2}^c + u_t^c, \\ X_t^r &= 2,0 + 0,7X_{t-1}^r - 0,2X_{t-2}^r + u_t^r, \end{aligned}$$

em que  $u_t^c$  e  $u_t^r$  têm médias zero e desvios-padrão iguais a um. A série das amplitudes é gerada de tal forma que  $X_t^r \geq 0 \forall t$  e, desta forma, asseguramos a condição  $X_t^U \geq X_t^L \forall t$ . Este PGD-I é um processo simples que favorece os modelos independentes

ajustados às séries dos centros e das amplitudes dos intervalos – os que levam o índice  $c, r$  na Tabela 5.1.

**Segundo processo gerador dos dados (PGD-II):** O segundo PGD também consiste de dois processos temporais gerados de modo independente. Agora, os dois processos representam a série dos limites superiores,  $X_t^U$ , e a série dos limites inferiores,  $X_t^L$ , dos intervalos. Da mesma forma que o PGD-I, no PGD-II a série do limite superior não leva em consideração informações da série do limite inferior e vice-versa. As STI no PGD-II são geradas por processos independentes na forma

$$\mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} = \begin{bmatrix} 21,0 + 0,4X_{t-1}^U + 0,3u_{t-1}^U + u_t^U \\ 18,0 + 0,4X_{t-1}^L + 0,3u_{t-1}^L + u_t^L \end{bmatrix},$$

em que  $u_t^U$  e  $u_t^L$  têm médias zero e desvios-padrão iguais a um. A equação que gera as séries dos limites superiores é similar à equação que gera as séries dos limites inferiores diferenciadas pelos interceptos que determinam as médias dos processos univariados e pelas séries defasadas consideradas em cada equação. Nas STI simuladas via o PGD-II, os modelos para ajuste das séries dos limites superior e inferior de modo independente – os que levam o índice  $s, i$  na Tabela 5.1 – devem levar vantagem.

**Terceiro processo gerador dos dados (PGD-III):** O terceiro PGD difere dos dois anteriores pelo fato de existir a influência de informação do limite inferior na geração da série dos limites superiores e vice-versa. Por isso, dizemos que a STI é gerada através de dois processos temporais dependentes. O PGD-III é dado por

$$\mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} = \begin{bmatrix} 0,63X_{t-1}^U - 0,35u_{t-1}^U + 0,38X_{t-1}^L + 0,21u_{t-1}^L + u_t^U \\ 0,37X_{t-1}^L + 0,20u_{t-1}^L + 0,62X_{t-1}^U - 0,37u_{t-1}^U + u_t^L \end{bmatrix}.$$

em que  $u_t^U$  e  $u_t^L$  têm médias zero e desvios-padrão iguais a um. Observe que na equação referente à série dos limites superiores, existem termos referentes ao limite inferior e na equação referente à série dos limites inferiores, existem termos referentes ao limite superior. Espera-se que o PGD-III favoreça os modelos em que as séries  $X_t^U$  e  $X_t^L$  são ajustadas de modo dependente.

A Tabela 5.3 apresenta a descrição resumida de cada um dos PGDs utilizados na geração das STI nos experimentos Monte Carlo.

Tabela 5.3: Nome e descrição dos PGDs utilizados na geração de STI.

PGD	Descrição
PGD-I	Dois processos temporais independentes, um representando a série univariada dos centros dos intervalos, $X_t^c$ , e outro representando a série univariadas das amplitudes dos intervalos, $X_t^r$ . As realizações da STI são dadas por $\mathbf{I}_t = [X_t^c - \frac{X_t^r}{2}; X_t^c + \frac{X_t^r}{2}]$ .
PGD-II	Dois processos temporais independentes, um representando a série dos limites superiores dos intervalos, $X_t^U$ , e outro representando série dos limites inferiores dos intervalos, $X_t^L$ .
PGD-III	Dois processos temporais dependentes, em que existe a influência de informação do limite inferior na geração da série dos limites superiores e vice-versa.

É desejável que os três PGDs descritos produzam sempre valor de  $X_t^U$  maior ou igual a  $X_t^L$ . No entanto, pode ocorrer ao acaso valor gerado de  $X_t^U$  menor que  $X_t^L$ . Caso isso seja observado durante a simulação, nós substituímos automaticamente a STI gerada por outra em que este fato não ocorra.

A comparação dos métodos e modelos foi efetuada via simulação Monte Carlo. Nós simulamos  $R = 500$  séries do tipo intervalo de acordo com processos geradores PGD-I, PGD-II e PGD-III. Para cada um dos PGDs, nós simulamos STI de tamanhos  $n = 220$  e  $n = 620$ . Para avaliar os modelos para previsão de intervalos futuros, nós dividimos as STI sintéticas em conjunto de treinamento com  $n - 20$  intervalos e conjunto de teste com 20 intervalos. Então, para avaliar os modelos em conjuntos de teste, nós consideramos dois subconjuntos: um referente aos primeiros 5 intervalos e outro referente ao total de 20 intervalos. Ou seja, consideramos dois conjuntos de teste, um de 5 passos-à-frente e um de 20 passos-à-frente. Assim, quando nos referimos ao horizonte de previsão de  $h$  passos-à-frente, nós realizamos uma sequência de  $h$  previsões: de 1 passo-à-frente até  $h$  passos-à-frente. Ou seja, quando consideramos o horizonte de previsão de 5 passos-à-frente, nós prevemos os próximos 5 intervalos futuros. Da mesma maneira, quando consideramos o horizonte de previsão de 20 passos-à-frente, nós prevemos os próximos 20 intervalos futuros. A ideia é observar os desempenhos dos modelos propostos para previsão de STI em séries com diferentes estruturas de correlação serial e diferentes tamanhos de amostra e horizontes de previsão e, ainda, verificar se estes modelos superam os modelos univariados no ajuste de STI.

Em cada réplica Monte Carlo, nós computamos as medidas de precisão de previsões  $U^I$  e  $ARV^I$  para cada modelo e, em seguida, calculamos o valor médio e o desvio-padrão das  $R$  réplicas. Além das medidas de precisão, nós computamos o

número médio de predições invertidas fornecidas por cada modelo. Com isso, apresentamos os resultados obtidos com os intervalos onde houve predições invertidas e os resultados obtidos eliminando-se os intervalos onde pelo menos um modelo forneceu predição invertida. A intenção é verificar se as predições invertidas interferem no desempenho dos modelos.

Para avaliar a correlação serial dos resíduos, nós realizamos o teste Ljung–Box em cada réplica Monte Carlo baseado em 30 autocorrelações residuais obtidas para cada modelo. O teste foi realizado separadamente para os resíduos no limite superior e para os resíduos no limite inferior no conjunto de treinamento com a finalidade de avaliar a qualidade do ajuste. Com isso, computamos a quantidade de vezes em que a hipótese nula de não-correlação da série residual é rejeitada. Os resultados são apresentados como a porcentagem de vezes em que  $H_0$  foi rejeitada nas  $R$  séries intervalares ( $\tilde{H}_0^I$ ) para os resíduos do limite superior ou para os resíduos para o limite inferior obtidas através da expressão

$$\tilde{H}_0^I = \frac{\tilde{H}_0^U + \tilde{H}_0^L}{2R} \times 100,$$

em que  $\tilde{H}_0^U$  é o número de vezes que  $H_0$  foi rejeitada no limite superior,  $\tilde{H}_0^L$  é o número de vezes que  $H_0$  foi rejeitada no limite inferior e  $R$  é o número de réplicas Monte Carlo. O número 2 no denominador deve-se ao fato de estarmos somando as rejeições no limite superior e as rejeições no limite inferior.

As Tabelas A.1 a A.18 referem-se aos resultados obtidos na comparação dos modelos no ajuste de STI sintéticas e estão apresentadas no Apêndice A na página 164. Estas tabelas incluem os resultados para os três PGDs (PGD-I, PGD-II e PGD-III), em dois tamanhos de amostra ( $n = 220$  e  $n = 620$ ), no conjunto de treinamento ( $n_T = n - 20$ ) e dois conjuntos de teste correspondentes a dois horizontes de previsão distintos (5 e 20 passos-à-frente).

Na primeira coluna das Tabelas A.1, A.4, A.7, A.10, A.13 e A.16 referentes aos conjuntos de treinamento estão apresentados os resultados para o teste Ljung–Box das autocorrelações residuais. Apenas estas tabelas apresentam os resultados para o teste Ljung–Box, pois ele foi realizado apenas para o conjunto de treinamento. Nesta coluna está a porcentagem referente à quantidade de vezes em que a hipótese nula  $H_0$  de não-correlação da série residual foi rejeitada ou nos resíduos do limites

superior ou nos resíduos do limite inferior nas  $R$  réplicas Monte Carlo. Denotamos esta porcentagem por  $\tilde{H}_0^I$ .

A segunda coluna das tabelas apresenta o número médio de vezes em que o modelo forneceu predições invertidas (Pred. inv.) nas  $R$  réplicas Monte Carlo. Vale lembrar que os modelos iARMA-restrito e iARMA-combinado estão livres de tais predições.

Nas demais colunas estão os valores médios e os respectivos desvios-padrão (em parênteses) das medidas de precisão  $U^I$  e  $ARV^I$  computados nas  $R$  réplicas Monte Carlo primeiro para todos os intervalos ajustados e previstos e, em seguida, apenas para os intervalos onde nenhum dos modelos forneceu predição invertida.

Cada uma das tabelas com os resultados está dividida em três blocos, em que cada bloco refere-se a um grupo de modelos conforme descritos anteriormente na Tabela 5.1. O primeiro bloco contém os resultados dos modelos  $ARMA_{c,r}$ ,  $MLP_{c,r}$  e  $Híbrido_{c,r}$ , e refere-se aos modelos ajustados à série dos centros e à série das amplitudes dos intervalos de modo independente. O segundo contém os resultados dos modelos  $AES^I$ ,  $AED^I$ ,  $HW^I$ , iARMA-irrestrito, iARMA-restrito, iARMA-combinado, iARIMA,  $MLP^I$  e  $Híbrido^I$ , e refere-se aos modelos em que os limites dos intervalos são ajustados simultaneamente de modo dependente. O terceiro bloco refere-se aos resultados dos modelos  $AES_{s,i}$ ,  $AED_{s,i}$ ,  $HW_{s,i}$ ,  $ARMA_{s,i}$ ,  $ARIMA_{s,i}$ ,  $MLP_{s,i}$  e  $Híbrido_{s,i}$ , que são os modelos univalorados ajustados à série dos limites superiores e à série dos limites inferiores de modo independente. No total são confrontados os resultados fornecidos por 19 modelos levando-se em conta a presença ou não de predições invertidas.

Para facilitar a comparação dos resultados e para avaliar a hipótese estatística de que os desempenhos de diferentes modelos nas STI sintéticas são iguais, nós construímos intervalos de 95% de confiança para os valores médios de  $U^I$  e de  $ARV^I$  nos experimentos Monte Carlo. Através do intervalo de confiança (IC) construído para a média de  $U^I$  e para a média de  $ARV^I$ , podemos ter indícios de que a média fornecida por um modelo “A” é igual a um valor especificado  $\mu_0$  ou que a média fornecida por um modelo “A” é igual à média fornecida por um modelo “B”.

O mecanismo para realização de teste de hipóteses através de IC é simples. Dado um IC obtido, deve-se rejeitar a hipótese nula de que a média seja igual a qualquer valor fora desse IC e deve-se não rejeitar a hipótese nula de que a média seja igual

a qualquer valor dentro do IC (e.g., Navidi, 2006) [75]. Por exemplo, se o IC obtido para a média do modelo “A” foi  $[0,900; 0,950]$  e estamos interessados em testar a hipótese nula  $H_0: \mu_A = 1$  contra  $H_1: \mu_A \neq 1$ , rejeitamos  $H_0$ , pois o IC  $[0,900; 0,950]$  não contém o valor 1. Por outro lado, para o IC  $[0,990; 1,020]$ , não rejeitamos  $H_0$ , pois este IC contém o valor 1.

O mecanismo para comparação de médias de diferentes modelos através de IC é semelhante ao explicado acima. Se o IC obtido para a média do modelo “A” contiver o valor da média amostral fornecida pelo modelo “B”, dizemos que não há evidências para rejeitar a hipótese nula de igualdade das médias (e.g., Smith, 1997) [85]. Neste caso, as hipóteses testadas são  $H_0: \mu_A = \mu_B$  contra  $H_1: \mu_A \neq \mu_B$ . Como exemplo, se o IC obtido para a média do modelo “A” foi  $[0,900; 0,950]$  e a média amostral do modelo “B” foi 0,920, não rejeita-se a hipótese de igualdade das médias, pois o IC contém o valor 0,920. Caso contrário, rejeita-se  $H_0$ .

Os intervalos de 95% de confiança construídos para os valores médios de  $U^I$  e de  $ARV^I$  nos experimentos com STI sintéticas estão apresentados nas Figuras A.1 a A.18 no Apêndice A na página 164. Os ICs nestas figuras são indicados por números conforme a Tabela 5.4, assim, a ordem de disposição dos ICs nos gráficos é a mesma das tabelas com os resultados. As linhas verticais tracejadas nos gráficos separam as classes de modelos conforme explicadas antes na Tabela 5.1. Os dois gráficos na parte superior das figuras correspondem aos ICs para os resultados com os intervalos onde houve predições invertidas e os dois na parte inferior correspondem aos ICs para os resultados desconsiderando as predições invertidas. Os dois gráficos da esquerda são relativos à medida  $U^I$  e os dois da direita são relativos à medida  $ARV^I$ .

A análise dos desempenhos dos métodos e modelos é realizada de forma global em relação aos PGDs considerados, no entanto, quando os resultados apresentarem comportamentos distintos em diferentes PGDs, nós realizamos a análise de maneira separada por PGD. Alguns resultados particulares que merecem destaque também são comentados. A descrição resumida de cada PGD foi apresentada na Tabela 5.3 na página 136. Os resultados que mais chamaram a atenção nos experimentos Monte Carlo estão descritos a seguir.

Com relação à adequação dos modelos ajustados e com base nos resultados do teste Ljung–Box realizados no conjunto de treinamento (coluna  $\tilde{H}_0^I$ ), observa-se que

Tabela 5.4: Número referente a cada modelo nas Figuras A.1 a A.18 e nas Figuras A.19 a A.33 apresentadas no Apêndice A.

Número	Modelo
01	ARMA <sub>c,r</sub>
02	MLP <sub>c,r</sub>
03	Híbrido <sub>c,r</sub>
04	AES <sup>I</sup>
05	AED <sup>I</sup>
06	HW <sup>I</sup>
07	iARMA-irrestrito
08	iARMA-restrito
09	iARMA-combinado
10	iARIMA
11	MLP <sup>I</sup>
12	Híbrido <sup>I</sup>
13	AES <sub>s,i</sub>
14	AED <sub>s,i</sub>
15	HW <sub>s,i</sub>
16	ARMA <sub>s,i</sub>
17	ARIMA <sub>s,i</sub>
18	MLP <sub>s,i</sub>
19	Híbrido <sub>s,i</sub>

todos os modelos apresentam taxas de rejeição da hipótese nula de não-correlação da série residual relativamente baixas, apesar de em alguns casos estas taxas ultrapassem os 10%. Para o PGD-I, os modelos que apresentaram as maiores taxas de rejeição  $\tilde{H}_0^I$ , tanto para  $n = 220$  quanto para  $n = 620$ , foram os métodos intervalares AES<sup>I</sup>, AED<sup>I</sup> e HW<sup>I</sup>, todas entre 11,6% e 14,6% (Tabelas A.1 e A.4). Neste processo gerador, as taxas observadas para estes métodos intervalares superam as taxas fornecidas pelos correspondentes métodos univariados ajustados às séries dos limites superiores e dos limites inferiores de modo independente, i.e., AES<sub>s,i</sub>, AED<sub>s,i</sub> e HW<sub>s,i</sub>. Ainda no PGD-I, o método Híbrido<sub>c,r</sub> é o que forneceu menor taxa de rejeição; 5,8% para STI de tamanho  $n = 220$  (Tabela A.1) e 5,0% para STI de tamanho  $n = 620$  (Tabela A.4).

Para o PGD-II (Tabelas A.7 e A.10), as maiores taxas de rejeição  $\tilde{H}_0^I$  são observadas nos modelos ajustados independentemente às séries dos centros e das amplitudes dos intervalos, ARMA<sub>c,r</sub>, MLP<sub>c,r</sub> e Híbrido<sub>c,r</sub>. As taxas para estes modelos ultrapassam os 12,2% em ambos os tamanho de amostras considerados. Como no PGD-I,

para o PGD-II, o método HW<sup>I</sup> fornece taxas acima dos 10%: 10,6% para  $n = 220$  (Tabela A.7) e 10,4% para  $n = 620$  (Tabela A.10).

Já para as STI sintéticas geradas de acordo com o PGD-III (Tabelas A.13 e A.16), as maiores taxas de rejeição  $\tilde{H}_0^I$  são observadas para os modelos ARMA<sub>c,r</sub> (11,4% e 10,4%), AES<sup>I</sup> (10,2% e 10,2%) e ARMA<sub>s,i</sub> (11,8% e 10,2%).

Ainda a respeito das taxas de rejeição da hipótese nula de não-correlação da série residual, no confronto dos três modelos, iARMA-restrito, iARMA-irrestrito e iARMA-combinado, podemos ver que as taxas são similares. No entanto, as taxas fornecidas pelo modelo combinado são as mais baixas em todas as situações consideradas. Estas três versões do modelo iARMA apresentam taxas de rejeição  $\tilde{H}_0^I$  inferiores às taxas dos modelos ARMA<sub>c,r</sub> e ARMA<sub>s,i</sub>, com algumas poucas exceções como, e.g., para o PGD-I (Tabelas A.1 e A.4) onde as taxas do modelo ARMA<sub>c,r</sub> são menores. Isto se deve ao fato de que o PGD-I favorece a estimação dos modelos baseados nas séries dos centros e das amplitudes dos intervalos; ver Tabela 5.3 na página 136. Entretanto, o modelo ARMA<sub>s,i</sub> não é melhor que as versões iARMA no PGD-II quanto as taxas de rejeição  $\tilde{H}_0^I$ .

Um ponto de suma relevância na análise dos resultados é a avaliação das previsões invertidas proporcionadas pelos modelos em questão. Um bom modelo para ajuste e previsão de STI deve evitar tais previsões, visto que elas podem comprometer o desempenho do modelo adotado. Neste sentido, alguns resultados merecem destaque e estão descritos a seguir.

A primeira constatação importante sobre as previsões invertidas refere-se ao modelo ARMA<sub>c,r</sub>. Para a nossa surpresa, este modelo apresentou valor médio para o número de previsões invertidas nulo em todas as situações experimentais consideradas. Apesar de não ser esperado, a razão para tal resultado é simples e deve-se ao modelo ARMA<sub>c,r</sub> ajustado à série univariada das amplitudes,  $X_t^r$ . No ajuste da série  $X_t^r$ , o modelo forneceu previsões univariadas sempre positivas para as amplitudes, apesar de este não ser isento de tal problema, pois não considera quaisquer restrições. Fornecendo sempre previsões positivas para a série das amplitudes, o modelo ARMA<sub>c,r</sub> está livre de previsões invertidas devido ao fato dos intervalos ajustados serem obtidos pela composição  $\hat{\mathbf{I}}_t = [\hat{X}_t^c - \frac{\hat{X}_t^r}{2}; \hat{X}_t^c + \frac{\hat{X}_t^r}{2}]$ . Entretanto, apesar destes resultados observados nestes experimentos, vale destacar que o modelo ARMA<sub>c,r</sub> pode fornecer previsões invertidas, basta que a série das amplitudes

tenha valor ajustado negativo. Em contrapartida, uma solução simples baseada em restrições nos parâmetros do modelo  $\text{ARMA}_{c,r}$  ajustado à série das amplitudes será capaz de tornar tal modelo livre de predições invertidas. Não adotamos esta solução, pois o objetivo é considerar o modelo sem restrições como apresentado originalmente em nosso artigo Maia et al. (2008) [67].

A rede neural  $\text{MLP}_{c,r}$  e o modelo  $\text{Híbrido}_{c,r}$  apresentaram valores médios não-nulos para o número de predições invertidas em poucas situações. Por exemplo, para STI de tamanho  $n = 620$  geradas pelos processos PGD-I, PGD-II e PGD-III, os valores médios de predições invertidas obtidos pela rede  $\text{MLP}_{c,r}$  foram, respectivamente, 2, 4 e 3 (Tabelas A.4, A.10 e A.16). O modelo  $\text{Híbrido}_{c,r}$  apresentou apenas dois valores não-nulos, ambos em STI de tamanho  $n = 620$  no PGD-II: um no conjunto de treinamento (Tabela A.10) e outro no conjunto de teste referente a 20 passos-à-frente (Tabela A.12). A razão para este reduzido valor de predições invertidas para estes dois modelos é a mesma explicada anteriormente para o modelo  $\text{ARMA}_{c,r}$ .

A ausência de predições invertidas no modelo  $\text{ARMA}_{c,r}$  e os baixos números médios de predições invertidas fornecidos pelos modelos  $\text{MLP}_{c,r}$  e  $\text{Híbrido}_{c,r}$  que ajustam a série dos centros e a série das amplitudes de modo independente demonstram a vantagem na modelagem de STI utilizando a série das amplitudes como variável. Para estes três modelos, estimativas não-negativas da série das amplitudes asseguram predições intervalares não-invertidas. Assim, implementar restrições aos parâmetros do modelo ajustado à série das amplitudes é uma solução viável para garantir a coerência matemática inerente aos dados simbólicos intervalares em que  $X_t^U \geq X_t^L$ ,  $t = 1, \dots, n$ . Esta solução foi utilizada com sucesso no ajuste de modelos de regressão em Lima Neto e De Carvalho (2010) [64].

Considerando os modelos que denominamos intervalares, ou seja, os modelos que ajustam as séries  $X_t^U$  e  $X_t^L$  de modo dependente, chamamos a atenção para alguns resultados sobre as predições invertidas. Primeiro, com relação às elevadas médias do número de predições invertidas observados nos resultados dos métodos de alisamento exponencial  $\text{AES}^I$ ,  $\text{AED}^I$  e  $\text{HW}^I$ . As médias fornecidas por estes métodos intervalares superam até mesmo as médias dos métodos univalorados  $\text{AES}_{s,i}$ ,  $\text{AED}_{s,i}$  e  $\text{HW}_{s,i}$  ajustados independentemente à série dos limites superiores e à série dos limites inferiores. Na soma de todas as médias de predições invertidas em todas

as situações, os métodos AES<sup>I</sup>, AED<sup>I</sup> e HW<sup>I</sup> apresentaram piores resultados dentre todos os modelos considerados.

Com relação aos modelos iARMA, sabe-se que as versões iARMA-restrito e iARMA-combinado são isentos de predições invertidas. A versão iARMA-irrestrito, porém, forneceu médias altas para o número de predições invertidas em várias situações. Por exemplo, no PGD-I com STI de tamanho  $n = 220$ , o modelo iARMA-irrestrito apresentou média de 10 predições invertidas no conjunto de treinamento (Tabela A.1) e média de 1 predição invertida no conjunto de teste relativo a 20 passos-à-frente (Tabela A.3). Esta característica realça a importância do uso das restrições impostas na estimação dos parâmetros do modelo iARMA.

Analisando os resultados referentes ao número médio de predições invertidas proporcionadas pelos modelos que ajustam a série dos limites superiores e a série dos limites inferiores de modo independente, i.e., AES<sub>s,i</sub>, AED<sub>s,i</sub>, HW<sub>s,i</sub>, ARMA<sub>s,i</sub>, ARIMA<sub>s,i</sub>, MLP<sub>s,i</sub> e Híbrido<sub>s,i</sub>, nota-se que os métodos de alisamento exponencial AES<sub>s,i</sub>, AED<sub>s,i</sub> e HW<sub>s,i</sub> são os de piores desempenhos. Como exemplo, estes três métodos têm, respectivamente, médias de inversão de 12, 9 e 11 predições no conjunto de treinamento do PGD-I com  $n = 220$  (Tabela A.1) e de 17, 19 e 12 predições no conjunto de treinamento do PGD-I com  $n = 620$  (Tabela A.4). Dentre estes três métodos, o AES<sub>s,i</sub> é o que mais fornece predições invertidas, seguido do método HW<sub>s,i</sub>. Por outro lado, dentre todos os modelos que ajustam a série dos limites superiores e a série dos limites inferiores de modo independente, o método Híbrido<sub>s,i</sub> é o mais eficiente com menos predições invertidas, seguido pela rede MLP<sub>s,i</sub>. Os modelos ARMA<sub>s,i</sub> e ARIMA<sub>s,i</sub> têm médias de inversão similares.

Comparando os modelos iARMA-irrestrito e ARMA<sub>s,i</sub>, únicos modelos ARMA que forneceram predições invertidas, observa-se que a versão intervalar irrestrita, iARMA-irrestrito, apresenta desempenho inferior com relação ao número médio de inversões.

Dentre as redes neurais MLP<sub>c,r</sub>, MLP<sup>I</sup> e MLP<sub>s,i</sub>, a primeira é a que apresenta menores médias do número de predições invertidas e a MLP<sub>s,i</sub> é a rede com mais predições invertidas na média. Mais uma vez, constatamos que o uso da série das amplitudes pode ter influenciado positivamente no ajuste de STI, pelo menos no que se refere ao desempenho do modelo com relação ao número de predições invertidas.

Feito o exame da correlação serial dos resíduos através do teste Ljung–Box e a

análise das previsões invertidas, passamos a avaliar os modelos e métodos via as medidas de precisão de previsões  $U^I$  e  $ARV^I$ . Os ICs apresentados nas Figuras A.1 a A.18 nos auxiliam nesta avaliação.

A vantagem das medidas de precisão utilizadas é que com elas podemos avaliar se os modelos e métodos são úteis para ajuste e previsão de STI. Como explicado anteriormente na Seção 4.7, a medida  $U^I$  compara os resíduos do modelo ajustado com os resíduos obtidos via o modelo ingênuo passeio aleatório, ou seja, em que o intervalo mais recente é considerado a previsão do intervalo futuro da STI. Assim, ao mesmo tempo em que comparamos diferentes modelos, esta medida nos indica se o modelo adotado é melhor que o passeio aleatório e, então, podemos concluir se o modelo é bom. A estatística  $ARV^I$  também nos proporciona uma comparação semelhante, porém, esta medida considera o intervalo médio da STI como previsão do intervalo futuro.

Baseado nas ideias acima, podemos concluir que os modelos propostos para previsão de STI são alternativas úteis para modelagem destas séries levando-se em conta os resultados observados nos conjuntos de treinamentos dos três PGDs (Tabelas A.1, A.4, A.7, A.10, A.13 e A.16). Para todos os PGDs considerados nos experimentos Monte Carlo, nos conjuntos de treinamento e independente do tamanho da amostra, os modelos adotados têm desempenhos superiores ao passeio aleatório e à média da STI com valores médios de  $U^I$  e de  $ARV^I$  menores que a unidade, i.e.,  $U^I < 1$  e  $ARV^I < 1$ . Os ICs construídos nos conjuntos de treinamento para as médias amostrais de  $U^I$  e  $ARV^I$  ao nível de 95% de confiança indicam a significância estatística destas conclusões, pois os ICs não contêm a unidade e estão abaixo dela. Estas conclusões são válidas também para os resultados que desconsideram os intervalos onde houve previsões invertidas.

De modo geral, para a classe dos modelos em que a série dos centros e a série das amplitudes dos intervalos são ajustadas através de dois modelos independentes,  $ARMA_{c,r}$ ,  $MLP_{c,r}$  e  $Híbrido_{c,r}$ , o modelo  $Híbrido_{c,r}$  é o que apresenta melhor desempenho quanto às estatísticas  $U^I$  e  $ARV^I$ . Em poucos casos, como para o PGD-I com  $n = 220$  no horizonte de previsão 20 passos-à-frente (Figura A.3), os ICs não indicam diferença significativa na média do  $Híbrido_{c,r}$  para os demais modelos desta classe.

Para a classe de modelos intervalares –  $AES^I$ ,  $AED^I$ ,  $HW^I$ , as três versões

iARMA, iARIMA, MLP<sup>I</sup> e Híbrido<sup>I</sup> –, chamamos a atenção para os desempenhos dos métodos de alisamento exponencial. Os métodos AES<sup>I</sup>, AED<sup>I</sup> e HW<sup>I</sup> têm as maiores médias para as medidas de precisão U<sup>I</sup> e ARV<sup>I</sup> desta classe. Além disso, os resultados não indicam qual o melhor método dentre eles três, pois a ordem de qualidade de desempenho se altera praticamente a cada situação considerada e de uma medida de precisão para a outra. Como exemplo, os ICs para U<sup>I</sup> nas Figuras A.1, A.2 e A.3 indicam melhor desempenho do HW<sup>I</sup>, enquanto os ICs para a estatística ARV<sup>I</sup> apontam melhor desempenho do AES<sup>I</sup>. Existem situações em que os ICs sugerem melhor desempenho do método AED<sup>I</sup>, como no PGD-II com  $n = 220$  baseado na medida U<sup>I</sup> (Figuras A.7, A.8 e A.9). Assim, não é possível concluir qual é o melhor método intervalar dentre estes três inspirados em alisamento exponencial.

As três versões do modelo iARMA apresentam desempenhos similares com respeito às medidas U<sup>I</sup> e ARV<sup>I</sup>. Na maioria das vezes, os ICs não indicam diferenças significativas nas médias para ambas as medidas de precisão, embora, para o PGD-III com  $n = 220$  para a medida ARV<sup>I</sup> (Figuras A.13, A.14 e A.15), os ICs acusam diferenças significativas para melhor desempenho do modelo iARMA-restrito. A similaridade dos desempenhos das três versões do modelo iARMA sugere que a solução da estimação restrita de parâmetros implementada no modelo iARMA-restrito é eficiente e, conseqüentemente, o seu uso para ajuste e previsão de STI é vantajoso, visto que este modelo é isento de predições invertidas.

O modelo iARMA-combinado foi aplicado com a intenção de eliminar as predições invertidas produzidas pelo modelo iARMA-irrestrito, substituindo estas predições pelas fornecidas pelo modelo iARMA-restrito. De fato, o modelo combinado está livre de predições invertidas, no entanto, o seu desempenho não apresentou melhora significativa em relação ao modelo iARMA-irrestrito.

O modelo intervalar iARIMA, por sua vez, apresenta desempenho inferior aos desempenhos das três versões do modelo iARMA. Observemos, por exemplo, os resultados obtidos para o PGD-I com  $n = 220$  apresentados nas Tabelas A.1, A.2 e A.3 e nas Figuras A.1, A.2 e A.3. Neles, notamos a inferioridade do modelo iARIMA em relação aos modelos iARMA-restrito, iARMA-irrestrito e iARMA-combinado.

Da mesma maneira que o modelo Híbrido<sub>c,r</sub> teve melhor desempenho dentre os modelos da classe de ajuste dos centros e das amplitudes de modo independente, na classe de modelos intervalares de ajuste dependente, o modelo Híbrido<sup>I</sup> superou

os demais. As médias de  $U^I$  e de  $ARV^I$  obtidas para este modelo são menores que as médias dos outros modelos na maioria das situações e, em muitos casos, os ICs indicam diferenças significativas. O modelo Híbrido<sup>I</sup> apresentou melhores desempenhos tanto no conjunto de treinamento para os dois tamanhos de amostra quanto nos conjuntos de teste de 5 e 20 passos-à-frente. Vale lembrar que este modelo combina o modelo iARMA-restrito com a rede neural MLP<sup>I</sup>.

A última classe de modelos, em que se ajusta um modelo univariado à série dos limites superiores,  $X_t^U$ , e outro modelo univariado à série dos limites inferiores,  $X_t^L$ , de modo independente, foi considerada com a finalidade de comparar seus resultados com os resultados obtidos para os correspondentes modelos desenvolvidos especificamente para modelagem de STI. Por esta razão, não faremos uma análise dos resultados fornecidos por esta classe de modelos de forma separada e, sim, confrontaremos com os respectivos modelos para STI. Ou seja, o objetivo é comparar os resultados entre os modelos  $ARMA_{c,r}$ , iARMA (as três versões) e  $ARMA_{s,i}$ , os resultados entre as redes neurais MLP<sub>c,r</sub>, MLP<sup>I</sup> e MLP<sub>s,i</sub>, os resultados entre os modelos AES<sup>I</sup> e AES<sub>s,i</sub>, os resultados entre os modelos AED<sup>I</sup> e AED<sub>s,i</sub> e assim por diante. Esta comparação visa identificar qual a melhor estratégia para se modelar e prever STI e, então, mostrar que os modelos intervalares de ajuste dependente são vantajosos.

De acordo com a comparação explicada acima, inicialmente consideramos os modelos  $ARMA_{c,r}$ ,  $ARMA_{s,i}$  e as três versões do modelo iARMA. Como dito antes, as três versões iARMA têm desempenhos similares, por isso, nos referimos a elas simplesmente como um único modelo: modelo iARMA. A primeira constatação nesta comparação é que, para o PGD-I (Figuras A.1 a A.6), o modelo  $ARMA_{c,r}$  é superior ao modelo  $ARMA_{s,i}$  em todas as situações consideradas. O PGD-I é aquele processo em que são geradas a série dos centros e a série das amplitudes,  $X_t^c$  e  $X_t^r$ , através de dois processos independentes. Por outro lado, para o PGD-II (Figuras A.7 a A.12), em que são geradas as séries  $X_t^U$  e  $X_t^L$  também independentes, o modelo  $ARMA_{s,i}$  supera o modelo  $ARMA_{c,r}$  na maioria das situações. Para o PGD-III, não há como apontar qual destes dois modelos tem melhor desempenho. Neste processo gerador, segundo a medida  $U^I$ , o modelo  $ARMA_{c,r}$  é melhor, porém, de acordo com a estatística  $ARV^I$  o modelo  $ARMA_{s,i}$  supera o modelo  $ARMA_{c,r}$ .

O resultado acima já era esperado, visto que o PGD-I favorece os modelos independentes ajustados às séries dos centros e das amplitudes dos intervalos, enquanto que o PGD-II favorece os modelos para ajuste das séries dos limites superior e inferior de modo independente.

Outra conclusão importante ainda acerca da comparação dos modelos  $\text{ARMA}_{c,r}$ ,  $\text{ARMA}_{s,i}$  e  $\text{iARMA}$  é que para os três PGDs considerados, tanto no conjunto de treinamento quanto nos conjuntos de teste, o modelo  $\text{iARMA}$  apresenta melhor desempenho. Isto nos leva a concluir que a estratégia adotada no modelo  $\text{iARMA}$  é útil na modelagem de STI.

Agora, comparamos os métodos intervalares inspirados em alisamento exponencial,  $\text{AES}^I$ ,  $\text{AED}^I$  e  $\text{HW}^I$ , com os respectivos métodos univalorados,  $\text{AES}_{s,i}$ ,  $\text{AED}_{s,i}$  e  $\text{HW}_{s,i}$ . Nota-se que, apesar dos métodos  $\text{AES}^I$ ,  $\text{AED}^I$  e  $\text{HW}^I$  não apresentarem bons desempenhos em relação à classe que se inserem, estes métodos são mais eficientes no ajuste e previsão de STI que os correspondentes métodos univalorados,  $\text{AES}_{s,i}$ ,  $\text{AED}_{s,i}$  e  $\text{HW}_{s,i}$ , respectivamente. Na maioria das situações, os ICs confirmam esta conclusão, embora, para alguns casos principalmente no PGD-II, as médias dos métodos  $\text{AES}^I$ ,  $\text{AED}^I$  e  $\text{HW}^I$  não apresentam diferenças significativas das médias dos métodos  $\text{AES}_{s,i}$ ,  $\text{AED}_{s,i}$  e  $\text{HW}_{s,i}$ . Veja, e.g., os ICs construídos para as médias de  $U^I$  para os métodos  $\text{AES}^I$  e  $\text{AES}_{s,i}$  nas Figuras A.7, A.8 e A.12.

Com respeito à comparação do modelo  $\text{iARIMA}$  com o modelo  $\text{ARIMA}_{s,i}$ , os resultados mostram claramente o melhor desempenho do modelo intervalar  $\text{iARIMA}$  em todas as situações. Os ICs indicam diferenças significativas entre as médias destes modelos para ambas as medidas de precisão consideradas. Assim, ajustar o modelo  $\text{iARIMA}$  à STI é mais eficiente que modelá-la através de dois modelos  $\text{ARIMA}$  ajustados independentemente às séries dos limites superiores e dos limites inferiores.

Na comparação entre as redes neurais  $\text{MLP}_{c,r}$ ,  $\text{MLP}^I$  e  $\text{MLP}_{s,i}$ , observa-se a superioridade da rede  $\text{MLP}^I$  em todas as situações consideradas. Mais uma vez, a metodologia empregada para desenvolvimento de modelos intervalares mostrou-se eficaz. Entre as redes  $\text{MLP}_{c,r}$  e  $\text{MLP}_{s,i}$  há um revezamento com relação ao melhor desempenho. Por exemplo, para o PGD-I (Figuras A.1 a A.6) a  $\text{MLP}_{c,r}$  é ligeiramente superior, enquanto que para o PGD-II (Figuras A.7 a A.12), em geral, a melhor é a  $\text{MLP}_{s,i}$ .

Também efetuamos a comparação entre os modelos híbridos inspirados na metodologia proposta por Zhang (2003) [94], Híbrido<sub>c,r</sub>, Híbrido<sup>I</sup> e Híbrido<sub>s,i</sub>. Os resultados para o PGD-I (Figuras A.1 a A.6) indicam, de modo geral, melhor desempenho do Híbrido<sub>c,r</sub> sobre o Híbrido<sub>s,i</sub> e o contrário para o PGD-II (Figuras A.7 a A.12). Dentre os três modelos híbridos, o modelo Híbrido<sup>I</sup> é superior no ajuste de STI. Estas conclusões são como as discutidas na comparação entre os modelos ARMA<sub>c,r</sub>, ARMA<sub>s,i</sub> e iARMA e também eram esperadas, visto que os desempenhos dos modelos híbridos são fortemente influenciados pelos desempenhos destes modelos. Vale lembrar que os modelos Híbrido<sub>c,r</sub>, Híbrido<sup>I</sup> e Híbrido<sub>s,i</sub> usam redes neurais MLP para ajustar os resíduos dos modelos ARMA<sub>c,r</sub>, iARMA-restrito e ARMA<sub>s,i</sub>, respectivamente.

Por fim, comparamos todos os 19 modelos globalmente quanto aos seus desempenhos no ajuste e previsão de STI de acordo com as medidas de precisão  $U^I$  e  $ARV^I$ . Na classe de modelos em que o intervalo é previsto de modo dependente, i.e., a previsão do limite superior leva em consideração informações relativas ao limite inferior e vice-versa, estão contidos os modelos de melhores desempenhos quanto à precisão dos intervalos ajustados e dos intervalos previstos. Estes modelos são o iARMA-irrestrito, o iARMA-restrito, o iARMA-combinado, o iARIMA, a MLP<sup>I</sup> e, especialmente, o modelo Híbrido<sup>I</sup>. Este último é o modelo que exhibe melhores resultados quanto às medidas de precisão dentre todos os modelos considerados. Estas conclusões são comprovadas através dos ICs construídos para as médias de  $U^I$  e de  $ARV^I$ . Nas Figuras A.1 a A.18, os ICs referentes aos modelos iARMA-irrestrito, iARMA-restrito, iARMA-combinado, iARIMA, MLP<sup>I</sup> e Híbrido<sup>I</sup> são, nesta ordem, os numerados de 07 a 12. Como podemos ver nestas figuras, de modo geral, os ICs para estes modelos estão abaixo dos ICs para os demais modelos indicando diferenças significativas nas médias das medidas  $U^I$  e de  $ARV^I$ .

## 5.2 Aplicação a STI reais

Esta seção demonstra a utilidade dos métodos e modelos apresentados nesta tese em aplicações a dados de STI reais. Nós consideramos cinco STI reais e aplicamos os mesmos 19 modelos descritos na Tabela 5.1 na página 119.

As três primeiras STI reais analisadas correspondem a dados de valores de negociações de ações de três companhias de diferentes segmentos do mercado. Nós consideramos os valores de ações das companhias *Petróleo Brasileiro S.A.* (Petrobras), *Google Inc.* e *Gol Linhas Aéreas Inteligentes S.A.* O limite superior e o limite inferior dos intervalos consistem, respectivamente, do valor mais alto de negociação da ação durante o dia e do valor mais baixo de negociação da ação durante o dia. Estas STI são compostas de dados diários e estão disponíveis na página <http://finance.yahoo.com> (acesso em março de 2008). Em análise de séries temporais univariadas, as séries financeiras, especialmente, as do mercado de ações, são tipicamente de difícil modelagem e previsão.

Outras duas STI reais equivalem aos dados de temperaturas registradas nas cidades do *Rio de Janeiro* e de *Nova Iorque*. Nestas STI, cada intervalo é composto da maior temperatura registrada durante o dia e da menor temperatura registrada durante o dia. As duas STI de temperaturas também são diárias e estão disponíveis na página <http://www.wunderground.com> (acesso em março de 2008). Apesar dos efeitos sazonais resultantes das estações do ano, séries de fatores climáticos, tais como, de temperatura são de mais fácil modelagem e previsão.

Nas Tabelas 5.5 e 5.6 estão, respectivamente, as descrições e as características das STI reais analisadas. Estas STI estão representadas graficamente nas Figuras 5.7, 5.8, 5.9, 5.10 e 5.11. Nestas figuras podemos observar algumas características no comportamento das STI analisadas.

Para avaliar os modelos em STI reais quanto à qualidade de previsão de intervalos futuros, nós dividimos as séries em conjunto de treinamento com  $n - 20$  intervalos e conjunto de teste com 20 intervalos. Como nas STI sintéticas, consideramos dois conjuntos de teste: um de 5 passos-à-frente e um de 20 passos-à-frente.

Em cada STI real analisada, nós computamos as medidas de precisão  $U^I$  e  $ARV^I$  para cada modelo. Desta vez, as medidas de precisão foram computadas uma única vez por STI, por isso, não fornecemos os desvios-padrão. Além das medidas de precisão, nós computamos o número de predições invertidas para cada modelo. Os resultados são apresentados com os intervalos onde houve predições invertidas e sem estes intervalos.

Também realizamos o teste Ljung–Box para a série dos resíduos no limite superior e para a série dos resíduos no limite inferior no conjunto de treinamento baseado

Tabela 5.5: Descrição das STI reais analisadas.

STI	Descrição
Petrobras	Série dos intervalos de preços de ações da companhia Petróleo Brasileiro S.A.
Google	Série dos intervalos de preços de ações da companhia Google, Inc.
Gol	Série dos intervalos de preços de ações da companhia Gol Linhas Aéreas Inteligentes S.A.
Rio de Janeiro	Série dos intervalos de temperaturas na cidade do Rio de Janeiro, Brasil.
Nova Iorque	Série dos intervalos de temperaturas na cidade de Nova Iorque, EUA.

Tabela 5.6: Período e tamanho das STI reais analisadas.

STI	Período	$n$
Petrobras	2 de julho de 2007 a 13 de março de 2008	177
Google	10 de agosto de 2006 a 10 de março de 2008	397
Gol	13 de dezembro de 2005 a 13 de março de 2008	565
Rio de Janeiro	1 de janeiro de 2006 a 31 de março de 2008	820
Nova Iorque	1 de janeiro de 2005 a 31 de março de 2008	1186

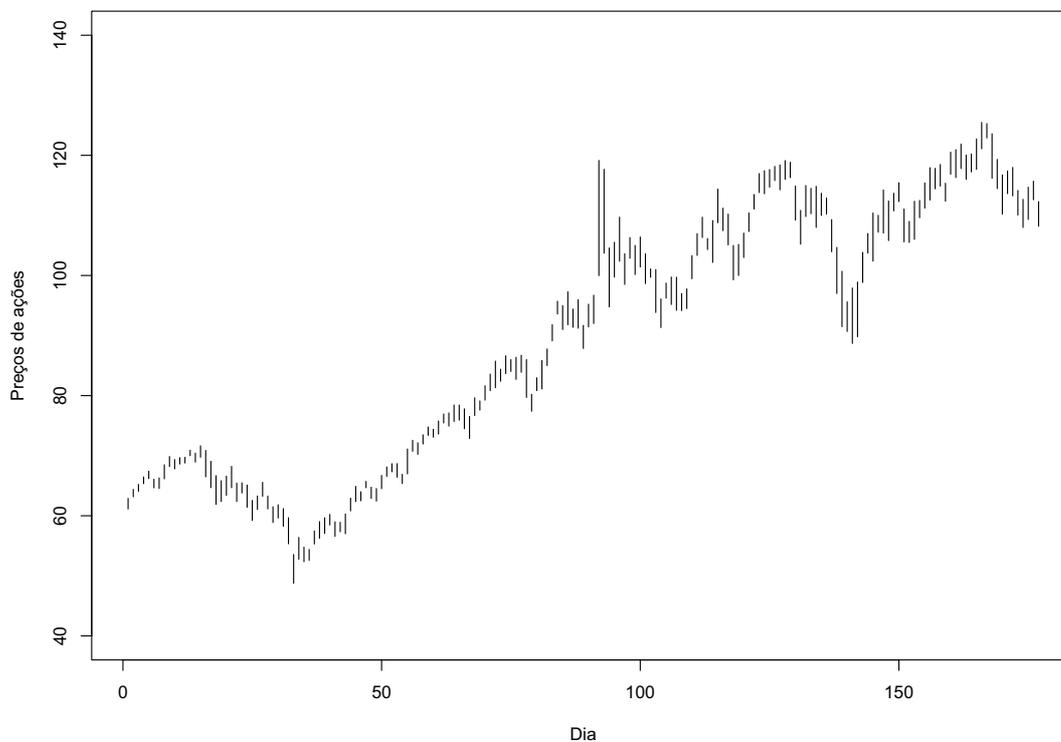
em 30 autocorrelações residuais obtidas para cada modelo. Assim, nós assinalamos os modelos que tiveram a hipótese nula de não-correlação da série residual rejeitada ou no limite superior ou no limite inferior no conjunto de treinamento.

As Tabelas A.19 a A.33 apresentadas no Apêndice A a partir da página 201 exibem os resultados obtidos para as cinco STI reais, no conjunto de treinamento ( $n_T = n - 20$ ) e dois conjuntos de teste correspondentes aos horizontes de previsão de 5 e 20 passos-à-frente.

Na primeira coluna das Tabelas A.19, A.22, A.25, A.28 e A.31 referentes aos conjuntos de treinamento estão assinalados (com a símbolo  $\bullet$ ) os modelos que tiveram a hipótese nula de não-correlação da série residual rejeitada ou no limite superior ou no limite inferior. A segunda coluna exhibe o número de predições invertidas para cada modelo e as demais colunas apresentam os resultados para as medidas de precisão  $U^I$  e  $ARV^I$ . As tabelas exibem os resultados destas medidas considerando as predições invertidas e sem considerá-las.

Para proporcionar a visualização dos resultados com base nas medidas  $U^I$  e  $ARV^I$  e, conseqüentemente, facilitar a comparação dos desempenhos dos modelos, nós construímos gráficos para os valores destas medidas fornecidos por cada um

Figura 5.7: STI real de preços diários de ações da Petrobras.



destes modelos. Os gráficos com os resultados para as STI reais estão apresentados nas Figuras A.19 a A.33. Os números 01 a 19 nas figuras referem-se aos modelos considerados e seguem a mesma ordem dos modelos nas tabelas dos resultados; ver Tabela 5.4 na página 140.

Os resultados referentes aos desempenhos dos modelos nas STI reais são análogos aos resultados observados para as STI sintéticas descritos na Seção 5.1.2. A seguir estão os comentários seguindo a mesma ordem dos descritos para as STI sintéticas.

Iniciamos com os comentários sobre o teste Ljung–Box da hipótese nula de não-correlação da série residual. A primeira observação sobre este teste é que o método AES<sup>I</sup> foi assinalado em quatro das cinco STI reais consideradas. Apenas para a STI de preços de ações da Google que esta hipótese não foi rejeitada para este método (Tabela A.22). Vale lembrar que este método está entre os que apresentaram maiores taxas de rejeição desta hipótese nas STI sintéticas. O método AES<sub>s,i</sub> foi assinalado em duas destas STI reais e os modelos MLP<sub>c,r</sub>, HW<sup>I</sup>, AES<sub>s,i</sub>, ARMA<sub>s,i</sub> e MLP<sub>s,i</sub> foram assinalados em apenas uma.

Figura 5.8: STI real de preços diários de ações da Google.

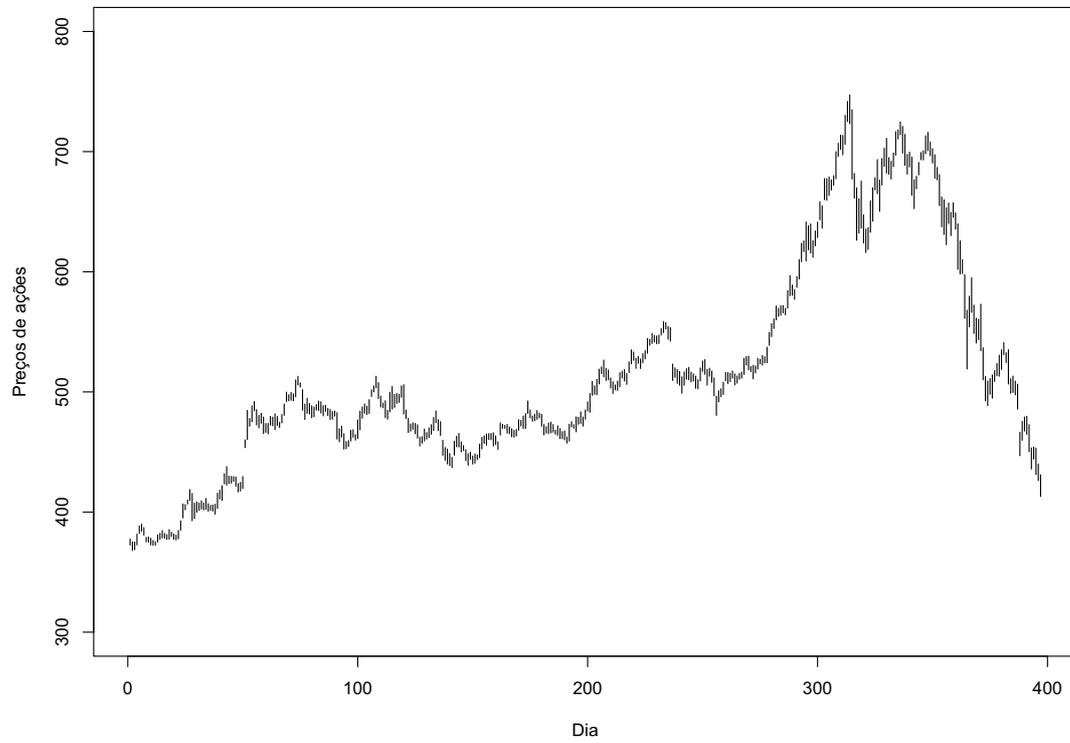


Figura 5.9: STI real de preços diários de ações da Gol.

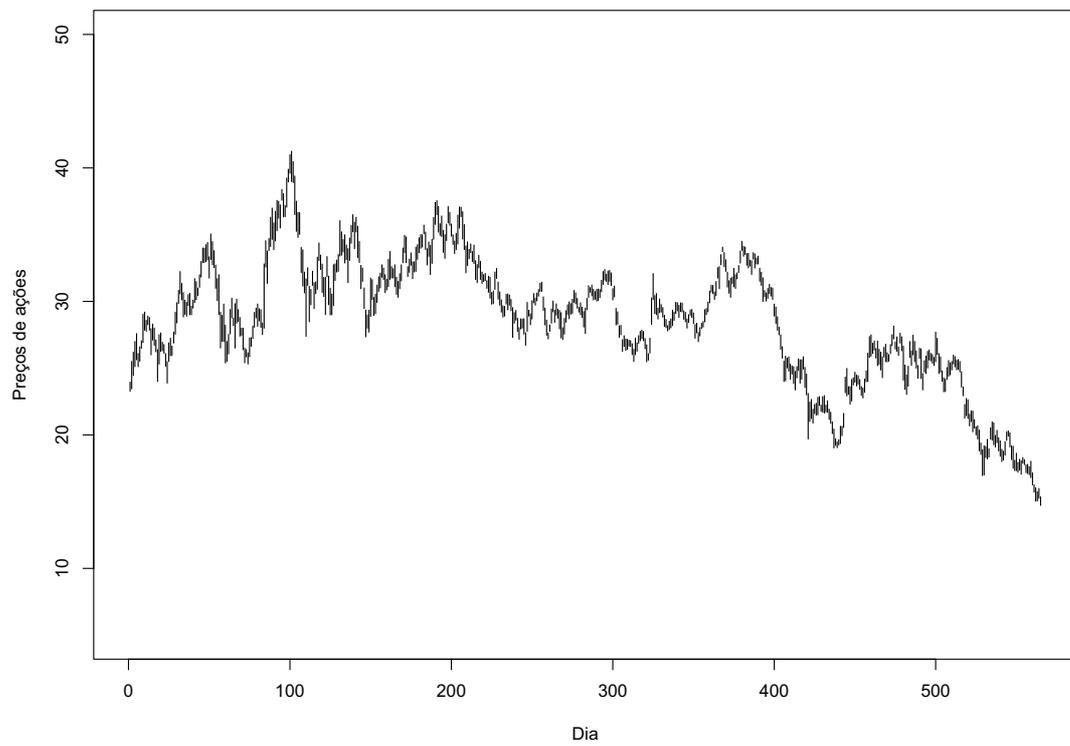


Figura 5.10: STI real de temperaturas diárias no Rio de Janeiro.

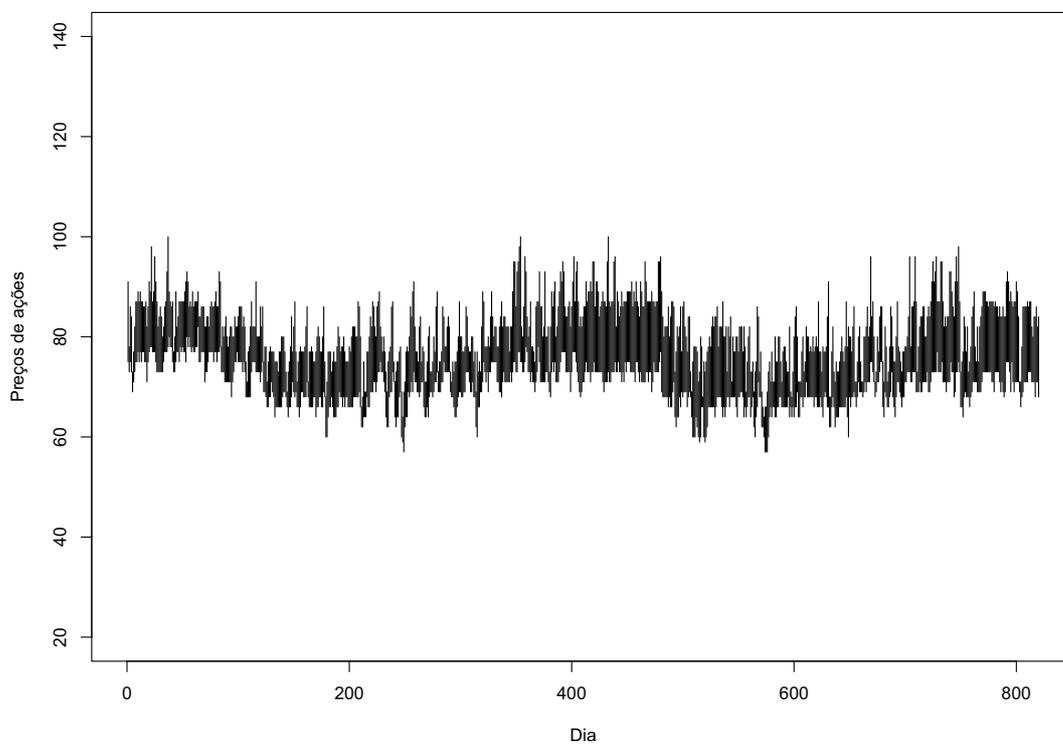
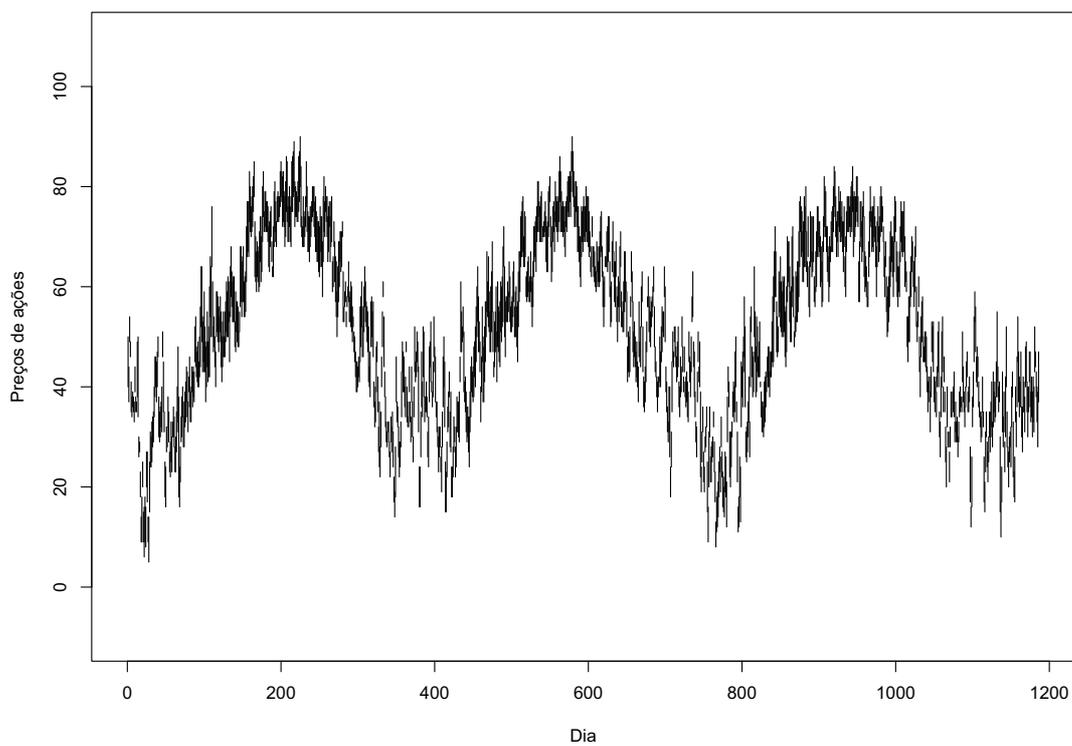


Figura 5.11: STI real de temperaturas diárias em Nova Iorque.



Quanto às predições invertidas produzidas pelos modelos, os resultados são interessantes. Dentre os modelos da classe do ajuste dos centros e das amplitudes de modo independente, o único que produziu predições invertidas foi a rede neural  $MLP_{c,r}$ : 2 inversões no conjunto de treinamento da STI da Google (Tabela A.22) e 1 no conjunto de treinamento da STI da Gol (Tabela A.25). O modelo  $ARMA_{c,r}$ , como nos experimentos com STI sintéticas, não deu origem a predições invertidas.

Para os modelos que se inserem na classe do ajuste dos limites superiores e inferiores de modo dependente, o modelo  $iARMA$ -irrestrito foi o que forneceu maior número de predições invertidas. No somatório dos resultados para as cinco STI reais, no conjunto de treinamento e nos conjuntos de teste, este modelo apresentou 47 predições invertidas, seguido do  $AES^I$  com 38 inversões e do  $iARIMA$  com 35. Os modelos  $AED^I$ ,  $HW^I$  e a rede  $MLP^I$  apresentam, respectivamente, 29, 27 e 14 predições invertidas no total. O modelo  $Híbrido^I$  não apresentou inversões.

Os modelos da classe de ajuste das séries dos limites superiores e dos limites inferiores de modo independente também produziram números elevados de predições invertidas. A ordem dos modelos que produziram maiores números de predição invertidas, nesta classe, são:  $AES_{s,i}$  com 39 inversões,  $ARMA_{s,i}$  com 35,  $AED_{s,i}$  com 23,  $Híbrido_{s,i}$  com 21,  $ARIMA_{s,i}$  com 20,  $MLP_{s,i}$  com 20 e  $HW_{s,i}$  com 13. Nesta classe, todos os sete modelos apresentaram predições invertidas.

Como podemos constatar nos comentários acima, o modelo  $iARMA$ -irrestrito foi, dentre os 19 modelos, o de pior desempenho com respeito ao número de predições invertidas nas cinco STI reais consideradas. Como dito antes, este modelo apresentou, no total, 47 predições invertidas. Este fato sugere que a solução adotada através da estimação restrita de parâmetros no modelo  $iARMA$ -restrito e, conseqüentemente, o seu uso para ajuste e previsão de STI é vantajoso. Lembramos que o modelo  $iARMA$ -restrito é livre de predições invertidas.

Os resultados para as STI reais relacionados aos desempenhos dos modelos acerca das medidas de precisão  $U^I$  e  $ARV^I$  revelam um cenário semelhante ao descrito nos experimentos com as STI sintéticas. Os gráficos para os valores destas medidas são apresentados no Apêndice A a partir da página 201 nas Figuras A.19 a A.33. Estes gráficos nos ajudam a examinar os comportamentos dos 19 modelos em aplicações às STI reais.

Como nos experimentos simulados, podemos concluir que os modelos propostos

para previsão de STI são alternativas úteis para modelagem destas séries levando-se em conta os resultados observados nos conjuntos de treinamentos das cinco STI reais consideradas (Tabelas A.19, A.22, A.25, A.28 e A.31). Esta conclusão vem do fato de que os modelos  $ARMA_{c,r}$ ,  $MLP_{c,r}$ ,  $Híbrido_{c,r}$ ,  $AES^I$ ,  $AED^I$ ,  $HW^I$ ,  $iARMA$ -irrestrito,  $iARMA$ -restrito,  $iARMA$ -combinado,  $iARIMA$ ,  $MLP^I$  e  $Híbrido^I$  forneceram valores para as medidas  $U^I$  e  $ARV^I$  abaixo da unidade nos conjuntos de treinamento para as cinco STI reais. Isto se estende para os resultados que descon sideram os intervalos onde houve predições invertidas.

Para a classe de ajuste independente dos centros e das amplitudes, o melhor desempenho é proveniente do modelo  $Híbrido_{c,r}$  e para a classe de ajuste dependente dos limites superiores e inferiores, o melhor desempenho vem do modelo  $Híbrido^I$ . Estes dois modelos geralmente superam os demais modelos de suas classes quanto aos valores das medidas  $U^I$  e  $ARV^I$ .

Os piores desempenhos dentre os modelos  $AES^I$ ,  $AED^I$ ,  $HW^I$ ,  $iARMA$ -irrestrito,  $iARMA$ -restrito,  $iARMA$ -combinado,  $iARIMA$ ,  $MLP^I$  e  $Híbrido^I$  são observados para os três métodos inspirados em alisamento exponencial:  $AES^I$ ,  $AED^I$  e  $HW^I$ . Estes métodos têm as maiores médias para as medidas de precisão  $U^I$  e  $ARV^I$  e, algumas vezes, eles são superados pelo modelos ajustados às séries dos centros e das amplitudes de modo independente e por alguns modelos ajustados às séries dos limites superiores e dos limites inferiores modo independente.

Por sua vez, os métodos  $AES^I$ ,  $AED^I$  e  $HW^I$  são melhores que os métodos univariados  $AES_{s,i}$ ,  $AED_{s,i}$  e  $HW_{s,i}$  que ajustam dois modelos independentes aos limites dos intervalos. Para a STI das temperaturas de Nova Iorque (Tabelas A.31, A.32 e A.33) que tem comportamento sazonal, o método  $HW^I$  foi muito melhor que os métodos  $AES^I$  e  $AED^I$ . Com base na medida  $U^I$  no conjunto de treinamento para esta STI (Figura A.31), o método  $HW^I$  superou as três versões do modelo  $iARMA$  e, ainda, o modelo  $iARIMA$  e a rede  $MLP^I$ .

Nos experimentos com STI reais, as três versões do modelo  $iARMA$  também têm desempenhos similares com respeito às medidas  $U^I$  e  $ARV^I$ . Este comportamento reforça a ideia de que a solução da estimação restrita adotada no modelo  $iARMA$ -restrito é eficiente. Além disso, os três modelos  $iARMA$  apresentam valores para as medidas  $U^I$  e  $ARV^I$  abaixo dos fornecidos pelos modelos  $ARMA_{c,r}$  e  $ARMA_{s,i}$  para as cinco STI reais analisadas, tanto no conjunto de treinamento quanto nas previsões

de 5 e 20 passos-à-frente. Estas conclusões sugerem que o uso do modelo iARMA-restrito é realmente vantajoso, pelo desempenho e por ele ser livre de predições invertidas.

Na comparação do modelo ARMA que ajusta as séries dos centros e das amplitudes independentemente,  $ARMA_{c,r}$ , com o modelo ARMA análogo que, por sua vez, ajusta as séries dos limites superiores e inferiores independentemente,  $ARMA_{s,i}$ , nota-se uma ligeira vantagem do modelo  $ARMA_{s,i}$  com relação aos valores de  $U^I$  e de  $ARV^I$ . No entanto, o modelo  $ARMA_{s,i}$  é vencido pela quantidade de predições invertidas, pois, como dito anteriormente, o  $ARMA_{c,r}$  não forneceu inversões.

Como nas STI sintéticas, no ajuste das STI reais, o modelo iARIMA apresenta desempenho inferior aos desempenhos das três versões do modelo iARMA. Como exceção, o modelo iARIMA foi superior aos três modelos iARMA na previsão 20 passos-à-frente para as STI de temperaturas do Rio de Janeiro e de Nova Iorque referindo-se à medida  $U^I$  (Figuras A.30 e A.33).

Na comparação do modelo iARIMA com o modelo  $ARIMA_{s,i}$ , os resultados com STI reais confirmam a superioridade do modelo intervalar iARIMA em todas as situações. Assim, ajustar o modelo iARIMA à STI é mais eficiente que modelá-la através de dois modelos ARIMA ajustados independentemente às séries dos limites superiores e dos limites inferiores.

A rede neural  $MLP^I$  também supera as redes neurais que ajustam duas séries de modo independente,  $MLP_{c,r}$  e  $MLP_{s,i}$ , em todas as STI reais. Na comparação das rede neurais  $MLP_{c,r}$  e  $MLP_{s,i}$  há um revezamento com relação ao melhor desempenho, porém, a rede  $MLP_{s,i}$  aparece mais vezes com melhor desempenho. Da mesma forma, o modelo Híbrido<sup>I</sup> tem melhor desempenho que os modelos Híbrido<sub>c,r</sub> e Híbrido<sub>s,i</sub>. No confronto dos modelos Híbrido<sub>c,r</sub> e Híbrido<sub>s,i</sub>, não há consenso sobre qual o melhor no ajuste das STI reais, pois eles se revezam quanto ao melhor desempenho acerca das medidas  $U^I$  e  $ARV^I$ .

Como podemos observar nos comentários feitos sobre os resultados nas cinco STI reais, é também na classe de modelos em que o intervalo é previsto de modo dependente que estão os modelos de melhores desempenhos. Como nas STI sintéticas, os melhores modelos nas STI reais são: o iARMA-irrestrito, o iARMA-restrito, o iARMA-combinado, o iARIMA, a  $MLP^I$  e, sobretudo, o modelo Híbrido<sup>I</sup> que tem o melhor desempenho dentre todos os 19 modelos considerados.

# Capítulo 6

## Conclusões e trabalhos futuros

Nesta tese são estendidos técnicas e modelos destinados à análise de séries temporais do tipo intervalo e abre espaço para novas discussões sobre o tema. A lista de trabalhos futuros relacionados não é extensa. De fato, citaremos algumas considerações que ainda poderiam estar inseridas no escopo desta tese.

Neste capítulo final da tese, nós apresentamos as conclusões alcançadas com o trabalho desenvolvido durante o doutorado e, em seguida, sugerimos alguns tópicos para trabalhos futuros.

### 6.1 Conclusões

A presente tese encontra-se dividida em duas partes: Parte I – *Métodos de ajuste independente* – e Parte II – *Métodos de ajuste dependente*. Os métodos diferem pela metodologia adotada para a modelagem de STI. Nos métodos da Parte I – *Métodos de ajuste independente* – são ajustados dois modelos de modo independente; um à série dos centros dos intervalos e outro à série das amplitudes. No ajuste da série dos centros não há interferência de informações da série das amplitudes e, da mesma forma, no ajuste da série das amplitudes não há interferência de informações da série dos centros. Por isso foram denominados de *Métodos de ajuste independente*. Na Parte II – *Métodos de ajuste dependente* –, a metodologia adotada implica no ajuste da STI através de mecanismos que modelam a série dos limites superiores e a série dos limites inferiores, em vez das séries dos centros e amplitudes. Esta é a primeira diferença com relação aos modelos da Parte I – *Métodos de ajuste*

*independente*. Outra diferença está no fato dos métodos da Parte II – *Métodos de ajuste dependente* – serem dependentes, ou seja, o modelo que ajusta a série do limite superior contém informações relativas ao limite inferior e vice-versa. Estes modelos foram construídos partindo-se do princípio de que a trajetória do limite superior pode ser afetada por realizações do limite inferior e que a trajetória do limite inferior pode ser afetada por realizações do limite superior. Sendo assim, os métodos tratados na segunda parte da tese foram denotados por *Métodos de ajuste dependente*.

No início da Parte I – *Métodos de ajuste independente* –, nós discutimos o modelo de regressão linear robusto para dados de natureza intervalar. O modelo apresentado é uma extensão do modelo de regressão linear da mediana para dados univariados e a metodologia adotada é análoga e inspirada nos modelos de regressão linear propostos por Lima Neto et al. (2004) [62] e Lima Neto e De Carvalho (2008) [63] para variáveis intervalares. Essa discussão teve por objetivo apresentar as origens dos modelos de ajuste independente propostos para STI aplicados aos centros e às amplitudes das variáveis intervalares. Vale destacar que estes modelos também podem ser aplicados, de modo independente, aos limites superiores e aos limites inferiores destas variáveis.

Em seguida foram apresentados os métodos de ajuste independente para modelagem e previsão de séries temporais do tipo intervalo no contexto da ADS baseados nas séries dos centros e das amplitudes. Os métodos consistem de adaptações dos modelos ARMA, de redes neurais MLP e de uma metodologia híbrida proposta por Zhang (2003) [94]. Estes métodos propostos para previsão de STI fazem parte do nosso artigo Maia et al. (2008)[67].

Na Parte II – *Métodos de ajuste dependente* –, utilizamos um procedimento diferente do apresentado na Parte I – *Métodos de ajuste independente* – para produzir extensões de métodos clássicos de séries temporais para análise de STI. O procedimento é inédito e consiste da substituição do escalar  $y_t$ , nos métodos usuais, pelo vetor intervalo  $\mathbf{I}_t$ . Consideramos que o intervalo é representado por um vetor bidimensional da forma  $\mathbf{I}_t = [X_t^L; X_t^U]$ , em que  $X_t^L$  é o próprio valor do limite inferior (*lower bound*) e  $X_t^U$  é o próprio valor do limite superior (*upper bound*).

Com base no procedimento descrito acima, iniciamos a segunda parte da tese com a proposta de novas formas de medir a autocorrelação e a autocovariância

amostrais em STI e apresentamos o correlograma para séries intervalares. Adicionalmente, apresentamos novas expressões para a correlação e para a covariância amostrais em variáveis aleatórias simbólicas do tipo intervalo. Nestas estatísticas, valores observados para o limite superior e para o limite inferior são considerados simultaneamente na descrição de dados intervalares e, sobretudo, na descrição de STI. Os resultados mostram que estas estatísticas são ferramentas úteis e podem auxiliar na análise de dados desta natureza.

Para extensão dos modelos de previsão, além da substituição do escalar  $y_t$  pelo vetor intervalo  $\mathbf{I}_t$ , outro aspecto foi considerado. Em modelos univalorados os parâmetros são dados por escalares, enquanto que nos métodos que propomos para STI, empregamos matrizes  $(2 \times 2)$  de parâmetros. Estas matrizes assumem papel fundamental para modelagem de STI, pois através dos seus elementos que é possível inserir informações do limite inferior na previsão do limite superior e vice-versa.

Na segunda parte da tese, nós apresentamos extensões para sete métodos e modelos clássicos de previsão de séries temporais univaloradas para modelagem de STI. Os três primeiros são inspirados em métodos de alisamento exponencial e os parâmetros são estimados através de técnicas de otimização com restrições em caixa para assegurar que os parâmetros tenham suporte entre 0 e 1. Estes métodos são: o alisamento exponencial simples, o alisamento exponencial duplo e o método de Holt–Winters.

O quarto modelo proposto para previsão de STI na Parte II – *Métodos de ajuste dependente* – é inspirado no modelo ARMA univalorado. Denotamos o modelo proposto por iARMA. Para este modelo, nós propomos duas estratégias para estimação das matrizes de parâmetros, ambas derivadas de métodos clássicos de estimação de modelos ARMA univalorados. A diferença entre as duas está no fato de no segundo procedimento usarmos técnicas de otimização com restrições nos parâmetros com a finalidade de assegurar que os intervalos ajustados satisfaçam a condição  $\hat{X}_t^U \geq \hat{X}_t^L$ .

Depois de estender os quatro métodos descritos acima e inseridos na Parte II – *Métodos de ajuste dependente* – da tese, nós apresentamos uma rede neural MLP para ajuste de STI de modo dependente e um método híbrido que usa, simultaneamente, o modelo iARMA e esta rede neural MLP para ajuste de STI.

Com a finalidade de avaliar se os modelos propostos com o objetivo específico de

modelar STI são eficientes, nós comparamos os seus desempenhos com os desempenhos de modelos univalorados de séries temporais. Os modelos univalorados foram ajustados às séries dos limites superiores e dos limites inferiores separadamente e independentemente. Ou seja, as STI foram encaradas como duas séries univaloradas independentes, uma representada pela série dos limites superiores dos intervalos,  $\{X_t^U\}$ , e a outra representada pela série dos limites inferiores dos intervalos,  $\{X_t^L\}$ .

Para avaliar os desempenhos dos modelos propostos para STI e efetuar a comparação entre eles e com os modelos univalorados, nós adaptamos duas medidas de precisão utilizada em dados univalorados e as chamamos de *estatística U de Theil intervalar* e *variância relativa intervalar média*. Além de computar estas medidas nos experimentos, nós examinamos a correlação serial dos resíduos fornecidos por cada um dos modelos ajustados. Para isso, nós executamos o teste Ljung–Box da hipótese nula de não-correlação da série residual para as séries dos resíduos no limite superior e no limite inferior separadamente.

Os modelos também foram avaliados a respeito da coerência matemática  $\hat{X}_t^U \geq \hat{X}_t^L$ . O interesse é avaliar se os modelos são capazes de considerar a relação de ordem existente nos dados de natureza intervalar e, então, produzir o mínimo de previsões invertidas.

Iniciamos nossos experimentos através de simulação Monte Carlo com STI sintéticas e, em seguida, conduzimos aplicações em STI reais. Os dados reais são compostos de cinco STI, sendo três STI de preços diários de ações de três companhias de diferentes segmentos do mercado e duas STI de temperaturas registradas durante o dia nas cidades de Nova Iorque e do Rio de Janeiro.

A primeira conclusão importante refere-se ao modelo ARMA ajustado de modo independente às séries dos centros e das amplitudes que chamamos de  $\text{ARMA}_{c,r}$ . A estimação deste modelo não incorpora qualquer restrição nos parâmetros para assegurar que  $\hat{X}_t^U \geq \hat{X}_t^L$ , porém, ele não forneceu previsões invertidas. A razão para tal resultado é simples e deve-se ao modelo  $\text{ARMA}_{c,r}$  ajustado à série das amplitudes. No ajuste da série das amplitudes, o modelo forneceu previsões sempre positivas, o que impediu previsões invertidas por tal modelo. Entretanto, apesar de não terem sido observadas, o modelo  $\text{ARMA}_{c,r}$  pode fornecer previsões invertidas, basta que a série das amplitudes tenha valor ajustado negativo. Alguns pesquisadores defendem o uso de modelos para as amplitudes no ajuste de dados intervalares pela facilidade

de impor restrições nos parâmetros.

Apesar dos modelos de alisamento exponencial estendidos para previsão de STI, AES<sup>I</sup>, AED<sup>I</sup> e HW<sup>I</sup>, apresentares desempenhos superiores aos respectivos métodos univalorados de ajuste independente, eles proporcionam predições invertidas sugerindo que a implementação de métodos que assegurem  $\hat{X}_t^U \geq \hat{X}_t^L$  faz-se necessária.

Com relação aos modelos iARMA, sabe-se que a versão iARMA-restrito, por construção, não fornece predições invertidas e ele apresentou bom desempenho no ajuste e previsão de STI com relação às medidas de precisão consideradas. Por outro lado, a versão iARMA-irrestrito forneceu muitas predições invertidas, porém, seu desempenho quanto às medidas de precisão é similar ao modelo restrito. Esta similaridade nos desempenhos sugere que a solução da estimação restrita de parâmetros implementada no modelo iARMA-restrito é eficiente e, conseqüentemente, o seu uso para ajuste e previsão de STI é vantajoso, visto que este modelo é isento de predições invertidas.

Observamos ainda que as três versões do modelo iARMA têm melhores desempenhos que os modelos de ajuste independente ARMA<sub>c,r</sub> e ARMA<sub>s,i</sub> e que o modelo iARIMA é superior ao modelo ARIMA<sub>s,i</sub>. Observamos também que a rede neural MLP<sup>I</sup> supera as redes neurais para ajuste independente, MLP<sub>c,r</sub> e MLP<sub>s,i</sub>. Estes fatos nos levam a concluir que a estratégia adotada nos modelos intervalares iARMA e iARIMA é útil na modelagem de STI.

Os modelos intervalares de ajuste dependente se mostraram úteis na previsão de STI e, sobretudo, foram superiores às suas respectivas versões univaloradas em que as séries dos limites superiores e dos limites inferiores são ajustadas individualmente através de modelos univalorados independentes.

Os desempenhos dos modelos híbridos demonstram a potencialidade das redes neurais do tipo MLP para modelagem de processos temporais não capturados pelos modelos lineares ARMA e iARMA. Ou seja, a rede MLP utilizada no ajuste independente e a MLP<sup>I</sup> no ajuste dependente, quando usadas na segunda etapa da metodologia híbrida na modelagem dos resíduos destes modelos lineares, foram capazes de melhorar a precisão das previsões.

Por fim, na comparação de todos os modelos considerados na previsão de STI acerca das medidas de precisão, concluímos que na classe de modelos em que o intervalo é previsto de modo dependente, i.e., a previsão do limite superior leva em

consideração informações relativas ao limite inferior e vice-versa, estão contidos os modelos de melhores desempenhos quanto à precisão dos intervalos ajustados e dos intervalos previstos. Estes modelos são o iARMA-irrestrito, o iARMA-restrito, o iARMA-combinado, o iARIMA, a MLP<sup>I</sup> e, especialmente, o modelo Híbrido<sup>I</sup> que tem o melhor desempenho dentre todos os modelos considerados.

Dadas as conclusões acima, nós recomendamos o uso do modelo híbrido que combina o modelo iARMA-restrito com a rede neural MLP<sup>I</sup> para previsão de STI. Adicionalmente, pela simplicidade, recomendamos o uso do modelo iARMA-restrito que mostrou-se ser uma alternativa também viável para previsão de STI.

## 6.2 Trabalhos futuros

A lista de sugestões para trabalhos futuros não é extensa e, alguns casos, nós já ensaiamos estudos preliminares.

1. A primeira sugestão para trabalhos futuro refere-se à implementação do método Holt–Winters para sazonalidade multiplicativa. Este método já está, teoricamente, desenvolvido, faltando sua implementação e teste.

No método Holt–Winters univalorado para variação sazonal multiplicativa, i.e., com amplitude sazonal que tende a crescer ou decrescer ao longo do tempo, as equações de atualização de  $\hat{N}_t$ ,  $\hat{T}_t$  e  $\hat{S}_t$  são dadas por

$$\begin{aligned}\hat{N}_t &= \alpha(y_t/S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}), \\ \hat{T}_t &= \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}, \\ \hat{S}_t &= \delta(y_t/L_t) + (1 - \delta)S_{t-s}.\end{aligned}$$

Em STI, as equações propostas para  $\mathbf{N}_t^I$ ,  $\mathbf{T}_t^I$  e  $\mathbf{S}_t^I$  são dadas por

$$\begin{aligned}\hat{\mathbf{N}}_t^I &= \mathcal{A}(\mathbf{I}_t \odot \hat{\mathbf{S}}_{t-s}^I) + (\mathbf{I} - \mathcal{A})(\hat{\mathbf{N}}_{t-1}^I + \hat{\mathbf{T}}_{t-1}^I), \\ \hat{\mathbf{T}}_t^I &= \mathcal{B}(\hat{\mathbf{N}}_t^I - \hat{\mathbf{N}}_{t-1}^I) + (\mathbf{I} - \mathcal{B})\hat{\mathbf{T}}_{t-1}^I, \\ \hat{\mathbf{S}}_t^I &= \mathcal{D}(\mathbf{I}_t \odot \hat{\mathbf{N}}_t^I) + (\mathbf{I} - \mathcal{D})\hat{\mathbf{S}}_{t-s}^I.\end{aligned}$$

O operador  $\odot$  representa a divisão do  $i$ -ésimo elemento de um vetor pelo  $i$ -ésimo elemento de outro vetor.

2. Imposição de restrições aos parâmetros de outros modelos com a finalidade de considerar a relação de ordem presente em dados do tipo intervalo e, então, assegurar  $\widehat{X}_t^U \geq \widehat{X}_t^L \forall t$ . As restrições podem ser imposta com base na série das amplitudes, de modo a obter  $\widehat{X}_t^r \geq 0$ .
3. Outro ponto considerado para trabalho futuro consiste da inserção de outras técnicas de estimação de parâmetros dos modelos.
4. Também sugerimos o estudo das propriedades de estacionariedade e invertibilidade dos modelos iARMA considerando as restrições nos parâmetros.
5. Explorar outros modelos, tais como, modelos de volatilidade e sistemas dinâmicos para previsão de STI.

# Apêndice A

## Tabelas e gráficos dos experimentos

Apresentamos neste apêndice as tabelas e os gráficos referentes aos resultados experimentais descritos no Capítulo 5.

As Tabelas A.1 a A.18 referem-se aos resultados obtidos na comparação dos modelos no ajuste de STI sintéticas.

Os intervalos de 95% de confiança construídos para os valores médios de  $U^I$  e de  $ARV^I$  nos experimentos com STI sintéticas estão apresentados nas Figuras A.1 a A.18.

As Tabelas A.19 a A.33 exibem os resultados obtidos para os experimentos envolvendo as cinco STI reais.

Os gráficos para os valores das medidas  $U^I$  e  $ARV^I$  na aplicação às STI reais são apresentados nas Figuras A.19 a A.33.

## A.1 Tabelas e gráficos dos experimentos com STI sintéticas

Tabela A.1: Resultados obtidos para o PGD-I com  $n = 220$ ; conjunto de treinamento.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	6,5	0	0,861 (0,029)	0,543 (0,083)	0,859 (0,032)	0,547 (0,085)
MLP <sub>c,r</sub>	9,7	1	0,863 (0,034)	0,551 (0,041)	0,855 (0,029)	0,560 (0,037)
Híbrido <sub>c,r</sub>	5,8	0	0,849 (0,036)	0,521 (0,098)	0,848 (0,039)	0,500 (0,100)
AES <sup>I</sup>	14,6	13	0,924 (0,098)	0,554 (0,118)	0,892 (0,103)	0,549 (0,111)
AED <sup>I</sup>	12,8	7	0,910 (0,107)	0,567 (0,110)	0,891 (0,110)	0,568 (0,110)
HW <sup>I</sup>	14,3	11	0,901 (0,097)	0,563 (0,121)	0,876 (0,102)	0,562 (0,121)
iARMA-irrestrito	8,6	10	0,813 (0,081)	0,500 (0,086)	0,815 (0,078)	0,504 (0,083)
iARMA-restrito	8,8	0	0,809 (0,074)	0,495 (0,082)	0,805 (0,069)	0,496 (0,079)
iARMA-combinado	8,2	0	0,813 (0,073)	0,494 (0,090)	0,815 (0,078)	0,504 (0,083)
iARIMA	10,2	8	0,838 (0,024)	0,509 (0,036)	0,827 (0,023)	0,511 (0,034)
MLP <sup>I</sup>	8,6	4	0,822 (0,034)	0,503 (0,066)	0,826 (0,033)	0,506 (0,064)
Híbrido <sup>I</sup>	9,7	0	0,799 (0,064)	0,466 (0,081)	0,791 (0,067)	0,463 (0,092)
AES <sub>s,i</sub>	10,4	12	0,951 (0,028)	0,572 (0,044)	0,930 (0,027)	0,568 (0,042)
AED <sub>s,i</sub>	9,2	9	0,960 (0,047)	0,603 (0,113)	0,938 (0,049)	0,590 (0,113)
HW <sub>s,i</sub>	10,8	11	0,937 (0,076)	0,618 (0,109)	0,926 (0,073)	0,623 (0,102)
ARMA <sub>s,i</sub>	8,9	9	0,919 (0,032)	0,598 (0,090)	0,912 (0,035)	0,599 (0,092)
ARIMA <sub>s,i</sub>	9,0	9	0,925 (0,031)	0,581 (0,101)	0,923 (0,036)	0,592 (0,104)
MLP <sub>s,i</sub>	8,3	6	0,905 (0,050)	0,570 (0,091)	0,899 (0,053)	0,570 (0,092)
Híbrido <sub>s,i</sub>	8,8	5	0,870 (0,039)	0,525 (0,089)	0,857 (0,040)	0,511 (0,089)

Figura A.1: Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com  $n = 220$ ; conjunto de treinamento.

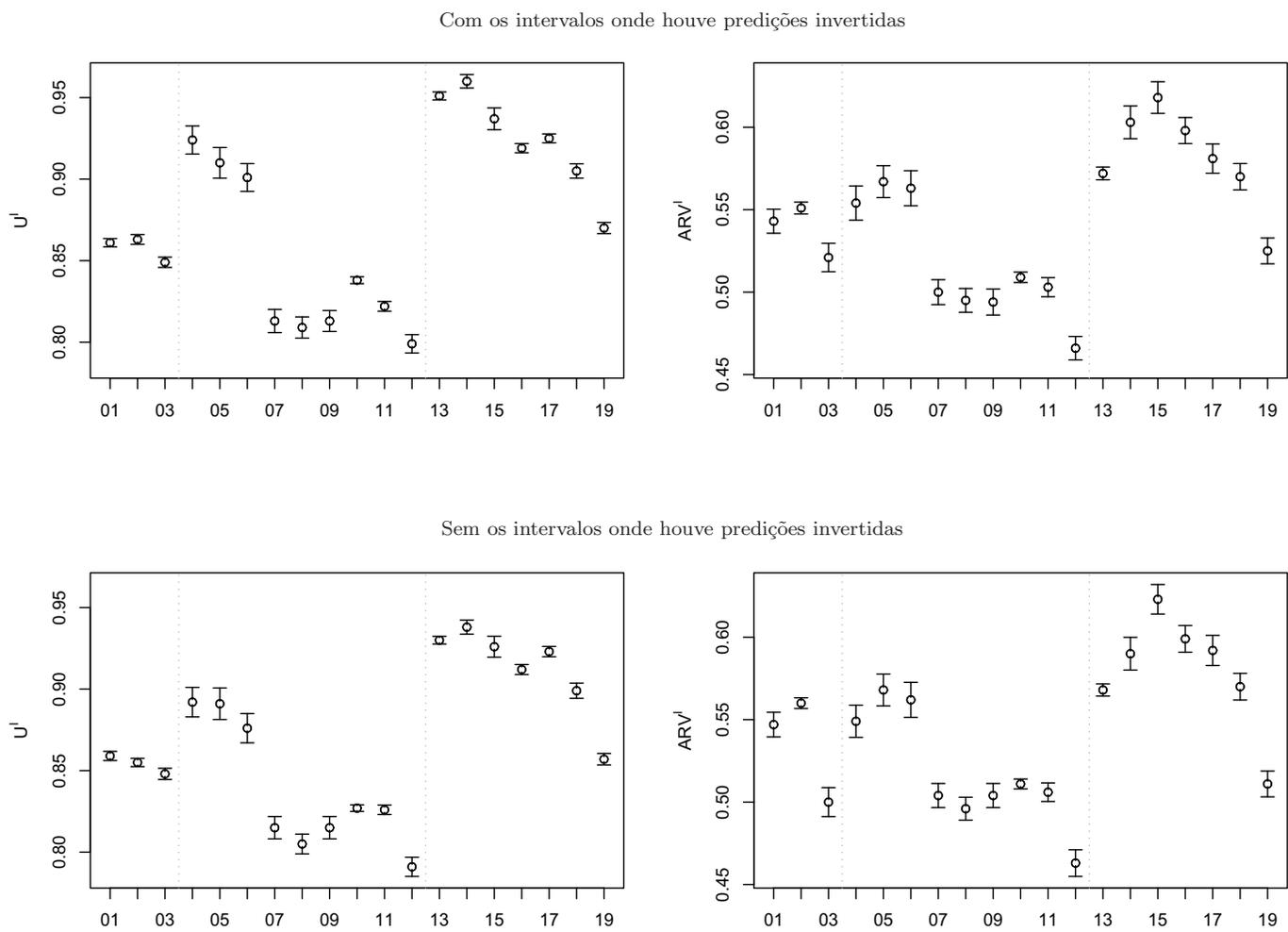


Tabela A.2: Resultados obtidos para o PGD-I com  $n = 220$ ; 5 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,923 (0,031)	0,532 (0,090)	0,923 (0,031)	0,532 (0,090)
MLP <sub>c,r</sub>	–	0	0,911 (0,037)	0,515 (0,048)	0,911 (0,037)	0,515 (0,048)
Híbrido <sub>c,r</sub>	–	0	0,876 (0,040)	0,502 (0,109)	0,876 (0,040)	0,502 (0,109)
AES <sup>I</sup>	–	0	0,972 (0,112)	0,542 (0,143)	0,972 (0,112)	0,542 (0,143)
AED <sup>I</sup>	–	0	0,939 (0,118)	0,565 (0,132)	0,939 (0,118)	0,565 (0,132)
HW <sup>I</sup>	–	0	0,925 (0,108)	0,568 (0,141)	0,925 (0,108)	0,568 (0,141)
iARMA-irrestrito	–	0	0,853 (0,099)	0,444 (0,096)	0,853 (0,099)	0,444 (0,096)
iARMA-restrito	–	0	0,866 (0,087)	0,448 (0,103)	0,866 (0,087)	0,448 (0,103)
iARMA-combinado	–	0	0,853 (0,099)	0,444 (0,096)	0,853 (0,099)	0,444 (0,096)
iARIMA	–	0	0,902 (0,032)	0,460 (0,043)	0,902 (0,032)	0,460 (0,043)
MLP <sup>I</sup>	–	0	0,878 (0,043)	0,458 (0,079)	0,878 (0,043)	0,458 (0,079)
Híbrido <sup>I</sup>	–	0	0,803 (0,083)	0,412 (0,108)	0,803 (0,083)	0,412 (0,108)
AES <sub>s,i</sub>	–	0	0,996 (0,028)	0,577 (0,047)	0,996 (0,028)	0,577 (0,047)
AED <sub>s,i</sub>	–	0	1,064 (0,061)	0,599 (0,119)	1,064 (0,061)	0,599 (0,119)
HW <sub>s,i</sub>	–	0	1,040 (0,078)	0,598 (0,141)	1,040 (0,078)	0,598 (0,141)
ARMA <sub>s,i</sub>	–	0	0,956 (0,032)	0,583 (0,109)	0,956 (0,032)	0,583 (0,109)
ARIMA <sub>s,i</sub>	–	0	1,045 (0,037)	0,587 (0,111)	1,045 (0,037)	0,587 (0,111)
MLP <sub>s,i</sub>	–	0	0,937 (0,051)	0,578 (0,109)	0,937 (0,051)	0,578 (0,109)
Híbrido <sub>s,i</sub>	–	0	0,913 (0,048)	0,510 (0,095)	0,913 (0,048)	0,510 (0,095)

Figura A.2: Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com  $n = 220$ ; 5 passos-à-frente.

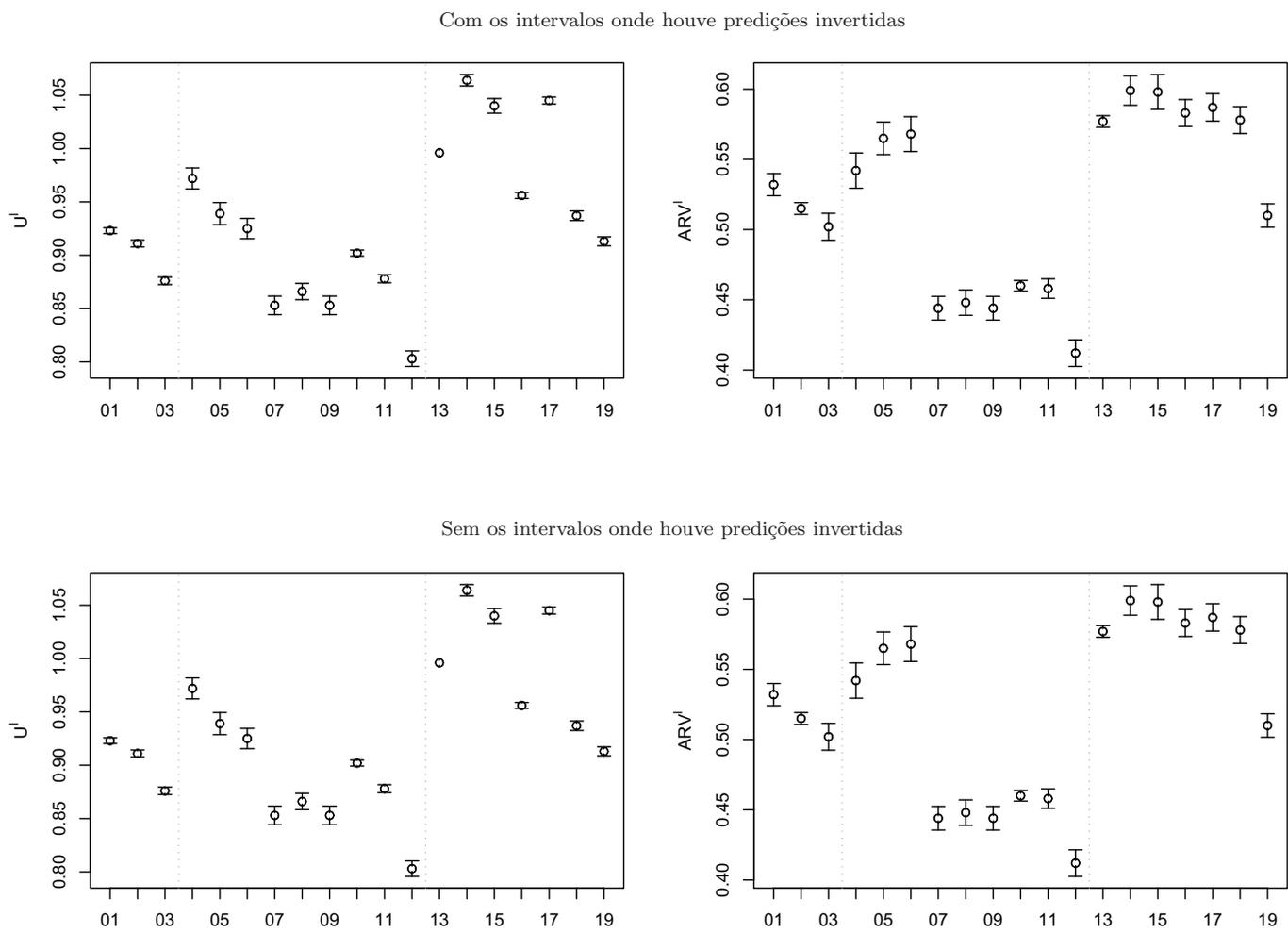


Tabela A.3: Resultados obtidos para o PGD-I com  $n = 220$ ; 20 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,013 (0,069)	0,527 (0,101)	0,998 (0,069)	0,525 (0,113)
MLP <sub>c,r</sub>	–	0	1,029 (0,070)	0,527 (0,045)	1,031 (0,071)	0,520 (0,033)
Híbrido <sub>c,r</sub>	–	0	0,998 (0,091)	0,513 (0,106)	0,996 (0,090)	0,515 (0,123)
AES <sup>I</sup>	–	0	1,062 (0,138)	0,531 (0,159)	1,057 (0,137)	0,533 (0,161)
AED <sup>I</sup>	–	1	1,026 (0,129)	0,551 (0,146)	1,030 (0,130)	0,554 (0,159)
HW <sup>I</sup>	–	4	1,018 (0,167)	0,549 (0,151)	1,017 (0,163)	0,550 (0,170)
iARMA-irrestrito	–	1	0,972 (0,109)	0,438 (0,093)	0,980 (0,110)	0,426 (0,081)
iARMA-restrito	–	0	0,968 (0,123)	0,439 (0,089)	0,972 (0,123)	0,431 (0,085)
iARMA-combinado	–	0	0,973 (0,118)	0,435 (0,095)	0,980 (0,110)	0,426 (0,081)
iARIMA	–	3	1,026 (0,067)	0,460 (0,044)	1,022 (0,067)	0,463 (0,050)
MLP <sup>I</sup>	–	1	0,992 (0,067)	0,462 (0,083)	0,993 (0,066)	0,464 (0,085)
Híbrido <sup>I</sup>	–	0	0,905 (0,117)	0,410 (0,099)	0,903 (0,112)	0,413 (0,091)
AES <sub>s,i</sub>	–	1	1,137 (0,043)	0,592 (0,051)	1,139 (0,043)	0,593 (0,052)
AED <sub>s,i</sub>	–	1	1,143 (0,074)	0,592 (0,137)	1,147 (0,074)	0,592 (0,138)
HW <sub>s,i</sub>	–	3	1,103 (0,121)	0,601 (0,135)	1,160 (0,120)	0,600 (0,134)
ARMA <sub>s,i</sub>	–	0	1,017 (0,043)	0,586 (0,104)	1,020 (0,043)	0,593 (0,135)
ARIMA <sub>s,i</sub>	–	0	1,093 (0,041)	0,573 (0,122)	1,089 (0,041)	0,569 (0,145)
MLP <sub>s,i</sub>	–	1	1,032 (0,086)	0,580 (0,121)	1,029 (0,086)	0,583 (0,137)
Híbrido <sub>s,i</sub>	–	0	1,003 (0,062)	0,536 (0,105)	1,007 (0,062)	0,535 (0,118)

Figura A.3: Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com  $n = 220$ ; 20 passos-à-frente.

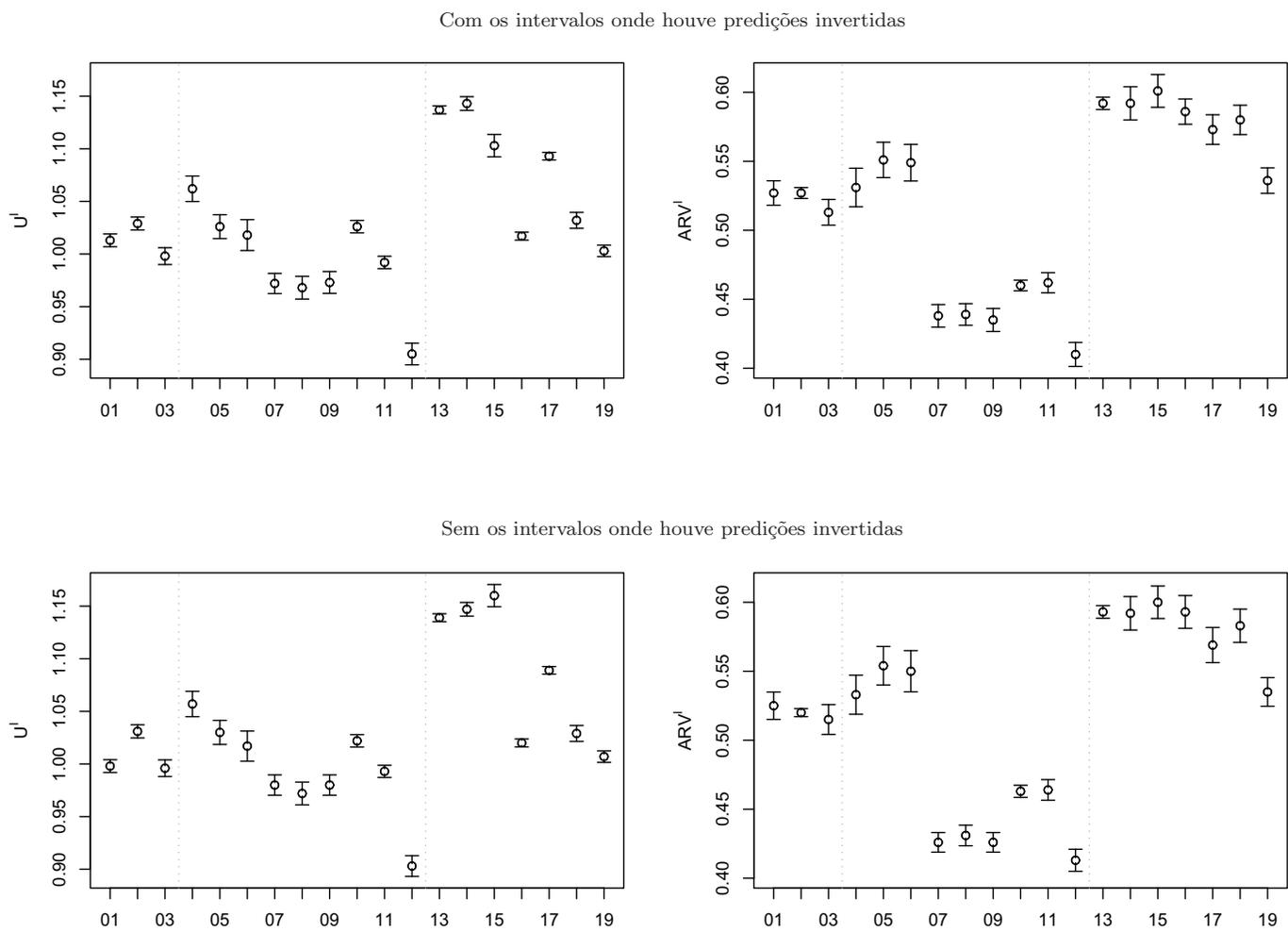


Tabela A.4: Resultados obtidos para o PGD-I com  $n = 620$ ; conjunto de treinamento.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$	6,5	0	0,913 (0,023)	0,457 (0,060)	0,916 (0,020)	0,452 (0,062)
$MLP_{c,r}$	8,4	2	0,903 (0,038)	0,461 (0,041)	0,899 (0,044)	0,468 (0,039)
Híbrido $_{c,r}$	5,0	0	0,888 (0,046)	0,447 (0,092)	0,893 (0,061)	0,442 (0,089)
$AES^I$	14,3	26	0,938 (0,092)	0,483 (0,101)	0,921 (0,089)	0,481 (0,103)
$AED^I$	11,6	17	0,914 (0,110)	0,479 (0,103)	0,917 (0,121)	0,481 (0,100)
$HW^I$	13,3	21	0,908 (0,098)	0,478 (0,112)	0,899 (0,098)	0,478 (0,112)
iARMA-irrestrito	8,4	12	0,840 (0,073)	0,449 (0,102)	0,834 (0,070)	0,440 (0,104)
iARMA-restrito	8,9	0	0,834 (0,071)	0,451 (0,105)	0,828 (0,072)	0,449 (0,105)
iARMA-combinado	7,6	0	0,836 (0,076)	0,443 (0,116)	0,834 (0,070)	0,440 (0,104)
iARIMA	9,2	9	0,847 (0,040)	0,443 (0,060)	0,852 (0,047)	0,439 (0,067)
$MLP^I$	8,2	7	0,842 (0,053)	0,424 (0,053)	0,840 (0,044)	0,419 (0,053)
Híbrido $^I$	9,1	0	0,809 (0,067)	0,415 (0,133)	0,807 (0,075)	0,412 (0,131)
$AES_{s,i}$	9,4	17	0,952 (0,045)	0,499 (0,061)	0,937 (0,054)	0,501 (0,062)
$AED_{s,i}$	8,6	19	0,964 (0,057)	0,498 (0,100)	0,960 (0,060)	0,495 (0,098)
$HW_{s,i}$	9,6	12	0,956 (0,090)	0,496 (0,122)	0,948 (0,114)	0,497 (0,116)
$ARMA_{s,i}$	9,3	11	0,950 (0,036)	0,497 (0,084)	0,952 (0,043)	0,507 (0,082)
$ARIMA_{s,i}$	8,7	14	0,959 (0,031)	0,492 (0,088)	0,955 (0,032)	0,496 (0,088)
$MLP_{s,i}$	7,2	9	0,925 (0,060)	0,479 (0,079)	0,918 (0,076)	0,489 (0,075)
Híbrido $_{s,i}$	8,0	6	0,876 (0,058)	0,453 (0,087)	0,871 (0,062)	0,450 (0,084)

Figura A.4: Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com  $n = 620$ ; conjunto de treinamento.

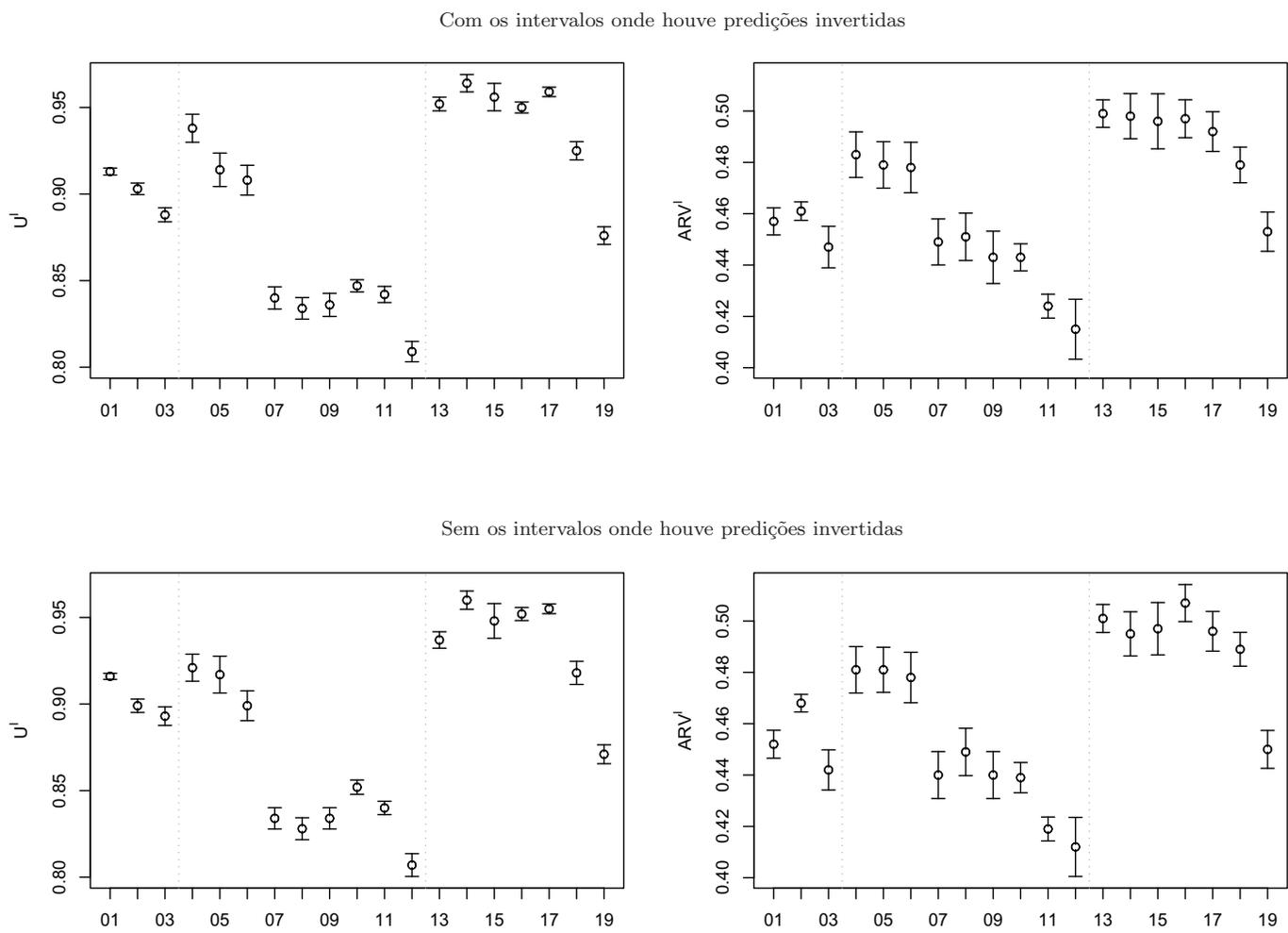


Tabela A.5: Resultados obtidos para o PGD-I com  $n = 620$ ; 5 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,903 (0,030)	0,460 (0,075)	0,904 (0,031)	0,462 (0,077)
MLP <sub>c,r</sub>	–	0	0,921 (0,048)	0,463 (0,053)	0,922 (0,050)	0,466 (0,055)
Híbrido <sub>c,r</sub>	–	0	0,911 (0,048)	0,439 (0,105)	0,911 (0,048)	0,435 (0,106)
AES <sup>I</sup>	–	1	0,957 (0,115)	0,480 (0,132)	0,961 (0,123)	0,480 (0,139)
AED <sup>I</sup>	–	0	0,937 (0,120)	0,481 (0,122)	0,938 (0,121)	0,483 (0,123)
HW <sup>I</sup>	–	0	0,940 (0,121)	0,459 (0,130)	0,941 (0,124)	0,460 (0,133)
iARMA-irrestrito	–	0	0,838 (0,084)	0,427 (0,146)	0,840 (0,088)	0,429 (0,150)
iARMA-restrito	–	0	0,833 (0,081)	0,430 (0,111)	0,836 (0,088)	0,430 (0,117)
iARMA-combinado	–	0	0,838 (0,084)	0,427 (0,146)	0,840 (0,088)	0,429 (0,150)
iARIMA	–	1	0,862 (0,053)	0,439 (0,025)	0,869 (0,060)	0,439 (0,033)
MLP <sup>I</sup>	–	0	0,852 (0,048)	0,421 (0,057)	0,852 (0,050)	0,422 (0,063)
Híbrido <sup>I</sup>	–	0	0,826 (0,089)	0,420 (0,105)	0,828 (0,091)	0,420 (0,100)
AES <sub>s,i</sub>	–	0	1,044 (0,045)	0,498 (0,035)	1,046 (0,040)	0,499 (0,039)
AED <sub>s,i</sub>	–	0	1,043 (0,067)	0,499 (0,122)	1,043 (0,069)	0,501 (0,124)
HW <sub>s,i</sub>	–	0	1,076 (0,105)	0,488 (0,125)	1,077 (0,108)	0,490 (0,130)
ARMA <sub>s,i</sub>	–	0	1,005 (0,043)	0,492 (0,096)	1,006 (0,044)	0,494 (0,098)
ARIMA <sub>s,i</sub>	–	0	0,999 (0,041)	0,478 (0,097)	0,998 (0,045)	0,482 (0,101)
MLP <sub>s,i</sub>	–	0	1,001 (0,075)	0,462 (0,086)	1,009 (0,081)	0,469 (0,092)
Híbrido <sub>s,i</sub>	–	0	0,967 (0,061)	0,432 (0,105)	0,970 (0,066)	0,434 (0,110)

Figura A.5: Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com  $n = 620$ ; 5 passos-à-frente.

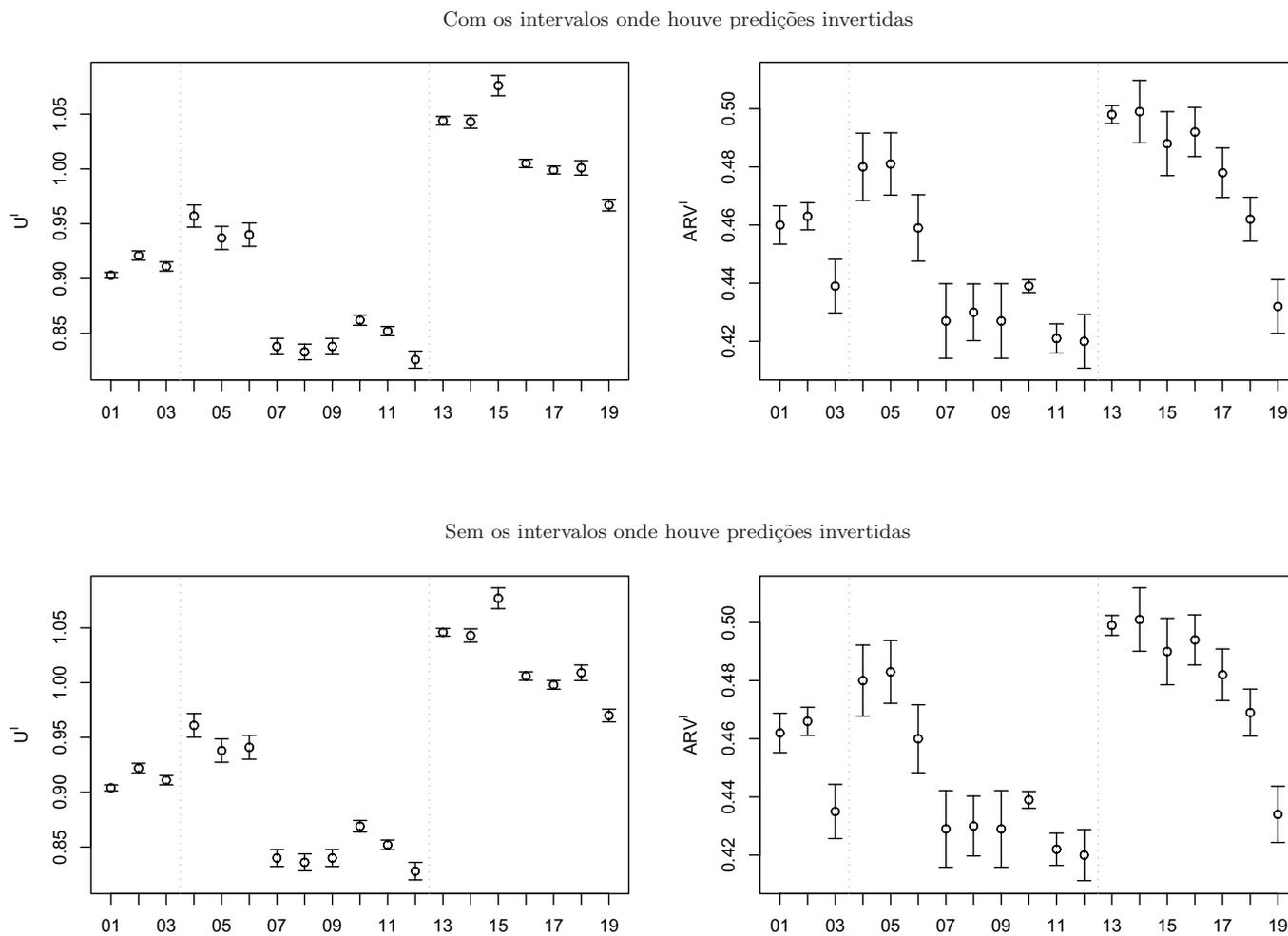


Tabela A.6: Resultados obtidos para o PGD-I com  $n = 620$ ; 20 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,122 (0,075)	0,463 (0,064)	1,113 (0,077)	0,472 (0,085)
MLP <sub>c,r</sub>	–	0	1,103 (0,085)	0,459 (0,044)	1,110 (0,085)	0,473 (0,046)
Híbrido <sub>c,r</sub>	–	0	1,039 (0,074)	0,439 (0,128)	1,045 (0,074)	0,441 (0,132)
AES <sup>I</sup>	–	4	1,131 (0,116)	0,482 (0,111)	1,129 (0,115)	0,483 (0,170)
AED <sup>I</sup>	–	2	1,073 (0,090)	0,473 (0,123)	1,068 (0,099)	0,475 (0,131)
HW <sup>I</sup>	–	3	1,043 (0,150)	0,502 (0,140)	1,059 (0,152)	0,509 (0,160)
iARMA-irrestrito	–	3	1,006 (0,130)	0,436 (0,103)	1,005 (0,089)	0,440 (0,109)
iARMA-restrito	–	0	1,008 (0,112)	0,446 (0,112)	1,009 (0,113)	0,456 (0,108)
iARMA-combinado	–	0	1,002 (0,093)	0,426 (0,188)	1,005 (0,089)	0,440 (0,109)
iARIMA	–	3	1,011 (0,069)	0,435 (0,073)	1,016 (0,070)	0,437 (0,071)
MLP <sup>I</sup>	–	2	0,991 (0,057)	0,406 (0,071)	0,989 (0,057)	0,410 (0,108)
Híbrido <sup>I</sup>	–	0	0,983 (0,108)	0,411 (0,113)	0,981 (0,102)	0,407 (0,114)
AES <sub>s,i</sub>	–	2	1,211 (0,045)	0,400 (0,039)	1,208 (0,044)	0,402 (0,063)
AED <sub>s,i</sub>	–	2	1,201 (0,105)	0,500 (0,117)	1,204 (0,105)	0,505 (0,127)
HW <sub>s,i</sub>	–	4	1,212 (0,116)	0,513 (0,093)	1,218 (0,104)	0,518 (0,091)
ARMA <sub>s,i</sub>	–	2	1,207 (0,051)	0,489 (0,106)	1,209 (0,051)	0,497 (0,128)
ARIMA <sub>s,i</sub>	–	0	1,134 (0,073)	0,483 (0,095)	1,159 (0,072)	0,490 (0,137)
MLP <sub>s,i</sub>	–	2	1,155 (0,090)	0,455 (0,088)	1,153 (0,091)	0,459 (0,142)
Híbrido <sub>s,i</sub>	–	0	1,082 (0,062)	0,432 (0,101)	1,083 (0,062)	0,429 (0,139)

Figura A.6: Intervalos de 95% de confiança para os resultados obtidos para o PGD-I com  $n = 620$ ; 20 passos-à-frente.

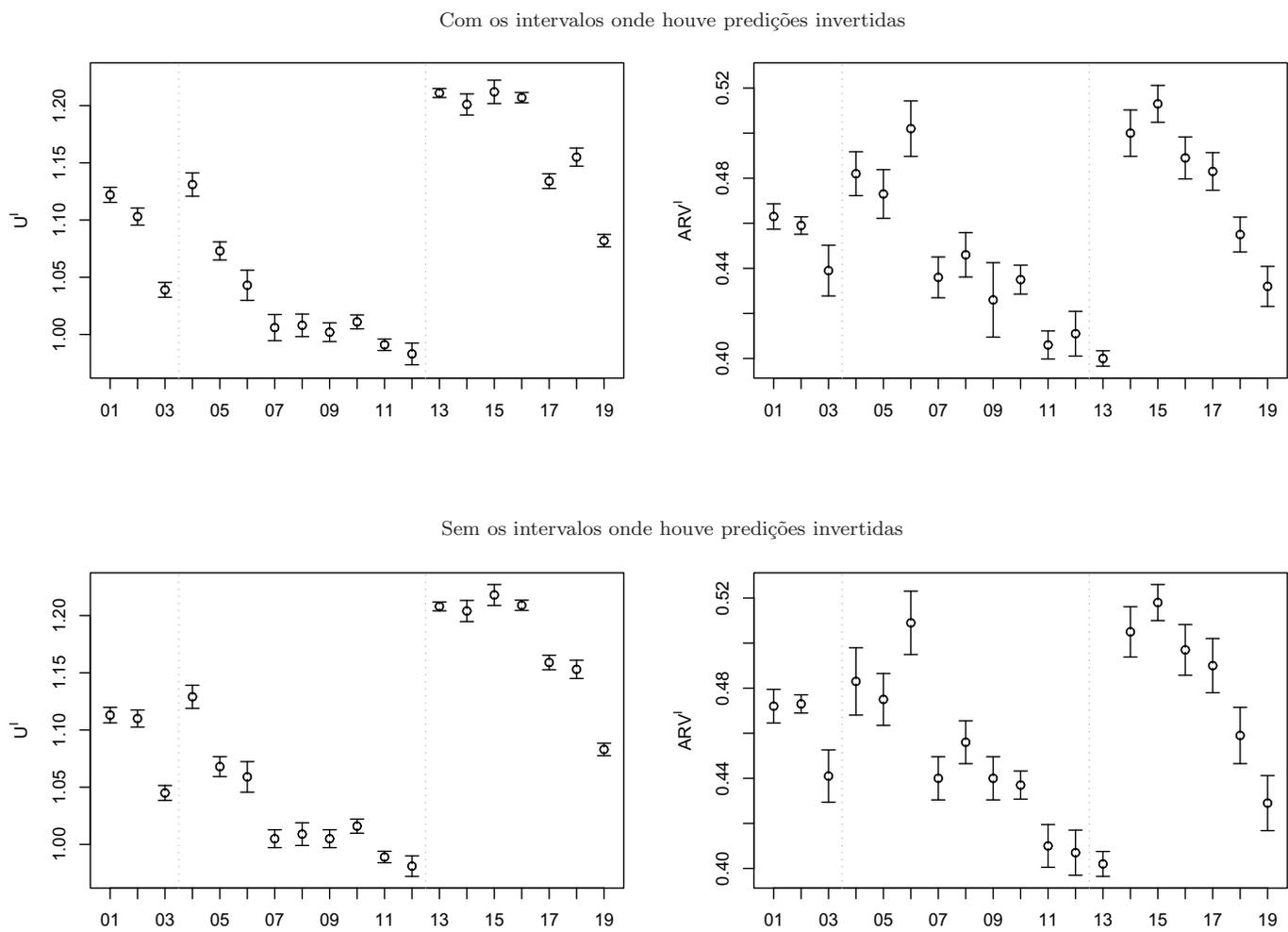


Tabela A.7: Resultados obtidos para o PGD-II com  $n = 220$ ; conjunto de treinamento.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	15,1	0	0,921 (0,043)	0,610 (0,091)	0,926 (0,046)	0,607 (0,093)
MLP <sub>c,r</sub>	12,3	0	0,951 (0,066)	0,600 (0,085)	0,928 (0,068)	0,593 (0,085)
Híbrido <sub>c,r</sub>	14,3	0	0,921 (0,053)	0,578 (0,044)	0,917 (0,048)	0,575 (0,041)
AES <sup>I</sup>	9,2	9	0,925 (0,124)	0,589 (0,124)	0,927 (0,132)	0,579 (0,125)
AED <sup>I</sup>	7,8	9	0,901 (0,097)	0,577 (0,121)	0,899 (0,104)	0,572 (0,122)
HW <sup>I</sup>	10,6	11	0,913 (0,171)	0,585 (0,139)	0,913 (0,182)	0,569 (0,139)
iARMA-irrestrito	9,1	8	0,819 (0,107)	0,474 (0,078)	0,808 (0,109)	0,477 (0,077)
iARMA-restrito	9,7	0	0,815 (0,061)	0,476 (0,075)	0,813 (0,060)	0,480 (0,074)
iARMA-combinado	8,8	0	0,818 (0,071)	0,471 (0,080)	0,808 (0,109)	0,477 (0,077)
iARIMA	9,1	7	0,848 (0,041)	0,492 (0,081)	0,833 (0,046)	0,498 (0,083)
MLP <sup>I</sup>	7,8	3	0,822 (0,046)	0,511 (0,098)	0,819 (0,051)	0,502 (0,100)
Híbrido <sup>I</sup>	9,3	0	0,799 (0,078)	0,454 (0,139)	0,779 (0,077)	0,452 (0,138)
AES <sub>s,i</sub>	8,7	7	0,930 (0,045)	0,600 (0,075)	0,931 (0,050)	0,609 (0,076)
AED <sub>s,i</sub>	8,6	6	0,916 (0,047)	0,612 (0,100)	0,920 (0,047)	0,611 (0,098)
HW <sub>s,i</sub>	9,1	2	0,925 (0,106)	0,603 (0,192)	0,934 (0,112)	0,591 (0,193)
ARMA <sub>s,i</sub>	9,9	2	0,874 (0,049)	0,586 (0,073)	0,851 (0,050)	0,566 (0,074)
ARIMA <sub>s,i</sub>	8,9	0	0,888 (0,057)	0,556 (0,078)	0,867 (0,059)	0,547 (0,078)
MLP <sub>s,i</sub>	5,9	6	0,868 (0,062)	0,570 (0,123)	0,863 (0,066)	0,569 (0,124)
Híbrido <sub>s,i</sub>	6,5	1	0,860 (0,042)	0,525 (0,059)	0,822 (0,037)	0,530 (0,056)

Figura A.7: Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com  $n = 220$ ; conjunto de treinamento.

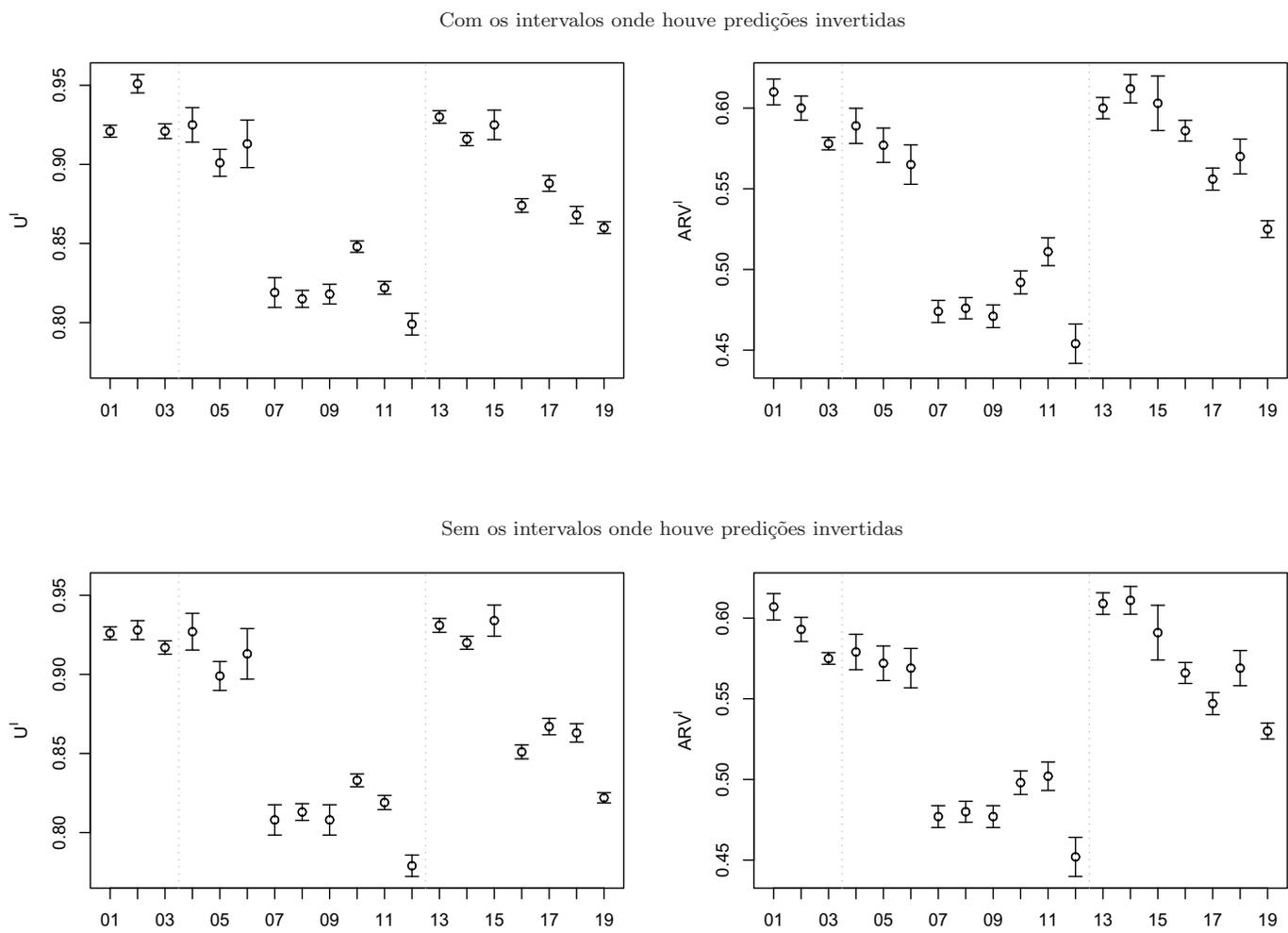


Tabela A.8: Resultados obtidos para o PGD-II com  $n = 220$ ; 5 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,986 (0,055)	0,535 (0,117)	0,986 (0,055)	0,535 (0,117)
MLP <sub>c,r</sub>	–	0	0,991 (0,078)	0,547 (0,102)	0,991 (0,078)	0,547 (0,102)
Híbrido <sub>c,r</sub>	–	0	0,981 (0,056)	0,515 (0,046)	0,981 (0,056)	0,515 (0,046)
AES <sup>I</sup>	–	0	1,007 (0,158)	0,542 (0,140)	1,007 (0,158)	0,542 (0,140)
AED <sup>I</sup>	–	0	0,968 (0,120)	0,538 (0,135)	0,968 (0,120)	0,538 (0,135)
HW <sup>I</sup>	–	0	1,013 (0,174)	0,521 (0,179)	1,013 (0,174)	0,521 (0,179)
iARMA-irrestrito	–	0	0,862 (0,112)	0,481 (0,098)	0,862 (0,112)	0,481 (0,098)
iARMA-restrito	–	0	0,867 (0,079)	0,483 (0,085)	0,867 (0,079)	0,483 (0,085)
iARMA-combinado	–	0	0,865 (0,112)	0,481 (0,098)	0,865 (0,112)	0,481 (0,098)
iARIMA	–	0	0,880 (0,079)	0,492 (0,089)	0,880 (0,079)	0,492 (0,085)
MLP <sup>I</sup>	–	0	0,847 (0,055)	0,499 (0,122)	0,847 (0,055)	0,499 (0,122)
Híbrido <sup>I</sup>	–	0	0,806 (0,082)	0,459 (0,162)	0,806 (0,082)	0,459 (0,162)
AES <sub>s,i</sub>	–	0	1,021 (0,055)	0,563 (0,095)	1,021 (0,055)	0,563 (0,095)
AED <sub>s,i</sub>	–	0	1,004 (0,051)	0,573 (0,106)	1,004 (0,051)	0,573 (0,106)
HW <sub>s,i</sub>	–	0	1,027 (0,126)	0,568 (0,214)	1,027 (0,126)	0,568 (0,214)
ARMA <sub>s,i</sub>	–	0	0,998 (0,055)	0,545 (0,089)	0,998 (0,055)	0,545 (0,089)
ARIMA <sub>s,i</sub>	–	0	0,998 (0,070)	0,547 (0,085)	0,998 (0,070)	0,547 (0,085)
MLP <sub>s,i</sub>	–	0	1,022 (0,077)	0,539 (0,129)	1,022 (0,077)	0,539 (0,129)
Híbrido <sub>s,i</sub>	–	0	0,973 (0,043)	0,516 (0,074)	0,973 (0,043)	0,516 (0,074)

Figura A.8: Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com  $n = 220$ ; 5 passos-à-frente.

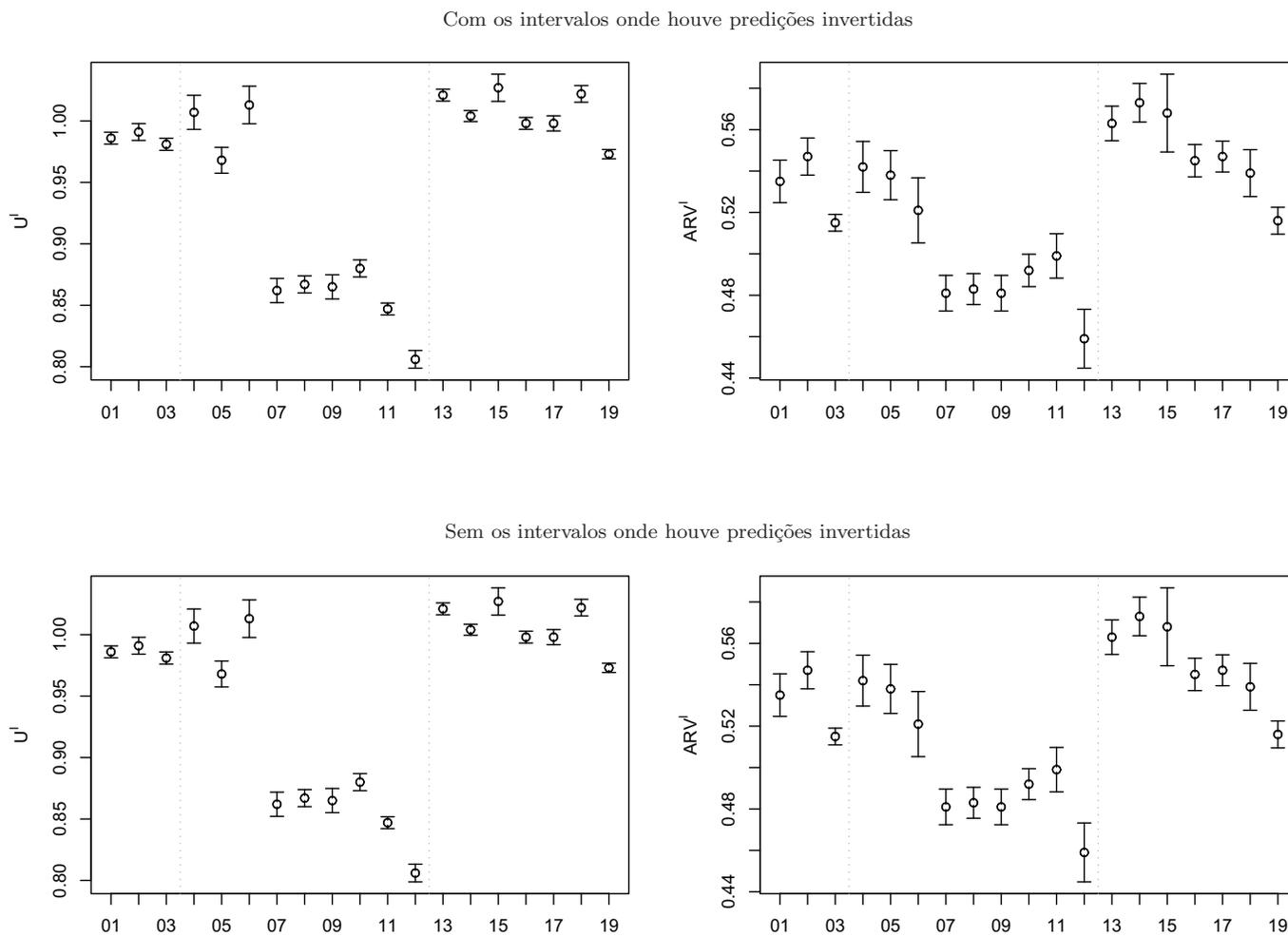


Tabela A.9: Resultados obtidos para o PGD-II com  $n = 220$ ; 20 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,153 (0,070)	0,630 (0,104)	1,153 (0,070)	0,638 (0,133)
MLP <sub>c,r</sub>	–	0	1,139 (0,104)	0,613 (0,104)	1,139 (0,104)	0,617 (0,110)
Híbrido <sub>c,r</sub>	–	0	1,063 (0,057)	0,561 (0,050)	1,064 (0,057)	0,559 (0,049)
AES <sup>I</sup>	–	1	1,110 (0,171)	0,585 (0,119)	1,112 (0,171)	0,590 (0,136)
AED <sup>I</sup>	–	2	1,082 (0,136)	0,567 (0,128)	1,081 (0,128)	0,568 (0,151)
HW <sup>I</sup>	–	1	1,099 (0,123)	0,558 (0,128)	1,098 (0,123)	0,546 (0,140)
iARMA-irrestrito	–	2	0,989 (0,147)	0,487 (0,134)	0,981 (0,146)	0,490 (0,136)
iARMA-restrito	–	0	0,988 (0,091)	0,483 (0,100)	0,983 (0,091)	0,489 (0,099)
iARMA-combinado	–	0	0,988 (0,115)	0,476 (0,105)	0,981 (0,146)	0,490 (0,136)
iARIMA	–	2	1,012 (0,089)	0,518 (0,111)	1,009 (0,087)	0,509 (0,128)
MLP <sup>I</sup>	–	0	1,031 (0,072)	0,489 (0,135)	1,023 (0,071)	0,493 (0,153)
Híbrido <sup>I</sup>	–	0	0,942 (0,089)	0,492 (0,192)	0,941 (0,088)	0,497 (0,191)
AES <sub>s,i</sub>	–	1	1,103 (0,069)	0,613 (0,103)	1,100 (0,069)	0,619 (0,124)
AED <sub>s,i</sub>	–	1	1,128 (0,061)	0,601 (0,137)	1,132 (0,061)	0,604 (0,136)
HW <sub>s,i</sub>	–	2	1,156 (0,142)	0,613 (0,103)	1,159 (0,143)	0,610 (0,110)
ARMA <sub>s,i</sub>	–	0	1,117 (0,090)	0,561 (0,094)	1,114 (0,088)	0,565 (0,102)
ARIMA <sub>s,i</sub>	–	1	1,115 (0,126)	0,547 (0,084)	1,112 (0,126)	0,554 (0,097)
MLP <sub>s,i</sub>	–	1	1,118 (0,109)	0,569 (0,133)	1,124 (0,110)	0,586 (0,160)
Híbrido <sub>s,i</sub>	–	0	1,089 (0,062)	0,517 (0,074)	1,094 (0,063)	0,521 (0,067)

Figura A.9: Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com  $n = 220$ ; 20 passos-à-frente.

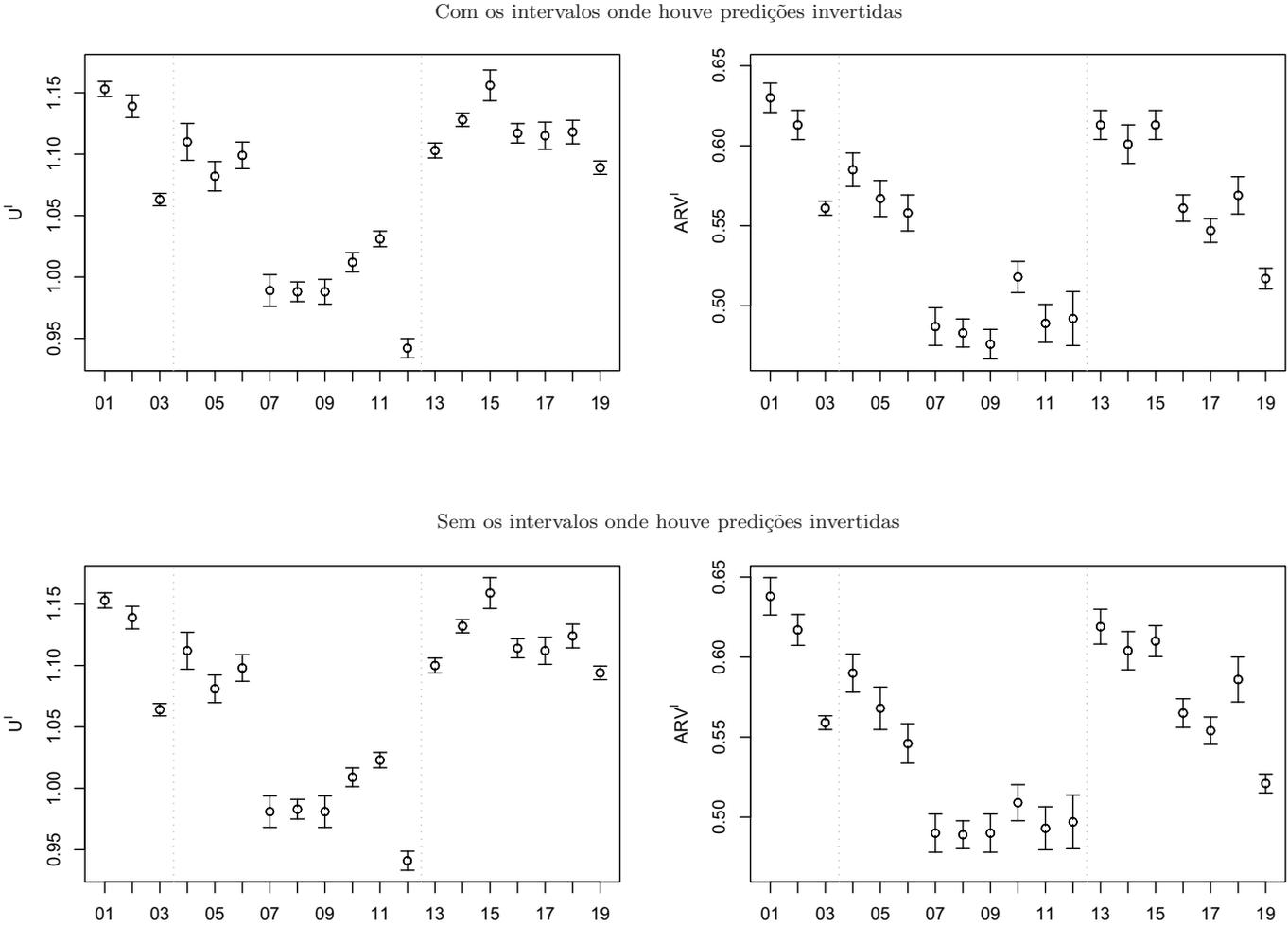


Tabela A.10: Resultados obtidos para o PGD-II com  $n = 620$ ; conjunto de treinamento.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$	14,6	0	0,979 (0,033)	0,522 (0,085)	0,982 (0,042)	0,529 (0,083)
$MLP_{c,r}$	12,3	4	0,984 (0,026)	0,536 (0,058)	0,990 (0,023)	0,541 (0,060)
Híbrido $_{c,r}$	13,5	1	0,923 (0,034)	0,495 (0,029)	0,929 (0,035)	0,500 (0,029)
$AES^I$	9,0	16	0,930 (0,111)	0,509 (0,122)	0,948 (0,106)	0,512 (0,126)
$AED^I$	7,1	21	0,918 (0,094)	0,508 (0,140)	0,915 (0,094)	0,502 (0,140)
$HW^I$	10,4	10	0,932 (0,129)	0,506 (0,133)	0,920 (0,128)	0,506 (0,134)
iARMA-irrestrito	8,6	15	0,847 (0,075)	0,470 (0,073)	0,838 (0,074)	0,468 (0,073)
iARMA-restrito	8,7	0	0,843 (0,037)	0,468 (0,064)	0,837 (0,042)	0,465 (0,063)
iARMA-combinado	7,8	0	0,840 (0,069)	0,463 (0,099)	0,838 (0,074)	0,466 (0,073)
iARIMA	9,3	18	0,850 (0,037)	0,470 (0,044)	0,846 (0,042)	0,470 (0,063)
$MLP^I$	7,8	4	0,807 (0,034)	0,473 (0,066)	0,817 (0,039)	0,470 (0,069)
Híbrido $^I$	9,4	0	0,800 (0,053)	0,442 (0,102)	0,803 (0,049)	0,450 (0,105)
$AES_{s,i}$	7,5	14	0,968 (0,031)	0,539 (0,069)	0,950 (0,036)	0,532 (0,068)
$AED_{s,i}$	8,0	11	0,930 (0,049)	0,507 (0,079)	0,925 (0,052)	0,518 (0,078)
$HW_{s,i}$	9,1	13	0,936 (0,062)	0,512 (0,131)	0,933 (0,058)	0,516 (0,134)
$ARMA_{s,i}$	9,5	7	0,905 (0,034)	0,499 (0,059)	0,904 (0,039)	0,489 (0,058)
$ARIMA_{s,i}$	7,8	6	0,921 (0,029)	0,518 (0,047)	0,914 (0,024)	0,524 (0,050)
$MLP_{s,i}$	8,6	3	0,915 (0,046)	0,478 (0,097)	0,915 (0,044)	0,486 (0,098)
Híbrido $_{s,i}$	7,7	2	0,889 (0,035)	0,467 (0,047)	0,880 (0,034)	0,470 (0,048)

Figura A.10: Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com  $n = 620$ ; conjunto de treinamento.

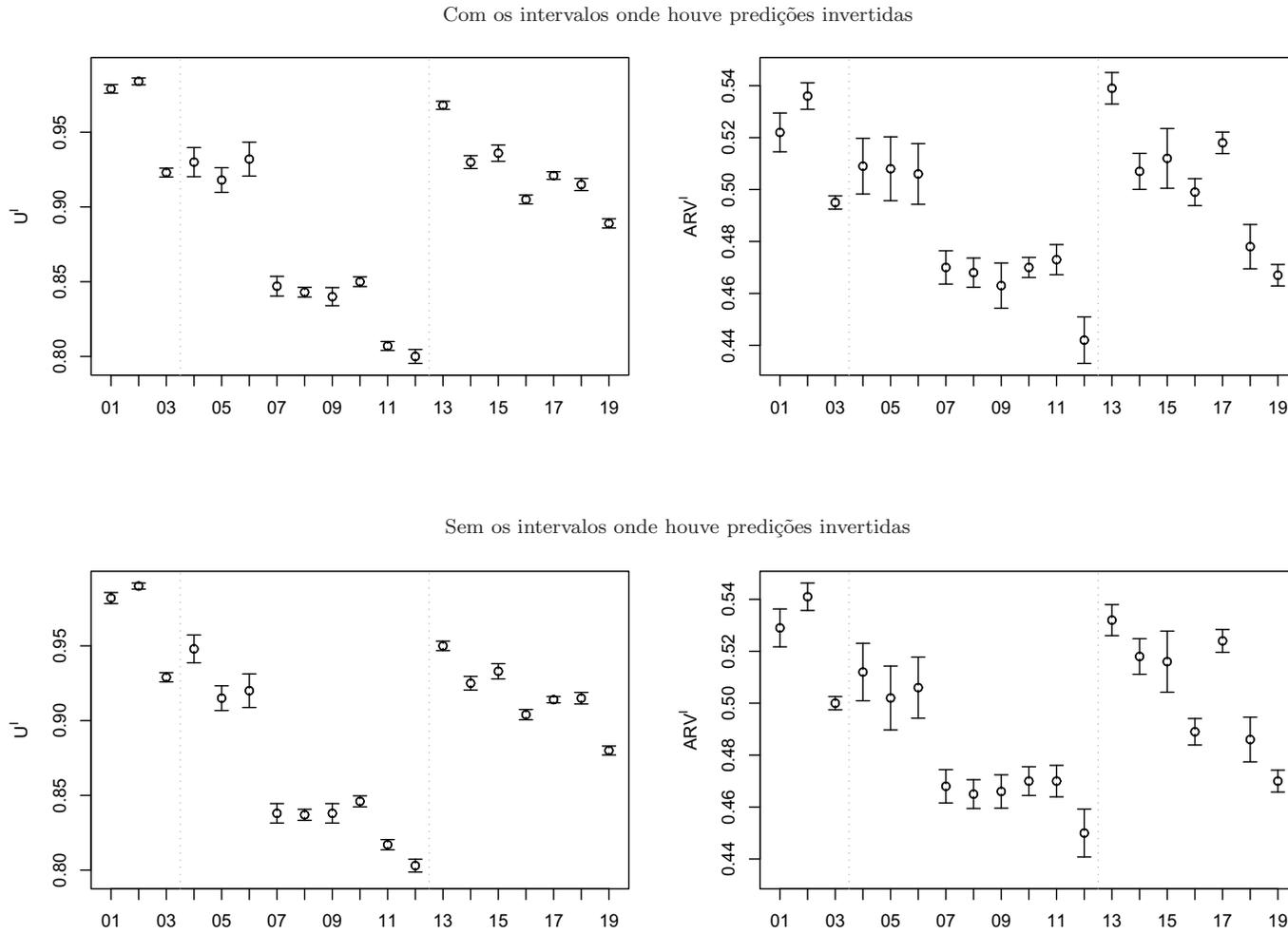


Tabela A.11: Resultados obtidos para o PGD-II com  $n = 620$ ; 5 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$	–	0	1,004 (0,042)	0,516 (0,093)	1,006 (0,045)	0,519 (0,096)
$MLP_{c,r}$	–	0	1,018 (0,026)	0,543 (0,066)	1,018 (0,028)	0,545 (0,068)
Híbrido $_{c,r}$	–	0	1,012 (0,039)	0,482 (0,035)	1,014 (0,041)	0,484 (0,037)
$AES^I$	–	0	0,993 (0,123)	0,499 (0,135)	0,987 (0,124)	0,499 (0,135)
$AED^I$	–	0	1,003 (0,112)	0,489 (0,166)	1,004 (0,113)	0,490 (0,167)
$HW^I$	–	0	0,975 (0,141)	0,514 (0,156)	0,976 (0,143)	0,515 (0,157)
iARMA-irrestrito	–	0	0,855 (0,090)	0,438 (0,083)	0,857 (0,092)	0,439 (0,085)
iARMA-restrito	–	0	0,862 (0,079)	0,436 (0,071)	0,862 (0,079)	0,435 (0,072)
iARMA-combinado	–	0	0,855 (0,090)	0,438 (0,083)	0,857 (0,092)	0,439 (0,085)
iARIMA	–	0	0,873 (0,041)	0,468 (0,047)	0,879 (0,046)	0,469 (0,048)
$MLP^I$	–	0	0,869 (0,036)	0,466 (0,070)	0,859 (0,037)	0,467 (0,071)
Híbrido $^I$	–	0	0,857 (0,060)	0,439 (0,111)	0,861 (0,063)	0,441 (0,113)
$AES_{s,i}$	–	1	0,987 (0,035)	0,516 (0,080)	0,990 (0,038)	0,519 (0,083)
$AED_{s,i}$	–	0	1,015 (0,065)	0,505 (0,091)	1,015 (0,066)	0,505 (0,092)
$HW_{s,i}$	–	0	1,041 (0,068)	0,530 (0,150)	1,041 (0,068)	0,530 (0,150)
$ARMA_{s,i}$	–	0	0,994 (0,035)	0,514 (0,065)	0,996 (0,037)	0,511 (0,067)
$ARIMA_{s,i}$	–	0	0,986 (0,032)	0,517 (0,058)	0,987 (0,033)	0,518 (0,060)
$MLP_{s,i}$	–	0	1,002 (0,053)	0,477 (0,109)	1,004 (0,056)	0,479 (0,111)
Híbrido $_{s,i}$	–	0	0,991 (0,043)	0,490 (0,061)	0,992 (0,044)	0,491 (0,063)

Figura A.11: Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com  $n = 620$ ; 5 passos-à-frente.

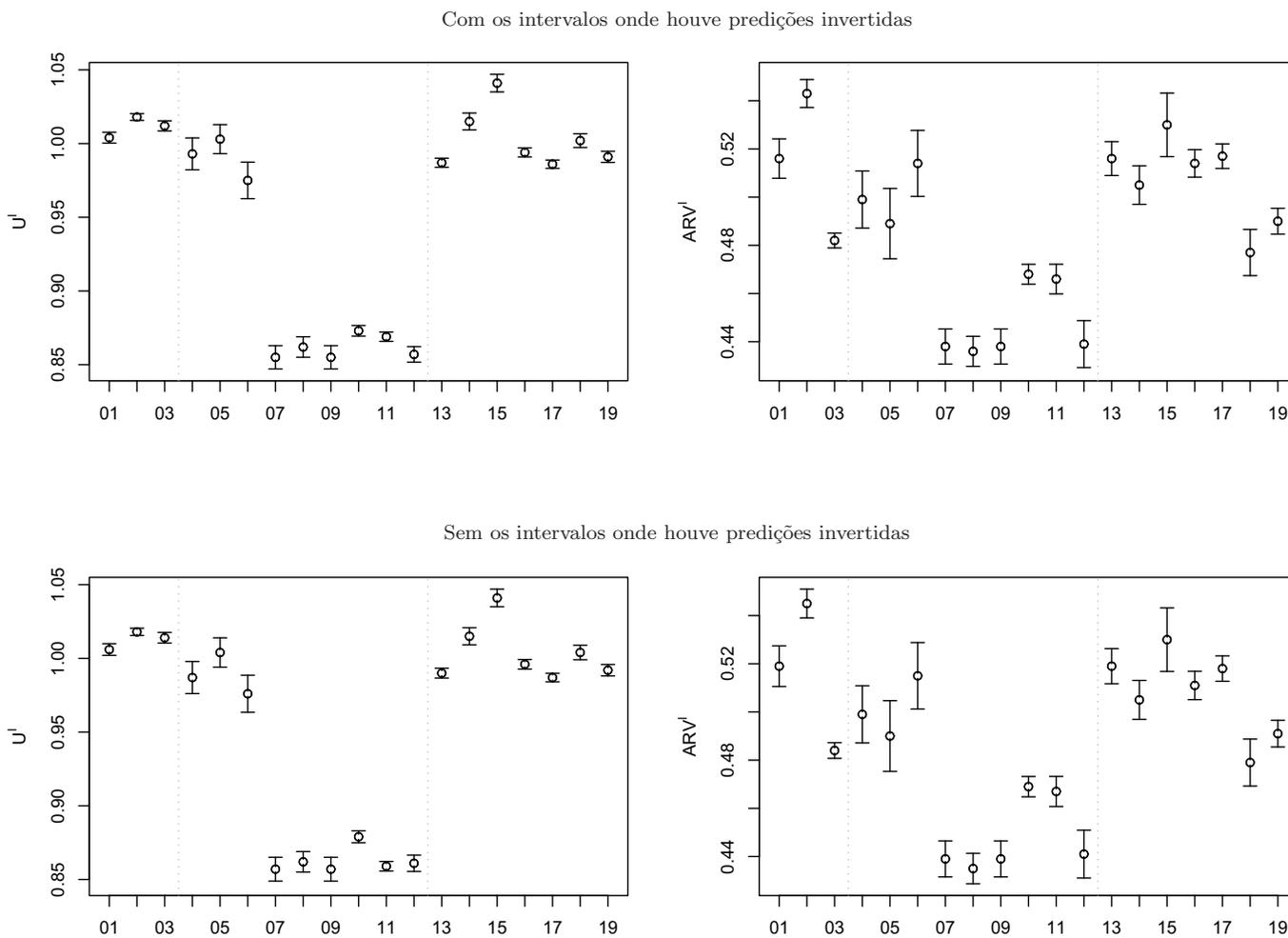


Tabela A.12: Resultados obtidos para o PGD-II com  $n = 620$ ; 20 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,299 (0,075)	0,542 (0,107)	1,320 (0,075)	0,549 (0,173)
MLP <sub>c,r</sub>	–	1	1,203 (0,044)	0,563 (0,076)	1,204 (0,044)	0,561 (0,124)
Híbrido <sub>c,r</sub>	–	1	1,124 (0,086)	0,518 (0,033)	1,155 (0,088)	0,522 (0,054)
AES <sup>I</sup>	–	2	1,151 (0,189)	0,502 (0,130)	1,159 (0,191)	0,503 (0,142)
AED <sup>I</sup>	–	3	1,087 (0,156)	0,501 (0,180)	1,085 (0,156)	0,515 (0,206)
HW <sup>I</sup>	–	1	1,055 (0,148)	0,501 (0,147)	1,061 (0,149)	0,519 (0,169)
iARMA-irrestrito	–	3	0,994 (0,117)	0,449 (0,089)	0,993 (0,117)	0,442 (0,120)
iARMA-restrito	–	0	0,987 (0,079)	0,443 (0,085)	0,972 (0,078)	0,449 (0,091)
iARMA-combinado	–	0	0,988 (0,099)	0,445 (0,125)	0,993 (0,117)	0,446 (0,120)
iARIMA	–	2	0,993 (0,060)	0,448 (0,054)	0,990 (0,060)	0,453 (0,084)
MLP <sup>I</sup>	–	0	0,994 (0,029)	0,428 (0,082)	0,996 (0,029)	0,423 (0,112)
Híbrido <sup>I</sup>	–	0	0,989 (0,096)	0,411 (0,128)	0,985 (0,095)	0,413 (0,178)
AES <sub>s,i</sub>	–	2	1,279 (0,048)	0,533 (0,097)	1,281 (0,049)	0,550 (0,154)
AED <sub>s,i</sub>	–	0	1,251 (0,111)	0,504 (0,085)	1,253 (0,111)	0,506 (0,093)
HW <sub>s,i</sub>	–	3	1,297 (0,113)	0,509 (0,114)	1,293 (0,113)	0,505 (0,121)
ARMA <sub>s,i</sub>	–	0	1,199 (0,057)	0,512 (0,068)	1,220 (0,057)	0,520 (0,108)
ARIMA <sub>s,i</sub>	–	1	1,265 (0,047)	0,502 (0,057)	1,267 (0,047)	0,515 (0,085)
MLP <sub>s,i</sub>	–	1	1,286 (0,067)	0,467 (0,117)	1,264 (0,067)	0,452 (0,169)
Híbrido <sub>s,i</sub>	–	0	1,142 (0,053)	0,498 (0,065)	1,139 (0,052)	0,498 (0,079)

Figura A.12: Intervalos de 95% de confiança para os resultados obtidos para o PGD-II com  $n = 620$ ; 20 passos-à-frente.

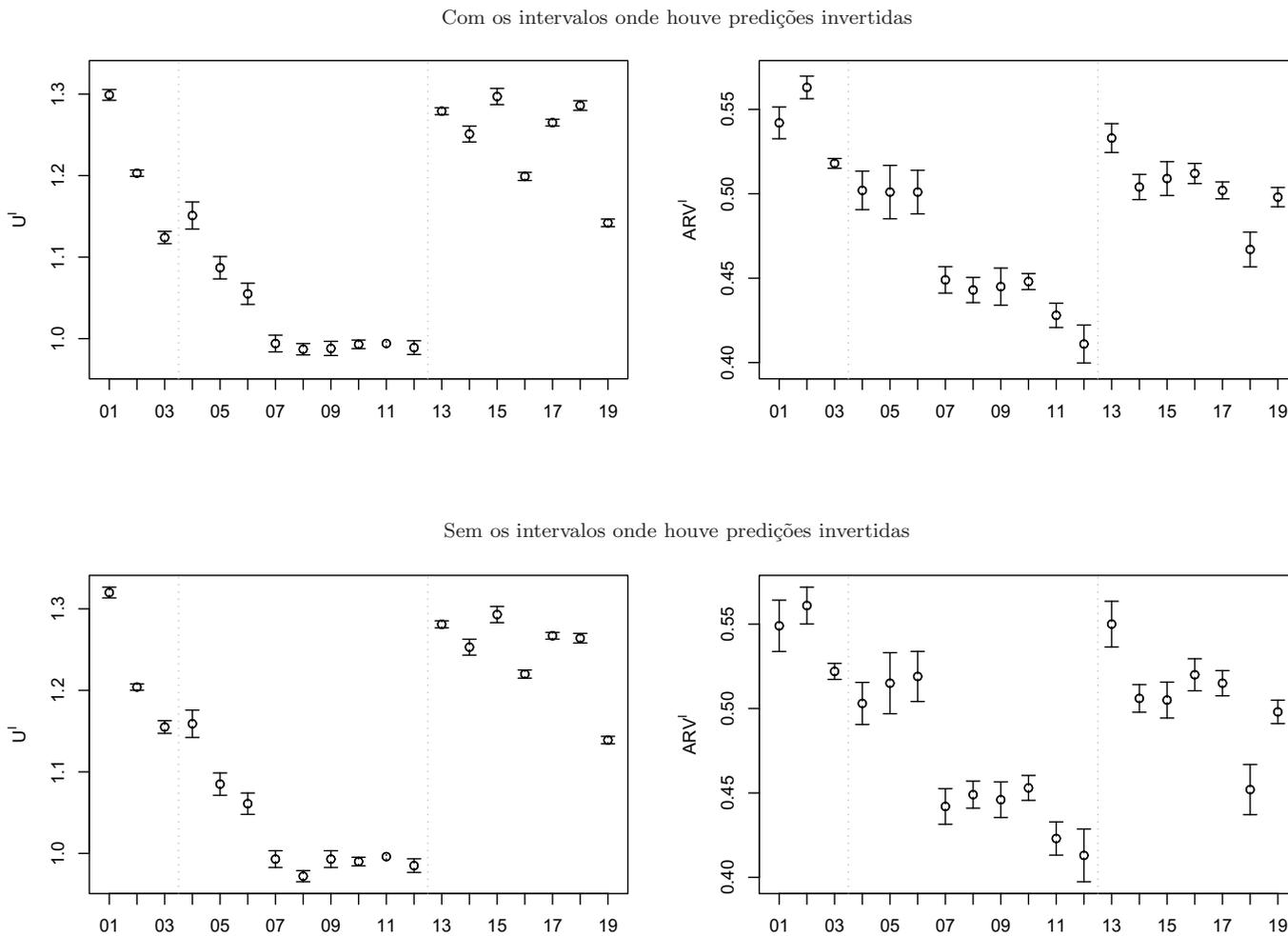


Tabela A.13: Resultados obtidos para o PGD-III com  $n = 220$ ; conjunto de treinamento.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	11,4	0	0,903 (0,033)	0,602 (0,066)	0,895 (0,033)	0,595 (0,067)
MLP <sub>c,r</sub>	8,3	1	0,840 (0,038)	0,538 (0,083)	0,831 (0,039)	0,528 (0,084)
Híbrido <sub>c,r</sub>	10,0	0	0,820 (0,032)	0,510 (0,076)	0,809 (0,031)	0,497 (0,076)
AES <sup>I</sup>	10,2	11	0,914 (0,111)	0,583 (0,102)	0,890 (0,113)	0,553 (0,101)
AED <sup>I</sup>	7,1	13	0,872 (0,109)	0,529 (0,140)	0,888 (0,115)	0,536 (0,141)
HW <sup>I</sup>	9,0	9	0,927 (0,123)	0,517 (0,129)	0,910 (0,124)	0,518 (0,128)
iARMA-irrestrito	6,7	4	0,786 (0,055)	0,496 (0,087)	0,761 (0,051)	0,487 (0,084)
iARMA-restrito	6,2	0	0,779 (0,020)	0,448 (0,061)	0,763 (0,018)	0,439 (0,059)
iARMA-combinado	6,2	0	0,782 (0,019)	0,485 (0,081)	0,761 (0,051)	0,487 (0,084)
iARIMA	8,2	10	0,795 (0,017)	0,479 (0,045)	0,789 (0,017)	0,484 (0,044)
MLP <sup>I</sup>	8,1	6	0,735 (0,045)	0,468 (0,088)	0,740 (0,047)	0,464 (0,089)
Híbrido <sup>I</sup>	5,0	0	0,716 (0,029)	0,405 (0,099)	0,712 (0,023)	0,413 (0,095)
AES <sub>s,i</sub>	8,6	8	0,935 (0,027)	0,621 (0,060)	0,944 (0,029)	0,591 (0,060)
AED <sub>s,i</sub>	7,0	5	0,927 (0,051)	0,615 (0,099)	0,905 (0,054)	0,594 (0,099)
HW <sub>s,i</sub>	9,3	7	0,951 (0,074)	0,600 (0,101)	0,906 (0,079)	0,585 (0,101)
ARMA <sub>s,i</sub>	11,8	6	0,911 (0,033)	0,553 (0,075)	0,899 (0,034)	0,542 (0,075)
ARIMA <sub>s,i</sub>	10,3	9	0,905 (0,024)	0,580 (0,072)	0,891 (0,025)	0,568 (0,072)
MLP <sub>s,i</sub>	8,7	3	0,828 (0,062)	0,498 (0,090)	0,803 (0,061)	0,503 (0,088)
Híbrido <sub>s,i</sub>	8,8	4	0,807 (0,021)	0,484 (0,068)	0,801 (0,019)	0,481 (0,066)

Figura A.13: Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com  $n = 220$ ; conjunto de treinamento.

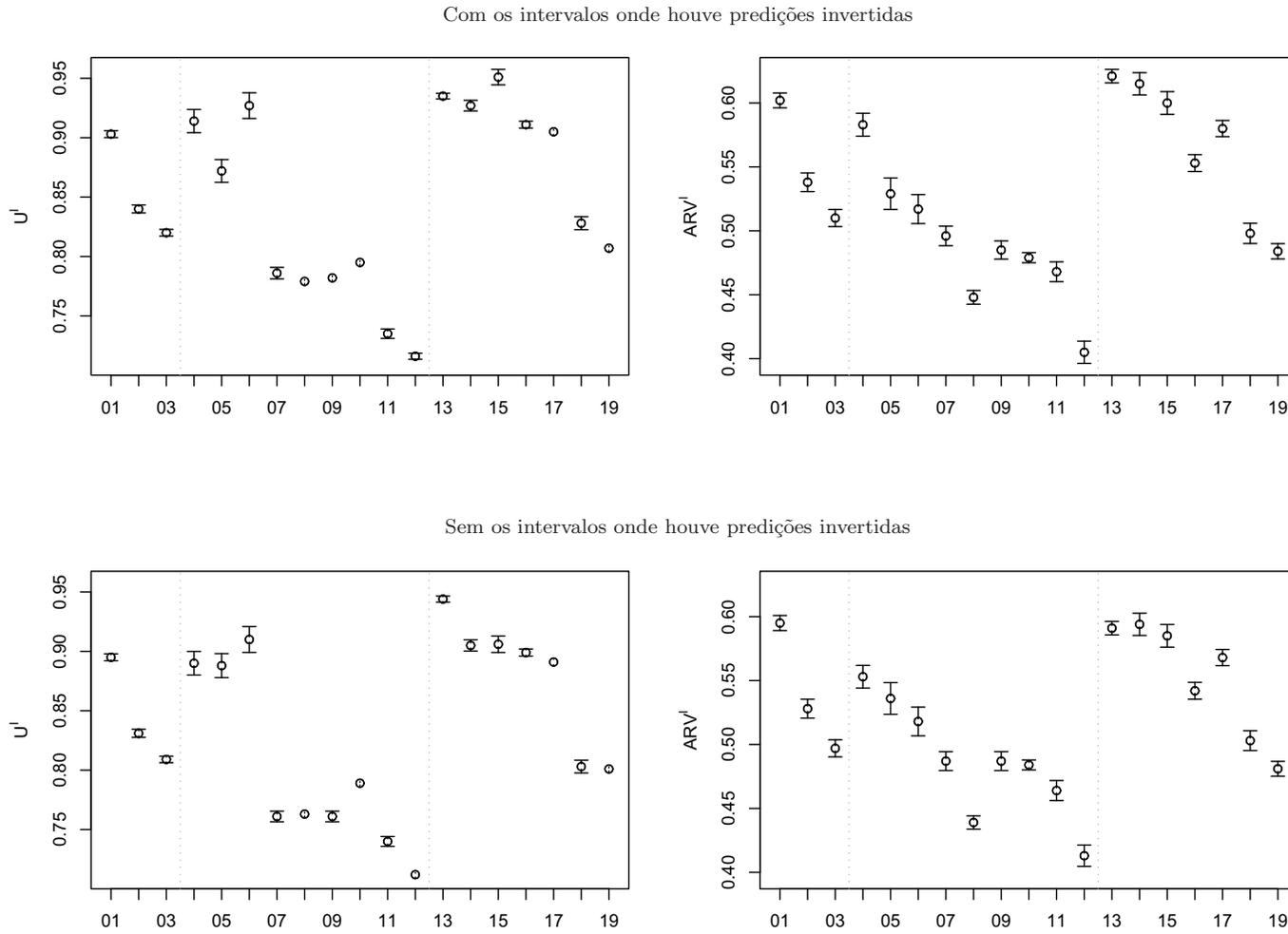


Tabela A.14: Resultados obtidos para o PGD-III com  $n = 220$ ; 5 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,026 (0,040)	0,585 (0,077)	1,026 (0,040)	0,585 (0,077)
MLP <sub>c,r</sub>	–	0	0,918 (0,046)	0,507 (0,100)	0,918 (0,046)	0,507 (0,100)
Híbrido <sub>c,r</sub>	–	0	0,904 (0,039)	0,514 (0,098)	0,904 (0,039)	0,514 (0,098)
AES <sup>I</sup>	–	0	0,910 (0,138)	0,543 (0,109)	0,910 (0,138)	0,543 (0,109)
AED <sup>I</sup>	–	0	0,951 (0,138)	0,526 (0,182)	0,951 (0,138)	0,526 (0,182)
HW <sup>I</sup>	–	0	0,938 (0,151)	0,524 (0,138)	0,938 (0,151)	0,524 (0,138)
iARMA-irrestrito	–	0	0,893 (0,060)	0,485 (0,095)	0,893 (0,060)	0,485 (0,095)
iARMA-restrito	–	0	0,852 (0,025)	0,449 (0,069)	0,852 (0,025)	0,449 (0,069)
iARMA-combinado	–	0	0,893 (0,060)	0,485 (0,095)	0,893 (0,060)	0,485 (0,095)
iARIMA	–	0	0,908 (0,021)	0,498 (0,054)	0,908 (0,021)	0,498 (0,054)
MLP <sup>I</sup>	–	0	0,809 (0,055)	0,455 (0,113)	0,809 (0,055)	0,455 (0,113)
Híbrido <sup>I</sup>	–	0	0,801 (0,037)	0,414 (0,123)	0,801 (0,037)	0,414 (0,123)
AES <sub>s,i</sub>	–	0	0,981 (0,033)	0,624 (0,068)	0,981 (0,033)	0,624 (0,068)
AED <sub>s,i</sub>	–	0	1,055 (0,055)	0,605 (0,104)	1,055 (0,055)	0,605 (0,104)
HW <sub>s,i</sub>	–	0	1,021 (0,093)	0,604 (0,118)	1,021 (0,093)	0,604 (0,118)
ARMA <sub>s,i</sub>	–	0	1,012 (0,040)	0,560 (0,086)	1,012 (0,040)	0,560 (0,086)
ARIMA <sub>s,i</sub>	–	0	1,034 (0,027)	0,567 (0,077)	1,034 (0,027)	0,567 (0,077)
MLP <sub>s,i</sub>	–	0	0,913 (0,074)	0,543 (0,109)	0,913 (0,074)	0,543 (0,109)
Híbrido <sub>s,i</sub>	–	0	0,901 (0,027)	0,525 (0,072)	0,901 (0,027)	0,525 (0,072)

Figura A.14: Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com  $n = 220$ ; 5 passos-à-frente.

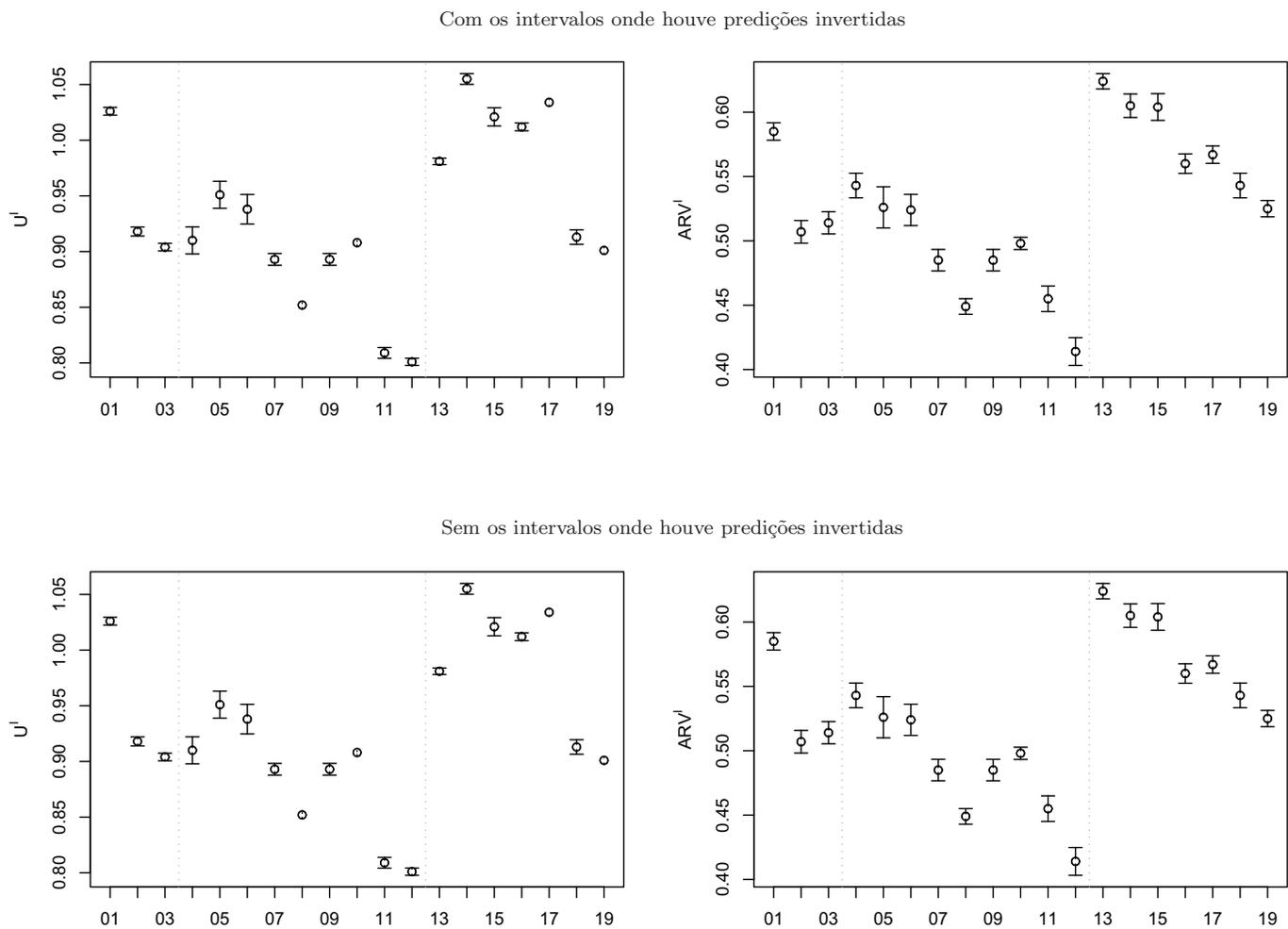


Tabela A.15: Resultados obtidos para o PGD-III com  $n = 220$ ; 20 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,135 (0,041)	0,576 (0,091)	1,135 (0,041)	0,584 (0,119)
MLP <sub>c,r</sub>	–	0	1,066 (0,068)	0,507 (0,098)	1,058 (0,068)	0,499 (0,149)
Híbrido <sub>c,r</sub>	–	0	0,997 (0,085)	0,516 (0,083)	0,989 (0,085)	0,517 (0,083)
AES <sup>I</sup>	–	3	1,074 (0,180)	0,563 (0,143)	1,043 (0,175)	0,568 (0,173)
AED <sup>I</sup>	–	2	1,077 (0,174)	0,533 (0,148)	1,077 (0,174)	0,551 (0,174)
HW <sup>I</sup>	–	1	1,008 (0,191)	0,559 (0,144)	1,029 (0,195)	0,565 (0,151)
iARMA-irrestrito	–	2	0,979 (0,097)	0,478 (0,101)	0,978 (0,097)	0,475 (0,097)
iARMA-restrito	–	0	0,991 (0,089)	0,470 (0,065)	0,992 (0,089)	0,466 (0,076)
iARMA-combinado	–	0	0,988 (0,094)	0,483 (0,094)	0,978 (0,097)	0,475 (0,097)
iARIMA	–	2	1,094 (0,046)	0,503 (0,054)	1,089 (0,046)	0,509 (0,060)
MLP <sup>I</sup>	–	0	0,870 (0,061)	0,444 (0,118)	0,871 (0,061)	0,460 (0,144)
Híbrido <sup>I</sup>	–	0	0,874 (0,041)	0,427 (0,117)	0,880 (0,041)	0,432 (0,136)
AES <sub>s,i</sub>	–	4	1,313 (0,077)	0,615 (0,065)	1,301 (0,076)	0,620 (0,108)
AED <sub>s,i</sub>	–	4	1,317 (0,111)	0,595 (0,133)	1,276 (0,108)	0,603 (0,142)
HW <sub>s,i</sub>	–	1	1,295 (0,109)	0,577 (0,133)	1,288 (0,109)	0,577 (0,160)
ARMA <sub>s,i</sub>	–	2	1,201 (0,094)	0,559 (0,083)	1,199 (0,094)	0,572 (0,098)
ARIMA <sub>s,i</sub>	–	0	1,250 (0,035)	0,580 (0,079)	1,249 (0,035)	0,591 (0,104)
MLP <sub>s,i</sub>	–	1	1,066 (0,094)	0,497 (0,103)	1,065 (0,093)	0,488 (0,094)
Híbrido <sub>s,i</sub>	–	2	1,082 (0,069)	0,479 (0,077)	1,089 (0,069)	0,483 (0,105)

Figura A.15: Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com  $n = 220$ ; 20 passos-à-frente.

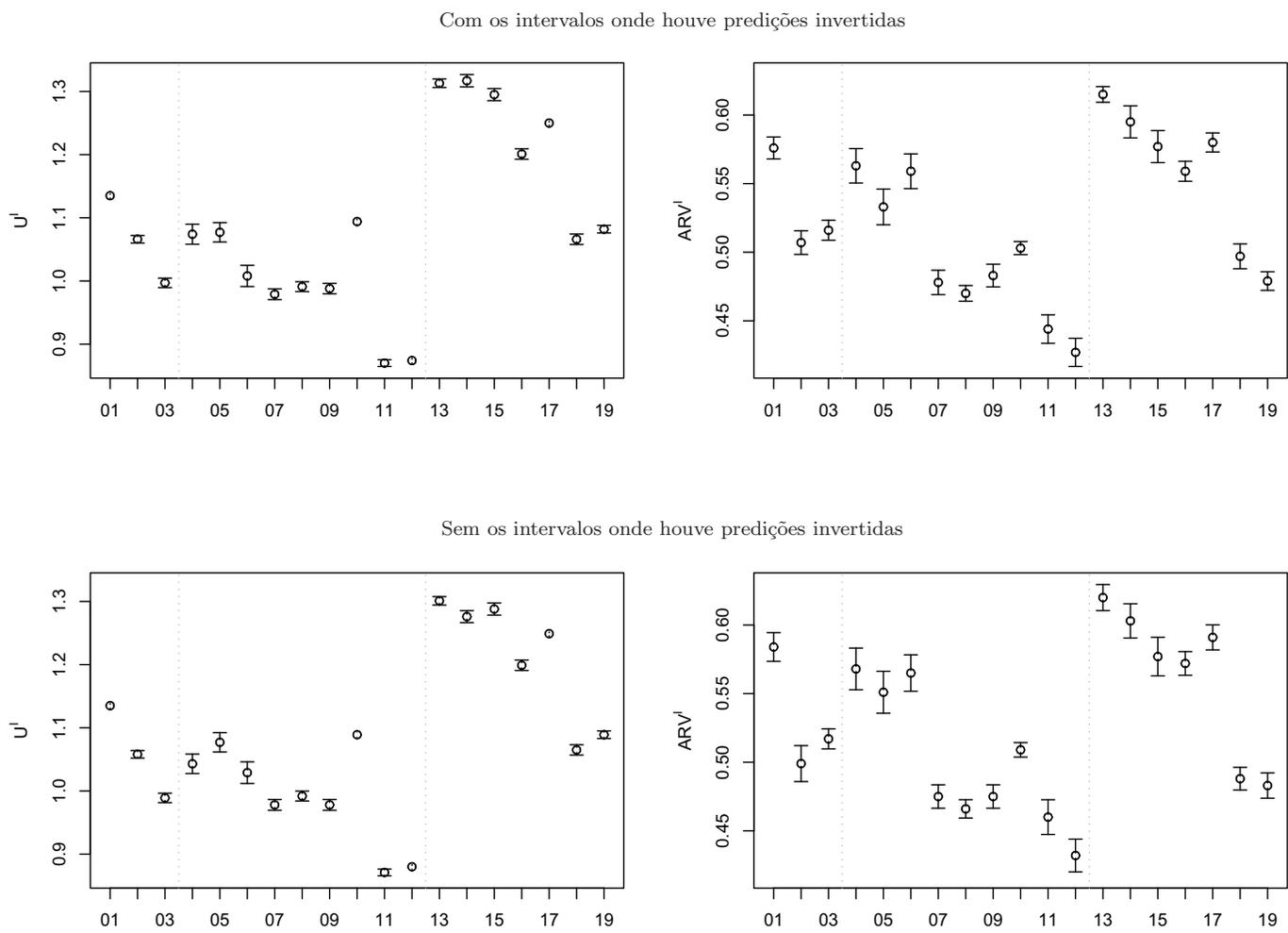


Tabela A.16: Resultados obtidos para o PGD-III com  $n = 620$ ; conjunto de treinamento.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	10,4	0	0,924 (0,016)	0,487 (0,030)	0,912 (0,009)	0,489 (0,034)
MLP <sub>c,r</sub>	7,0	3	0,854 (0,050)	0,404 (0,096)	0,866 (0,072)	0,407 (0,091)
Híbrido <sub>c,r</sub>	8,7	0	0,856 (0,063)	0,371 (0,094)	0,843 (0,103)	0,367 (0,085)
AES <sup>I</sup>	10,2	14	0,934 (0,106)	0,472 (0,168)	0,915 (0,100)	0,458 (0,172)
AED <sup>I</sup>	7,2	9	0,898 (0,087)	0,461 (0,183)	0,882 (0,081)	0,467 (0,186)
HW <sup>I</sup>	8,2	7	0,923 (0,125)	0,451 (0,204)	0,935 (0,126)	0,441 (0,203)
iARMA-irrestrito	5,8	9	0,803 (0,078)	0,398 (0,098)	0,807 (0,081)	0,370 (0,097)
iARMA-restrito	5,8	0	0,796 (0,069)	0,401 (0,124)	0,781 (0,067)	0,387 (0,125)
iARMA-combinado	5,4	0	0,781 (0,068)	0,382 (0,129)	0,807 (0,081)	0,370 (0,097)
iARIMA	7,4	11	0,805 (0,033)	0,411 (0,047)	0,821 (0,047)	0,399 (0,044)
MLP <sup>I</sup>	6,3	5	0,739 (0,050)	0,330 (0,088)	0,755 (0,075)	0,338 (0,082)
Híbrido <sup>I</sup>	4,6	0	0,723 (0,058)	0,367 (0,074)	0,701 (0,057)	0,331 (0,075)
AES <sub>s,i</sub>	7,5	17	0,940 (0,042)	0,489 (0,074)	0,941 (0,062)	0,493 (0,070)
AED <sub>s,i</sub>	7,3	9	0,950 (0,030)	0,515 (0,056)	0,964 (0,022)	0,515 (0,061)
HW <sub>s,i</sub>	8,8	13	0,983 (0,086)	0,508 (0,145)	0,963 (0,108)	0,506 (0,140)
ARMA <sub>s,i</sub>	10,2	11	0,946 (0,041)	0,461 (0,077)	0,960 (0,054)	0,462 (0,074)
ARIMA <sub>s,i</sub>	9,7	8	0,929 (0,037)	0,427 (0,062)	0,915 (0,049)	0,429 (0,060)
MLP <sub>s,i</sub>	8,3	4	0,859 (0,032)	0,444 (0,042)	0,868 (0,024)	0,425 (0,047)
Híbrido <sub>s,i</sub>	8,0	5	0,822 (0,049)	0,398 (0,069)	0,804 (0,065)	0,379 (0,065)

Figura A.16: Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com  $n = 620$ ; conjunto de treinamento.

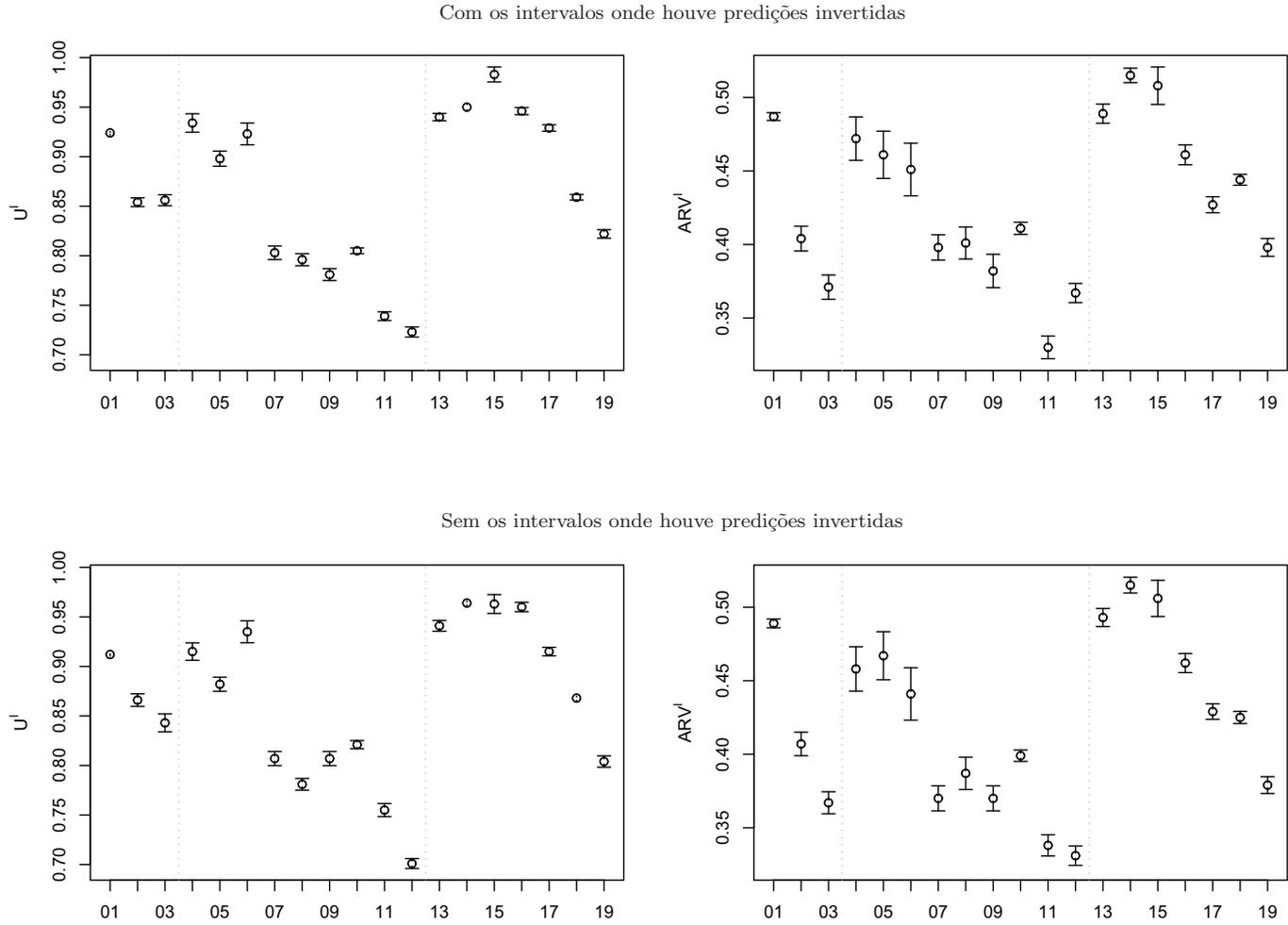


Tabela A.17: Resultados obtidos para o PGD-III com  $n = 620$ ; 5 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,968 (0,020)	0,470 (0,040)	0,971 (0,024)	0,474 (0,043)
MLP <sub>c,r</sub>	–	0	0,948 (0,059)	0,423 (0,101)	0,954 (0,066)	0,423 (0,107)
Híbrido <sub>c,r</sub>	–	0	0,899 (0,064)	0,405 (0,109)	0,899 (0,067)	0,405 (0,112)
AES <sup>I</sup>	–	0	0,980 (0,110)	0,471 (0,184)	0,980 (0,116)	0,471 (0,190)
AED <sup>I</sup>	–	0	0,983 (0,109)	0,467 (0,210)	0,983 (0,113)	0,472 (0,215)
HW <sup>I</sup>	–	0	0,931 (0,160)	0,444 (0,234)	0,937 (0,166)	0,450 (0,240)
iARMA-irrestrito	–	0	0,823 (0,082)	0,398 (0,119)	0,829 (0,088)	0,404 (0,125)
iARMA-restrito	–	0	0,825 (0,070)	0,411 (0,151)	0,830 (0,075)	0,416 (0,157)
iARMA-combinado	–	0	0,823 (0,082)	0,398 (0,119)	0,829 (0,088)	0,404 (0,125)
iARIMA	–	0	0,843 (0,033)	0,406 (0,057)	0,843 (0,037)	0,410 (0,060)
MLP <sup>I</sup>	–	0	0,826 (0,062)	0,386 (0,096)	0,831 (0,067)	0,386 (0,101)
Híbrido <sup>I</sup>	–	0	0,812 (0,071)	0,377 (0,086)	0,812 (0,072)	0,378 (0,087)
AES <sub>s,i</sub>	–	1	1,031 (0,049)	0,506 (0,082)	1,037 (0,055)	0,506 (0,088)
AED <sub>s,i</sub>	–	0	1,042 (0,032)	0,512 (0,067)	1,043 (0,033)	0,513 (0,068)
HW <sub>s,i</sub>	–	0	1,063 (0,095)	0,493 (0,155)	1,066 (0,098)	0,496 (0,158)
ARMA <sub>s,i</sub>	–	0	0,989 (0,053)	0,459 (0,096)	0,992 (0,056)	0,462 (0,100)
ARIMA <sub>s,i</sub>	–	1	0,980 (0,039)	0,419 (0,081)	0,988 (0,047)	0,426 (0,089)
MLP <sub>s,i</sub>	–	0	0,961 (0,033)	0,434 (0,051)	0,965 (0,037)	0,438 (0,055)
Híbrido <sub>s,i</sub>	–	0	0,920 (0,058)	0,403 (0,085)	0,926 (0,064)	0,403 (0,091)

Figura A.17: Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com  $n = 620$ ; 5 passos-à-frente.

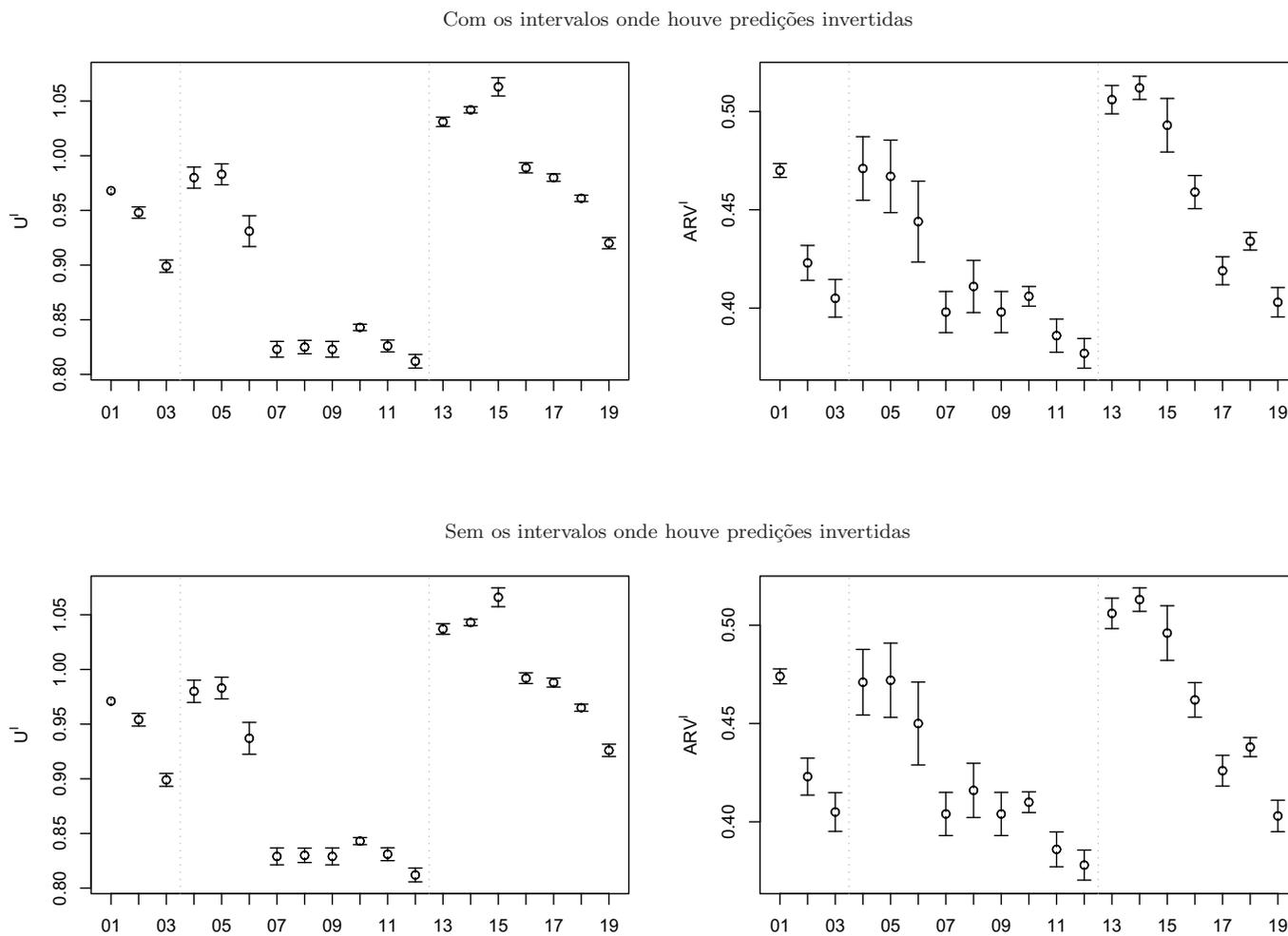
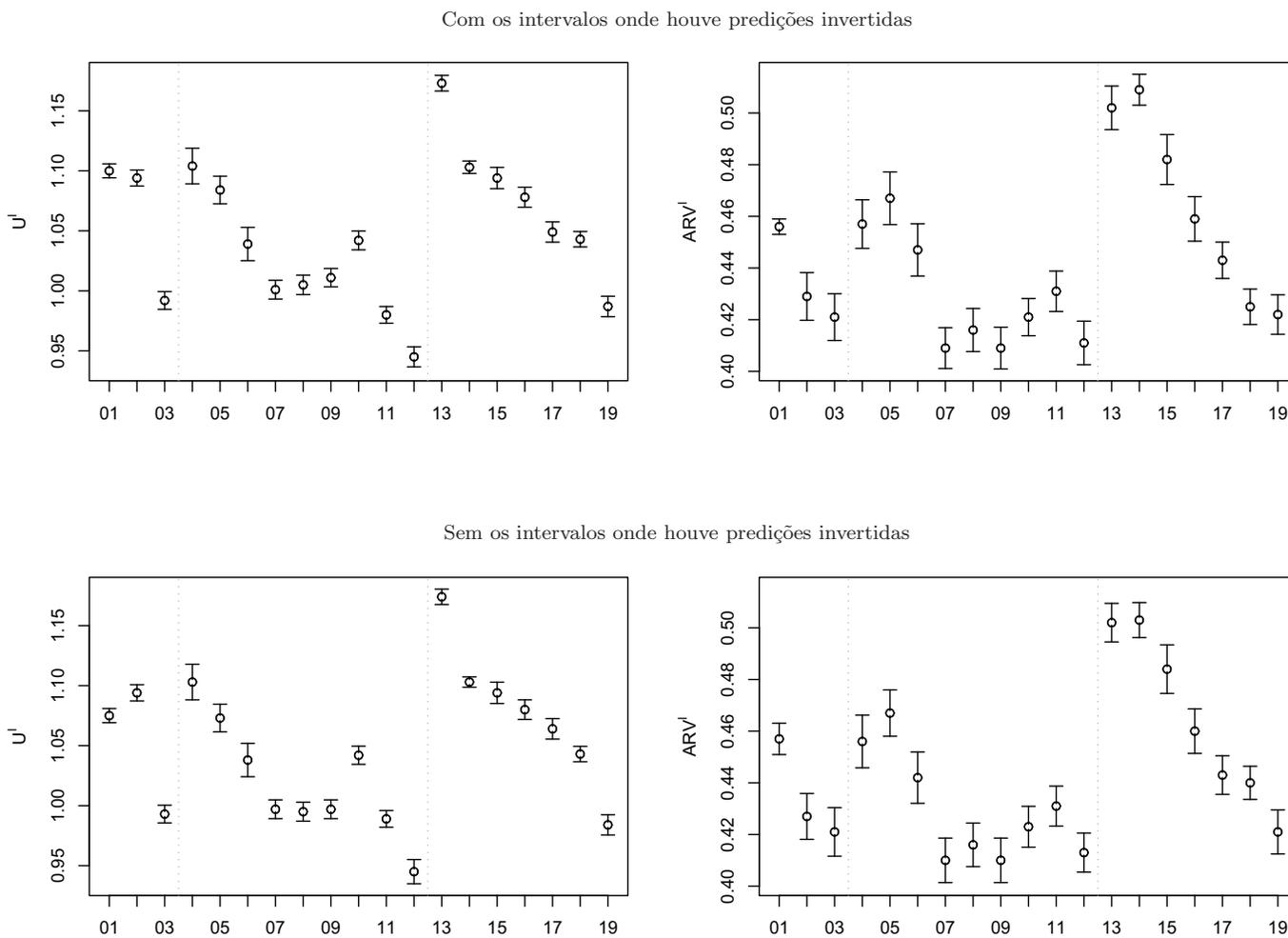


Tabela A.18: Resultados obtidos para o PGD-III com  $n = 620$ ; 20 passos-à-frente.

Modelo	$\tilde{H}_0^I$	Com os intervalos onde houve predições invertidas			Sem os intervalos onde houve predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,100 (0,066)	0,456 (0,034)	1,075 (0,067)	0,457 (0,069)
MLP <sub>c,r</sub>	–	0	1,094 (0,076)	0,429 (0,105)	1,094 (0,077)	0,427 (0,101)
Híbrido <sub>c,r</sub>	–	0	0,992 (0,084)	0,421 (0,103)	0,993 (0,084)	0,421 (0,107)
AES <sup>I</sup>	–	2	1,104 (0,169)	0,457 (0,107)	1,103 (0,169)	0,456 (0,116)
AED <sup>I</sup>	–	3	1,084 (0,131)	0,467 (0,116)	1,073 (0,131)	0,467 (0,102)
HW <sup>I</sup>	–	1	1,039 (0,158)	0,447 (0,115)	1,038 (0,158)	0,442 (0,113)
iARMA-irrestrito	–	1	1,001 (0,089)	0,409 (0,090)	0,997 (0,089)	0,410 (0,098)
iARMA-restrito	–	0	1,005 (0,092)	0,416 (0,095)	0,995 (0,090)	0,416 (0,096)
iARMA-combinado	–	0	1,011 (0,087)	0,409 (0,092)	0,997 (0,089)	0,410 (0,098)
iARIMA	–	1	1,042 (0,089)	0,421 (0,082)	1,042 (0,086)	0,423 (0,090)
MLP <sup>I</sup>	–	2	0,980 (0,079)	0,431 (0,089)	0,989 (0,079)	0,431 (0,088)
Híbrido <sup>I</sup>	–	0	0,945 (0,095)	0,411 (0,096)	0,945 (0,115)	0,413 (0,086)
AES <sub>s,i</sub>	–	2	1,173 (0,074)	0,502 (0,096)	1,174 (0,073)	0,502 (0,085)
AED <sub>s,i</sub>	–	2	1,103 (0,059)	0,509 (0,068)	1,103 (0,049)	0,503 (0,077)
HW <sub>s,i</sub>	–	1	1,094 (0,101)	0,482 (0,110)	1,094 (0,101)	0,484 (0,107)
ARMA <sub>s,i</sub>	–	3	1,078 (0,095)	0,459 (0,098)	1,080 (0,093)	0,460 (0,098)
ARIMA <sub>s,i</sub>	–	2	1,049 (0,096)	0,443 (0,080)	1,064 (0,097)	0,443 (0,085)
MLP <sub>s,i</sub>	–	0	1,043 (0,073)	0,425 (0,078)	1,043 (0,073)	0,440 (0,073)
Híbrido <sub>s,i</sub>	–	2	0,987 (0,097)	0,422 (0,087)	0,984 (0,096)	0,421 (0,097)

Figura A.18: Intervalos de 95% de confiança para os resultados obtidos para o PGD-III com  $n = 620$ ; 20 passos-à-frente.



## A.2 Tabelas e gráficos da aplicação a STI reais

Tabela A.19: Resultados obtidos para a STI de preços de ações da Petrobras; conjunto de treinamento.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$		0	0,992	0,650	0,994	0,647
$MLP_{c,r}$		0	0,989	0,622	0,985	0,618
Híbrido $_{c,r}$		0	0,971	0,627	0,973	0,623
$AES^I$	•	3	0,991	0,663	0,991	0,692
$AED^I$		2	0,982	0,647	0,982	0,665
$HW^I$		0	0,977	0,648	0,977	0,665
iARMA-irrestrito		2	0,918	0,511	0,914	0,513
iARMA-restrito		0	0,912	0,507	0,908	0,508
iARMA-combinado		0	0,915	0,509	0,914	0,513
iARIMA		0	0,919	0,510	0,917	0,512
$MLP^I$		0	0,925	0,527	0,923	0,529
Híbrido $^I$		0	0,901	0,503	0,903	0,500
$AES_{s,i}$		2	1,013	0,733	1,023	0,726
$AED_{s,i}$	•	0	1,021	0,734	1,027	0,732
$HW_{s,i}$		1	1,028	0,746	1,030	0,750
$ARMA_{s,i}$		2	0,978	0,639	0,980	0,624
$ARIMA_{s,i}$		1	0,980	0,642	0,987	0,636
$MLP_{s,i}$		0	0,991	0,665	0,993	0,672
Híbrido $_{s,i}$		0	0,968	0,616	0,968	0,617

Figura A.19: Valores das medidas de precisão obtidos para a STI da Petrobras; conjunto de treinamento.

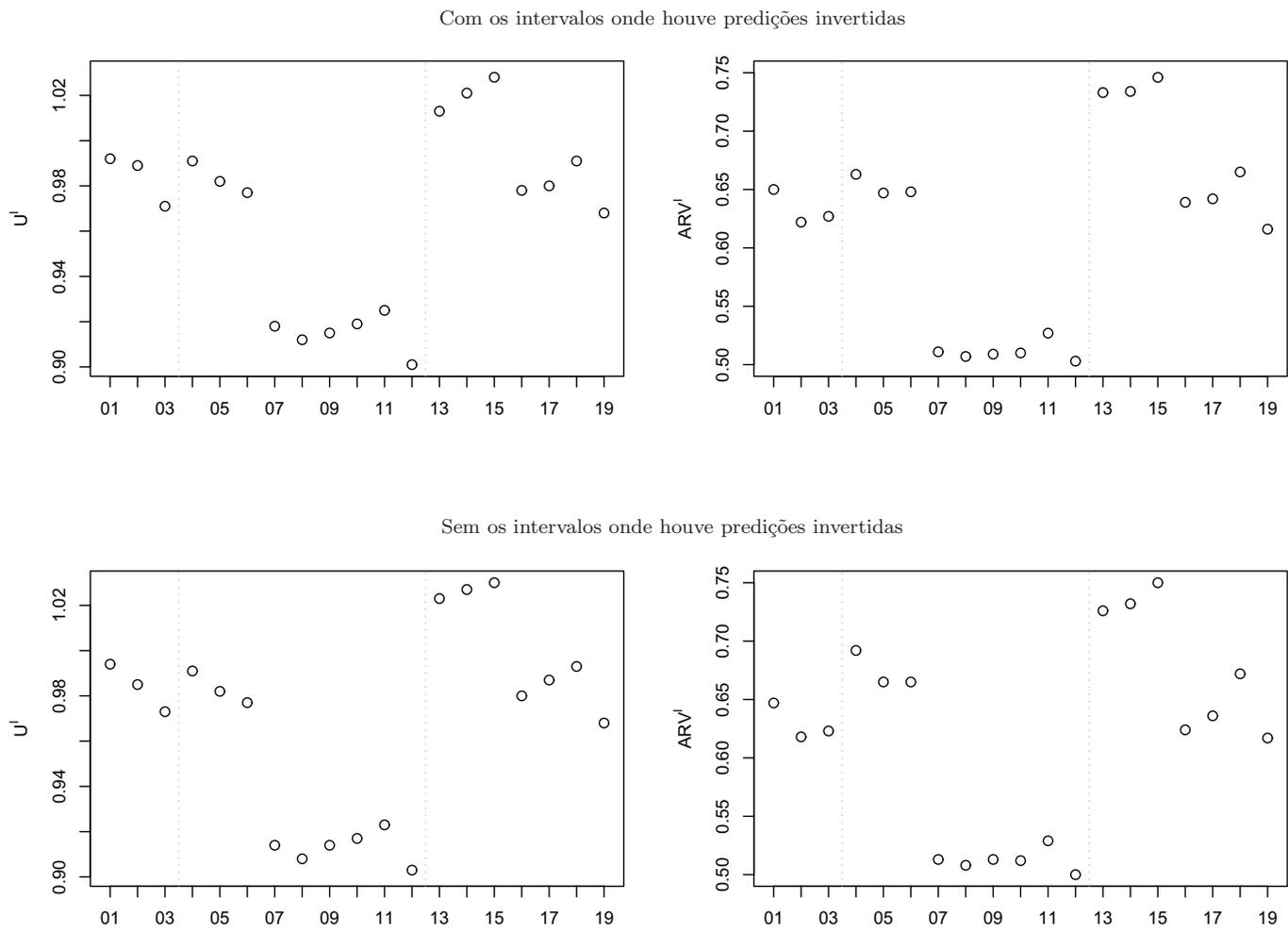


Tabela A.20: Resultados obtidos para a STI de preços de ações da Petrobras; 5 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,013	0,665	1,013	0,665
MLP <sub>c,r</sub>	–	0	1,008	0,658	1,008	0,658
Híbrido <sub>c,r</sub>	–	0	1,001	0,637	1,001	0,637
AES <sup>I</sup>	–	0	1,009	0,706	1,009	0,706
AED <sup>I</sup>	–	0	1,011	0,693	1,011	0,693
HW <sup>I</sup>	–	0	1,014	0,697	1,014	0,697
iARMA-irrestrito	–	0	0,934	0,605	0,934	0,605
iARMA-restrito	–	0	0,940	0,601	0,940	0,601
iARMA-combinado	–	0	0,934	0,605	0,934	0,605
iARIMA	–	0	0,943	0,622	0,943	0,622
MLP <sup>I</sup>	–	0	0,945	0,613	0,945	0,613
Híbrido <sup>I</sup>	–	0	0,936	0,582	0,936	0,582
AES <sub>s,i</sub>	–	0	1,080	0,786	1,080	0,786
AED <sub>s,i</sub>	–	0	1,089	0,777	1,089	0,777
HW <sub>s,i</sub>	–	0	1,102	0,752	1,102	0,752
ARMA <sub>s,i</sub>	–	0	1,010	0,650	1,010	0,650
ARIMA <sub>s,i</sub>	–	0	1,011	0,653	1,011	0,653
MLP <sub>s,i</sub>	–	0	1,002	0,660	1,002	0,660
Híbrido <sub>s,i</sub>	–	0	0,998	0,629	0,998	0,629

Figura A.20: Valores das medidas de precisão obtidos para a STI da Petrobras; 5 passos-à-frente.

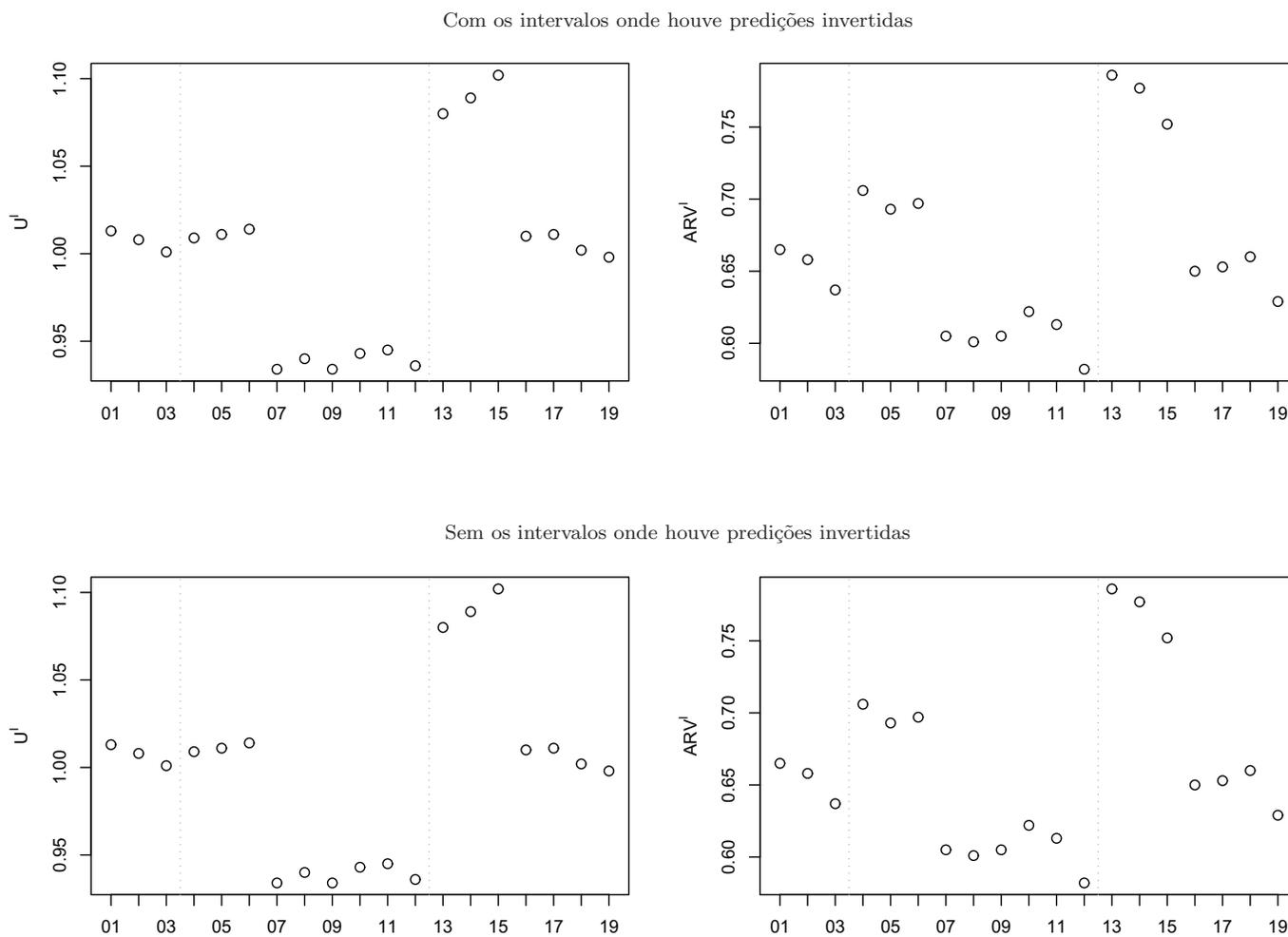


Tabela A.21: Resultados obtidos para a STI de preços de ações da Petrobras; 20 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,103	0,700	1,101	0,700
MLP <sub>c,r</sub>	–	0	1,089	0,697	1,098	0,691
Híbrido <sub>c,r</sub>	–	0	1,101	0,677	1,104	0,675
AES <sup>I</sup>	–	1	1,092	0,720	1,090	0,728
AED <sup>I</sup>	–	0	1,097	0,711	1,098	0,717
HW <sup>I</sup>	–	0	1,098	0,716	1,101	0,722
iARMA-irrestrito	–	1	1,001	0,621	1,008	0,618
iARMA-restrito	–	0	1,002	0,628	1,004	0,623
iARMA-combinado	–	0	1,002	0,623	1,008	0,618
iARIMA	–	0	1,021	0,637	1,019	0,636
MLP <sup>I</sup>	–	0	1,012	0,636	1,018	0,637
Híbrido <sup>I</sup>	–	0	0,986	0,615	0,985	0,615
AES <sub>s,i</sub>	–	0	1,236	0,802	1,243	0,793
AED <sub>s,i</sub>	–	0	1,229	0,791	1,234	0,790
HW <sub>s,i</sub>	–	0	1,263	0,787	1,251	0,784
ARMA <sub>s,i</sub>	–	0	1,091	0,699	1,098	0,697
ARIMA <sub>s,i</sub>	–	1	1,101	0,703	1,106	0,701
MLP <sub>s,i</sub>	–	0	1,098	0,689	1,095	0,689
Híbrido <sub>s,i</sub>	–	0	1,079	0,665	1,076	0,668

Figura A.21: Valores das medidas de precisão obtidos para a STI da Petrobras; 20 passos-à-frente.

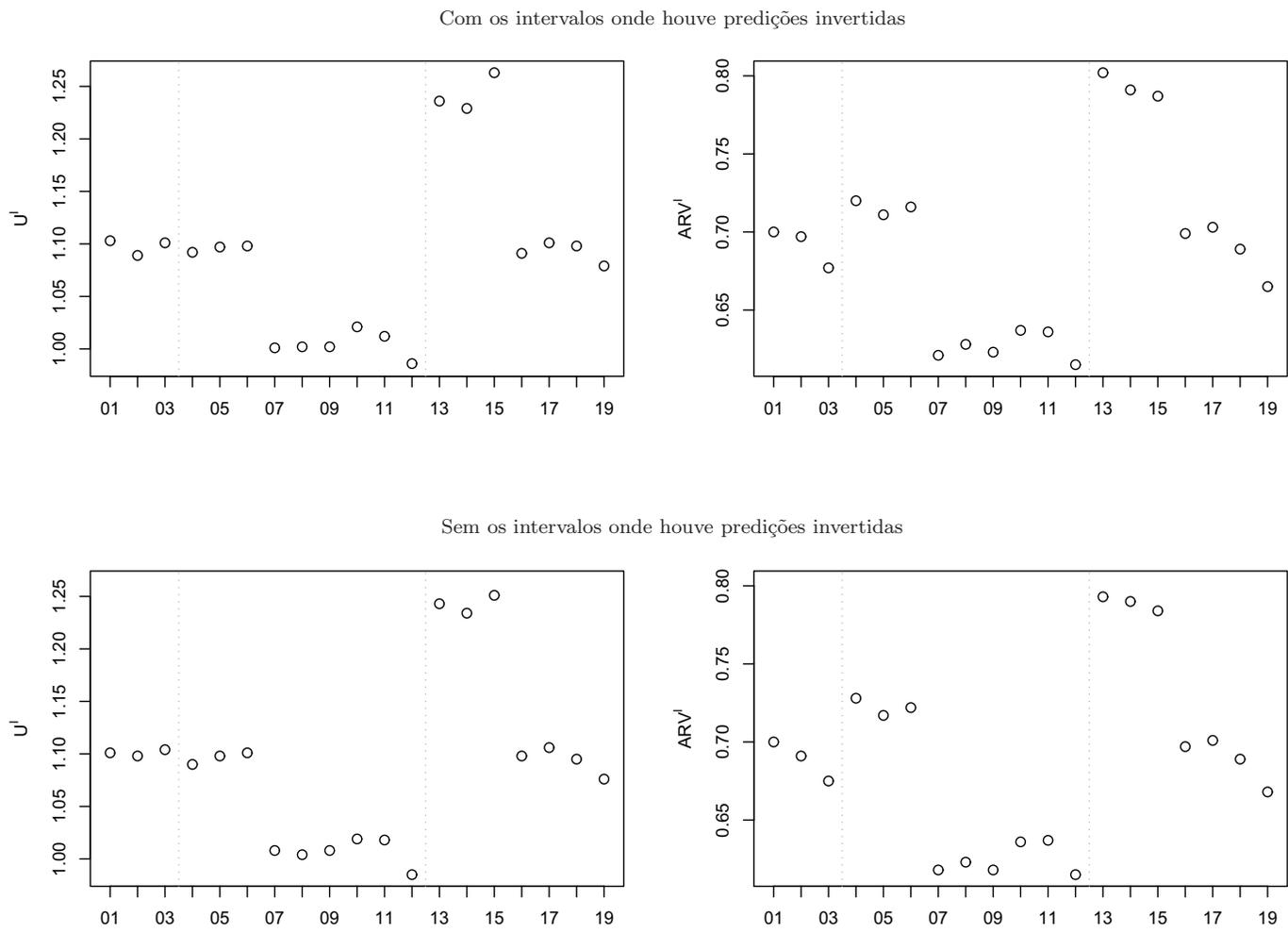


Tabela A.22: Resultados obtidos para a STI de preços de ações da Google; conjunto de treinamento.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$		0	0,979	0,672	0,972	0,670
$MLP_{c,r}$		2	0,990	0,680	0,987	0,683
Híbrido $_{c,r}$		0	0,981	0,669	0,978	0,669
$AES^I$		11	0,962	0,661	0,959	0,665
$AED^I$		9	0,956	0,641	0,955	0,644
$HW^I$	•	14	0,960	0,638	0,961	0,641
iARMA-irrestrito		13	0,935	0,610	0,934	0,615
iARMA-restrito		0	0,928	0,605	0,933	0,613
iARMA-combinado		0	0,927	0,604	0,934	0,615
iARIMA		10	0,931	0,620	0,925	0,624
$MLP^I$		3	0,927	0,616	0,922	0,621
Híbrido $^I$		0	0,914	0,603	0,918	0,609
$AES_{s,i}$		13	1,023	0,721	1,020	0,727
$AED_{s,i}$	•	11	1,030	0,749	1,027	0,752
$HW_{s,i}$		6	1,019	0,736	1,015	0,735
$ARMA_{s,i}$		4	0,983	0,703	0,985	0,711
$ARIMA_{s,i}$		2	0,980	0,708	0,980	0,718
$MLP_{s,i}$	•	2	0,979	0,675	0,976	0,689
Híbrido $_{s,i}$		4	0,961	0,681	0,956	0,680

Figura A.22: Valores das medidas de precisão obtidos para a STI da Google; conjunto de treinamento.

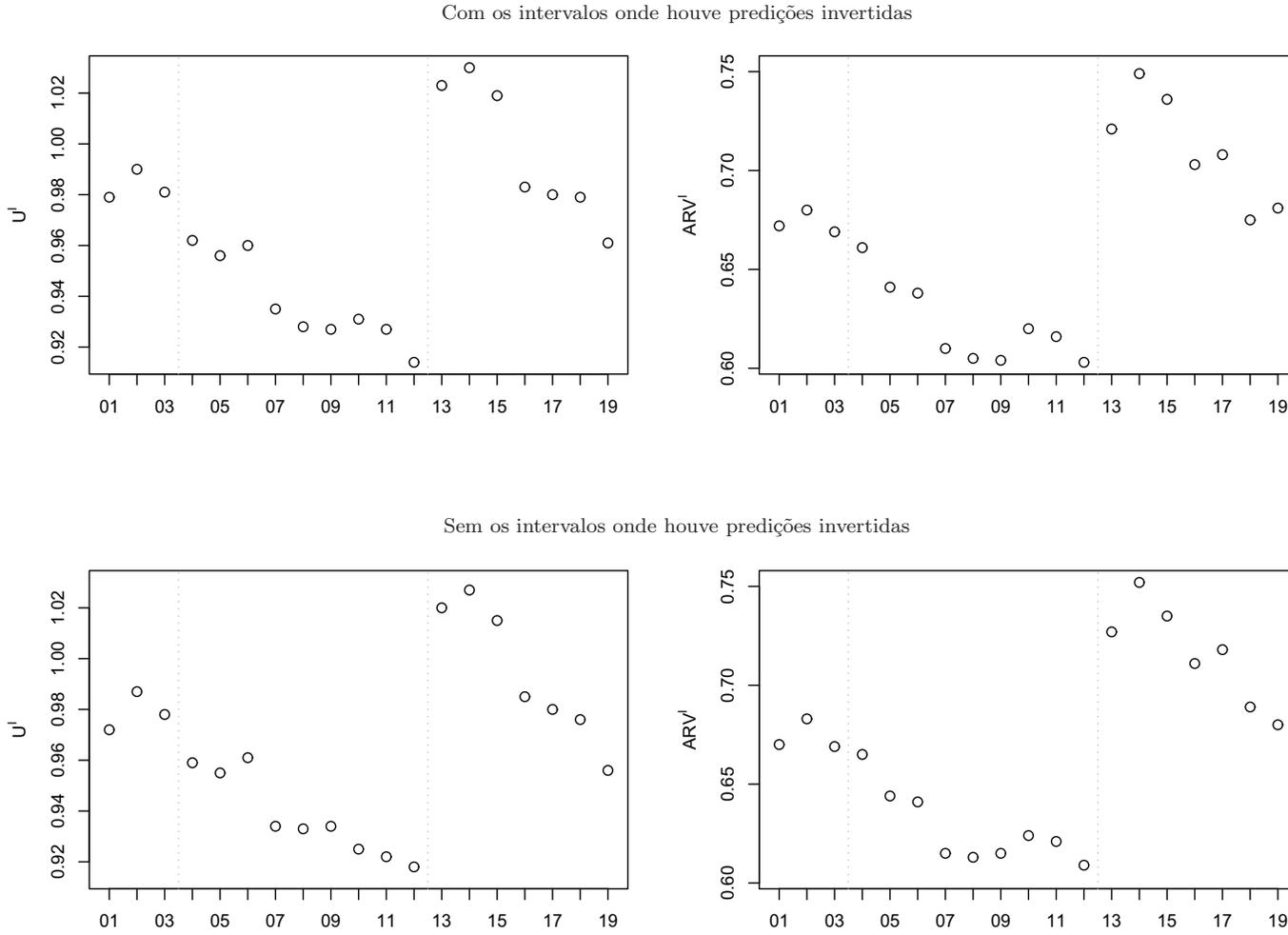


Tabela A.23: Resultados obtidos para a STI de preços de ações da Google; 5 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,984	0,729	0,986	0,730
MLP <sub>c,r</sub>	–	0	0,999	0,732	1,001	0,735
Híbrido <sub>c,r</sub>	–	0	1,001	0,723	1,002	0,726
AES <sup>I</sup>	–	0	1,010	0,727	1,010	0,730
AED <sup>I</sup>	–	0	1,009	0,714	1,009	0,716
HW <sup>I</sup>	–	1	1,013	0,714	1,011	0,717
iARMA-irrestrito	–	0	0,977	0,710	0,980	0,713
iARMA-restrito	–	0	0,970	0,704	0,973	0,705
iARMA-combinado	–	0	0,977	0,710	0,980	0,713
iARIMA	–	1	0,991	0,725	0,994	0,727
MLP <sup>I</sup>	–	0	0,946	0,712	0,948	0,714
Híbrido <sup>I</sup>	–	0	0,934	0,680	0,937	0,682
AES <sub>s,i</sub>	–	0	1,128	0,782	1,131	0,780
AED <sub>s,i</sub>	–	0	1,148	0,786	1,150	0,789
HW <sub>s,i</sub>	–	0	1,100	0,770	1,103	0,771
ARMA <sub>s,i</sub>	–	0	1,036	0,762	1,033	0,765
ARIMA <sub>s,i</sub>	–	0	1,026	0,763	1,029	0,766
MLP <sub>s,i</sub>	–	0	1,009	0,744	1,012	0,747
Híbrido <sub>s,i</sub>	–	0	1,002	0,743	1,005	0,746

Figura A.23: Valores das medidas de precisão obtidos para a STI da Google; 5 passos-à-frente.

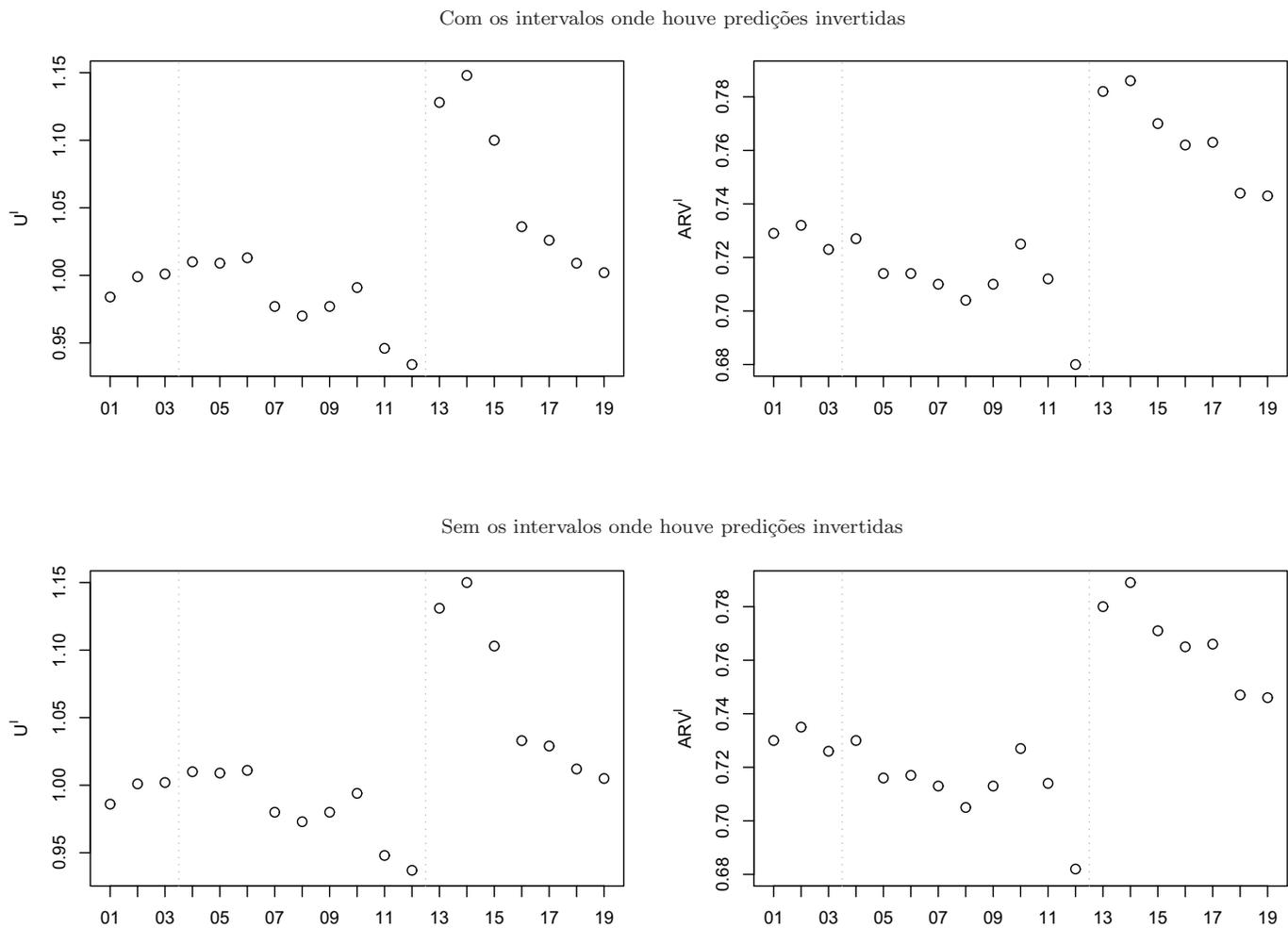


Tabela A.24: Resultados obtidos para a STI de preços de ações da Google; 20 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,158	0,743	1,161	0,742
MLP <sub>c,r</sub>	–	0	1,179	0,748	1,183	0,751
Híbrido <sub>c,r</sub>	–	0	1,155	0,731	1,160	0,729
AES <sup>I</sup>	–	1	1,171	0,742	1,175	0,747
AED <sup>I</sup>	–	0	1,172	0,729	1,178	0,723
HW <sup>I</sup>	–	1	1,199	0,727	1,198	0,731
iARMA-irrestrito	–	1	1,079	0,701	1,076	0,696
iARMA-restrito	–	0	1,079	0,700	1,082	0,696
iARMA-combinado	–	0	1,066	0,700	1,076	0,696
iARIMA	–	1	1,083	0,706	1,085	0,702
MLP <sup>I</sup>	–	0	1,178	0,700	1,179	0,701
Híbrido <sup>I</sup>	–	0	1,007	0,685	0,999	0,683
AES <sub>s,i</sub>	–	1	1,300	0,782	1,305	0,786
AED <sub>s,i</sub>	–	1	1,324	0,797	1,329	0,792
HW <sub>s,i</sub>	–	0	1,352	0,777	1,351	0,774
ARMA <sub>s,i</sub>	–	0	1,222	0,753	1,227	0,749
ARIMA <sub>s,i</sub>	–	0	1,217	0,744	1,215	0,738
MLP <sub>s,i</sub>	–	0	1,209	0,747	1,204	0,745
Híbrido <sub>s,i</sub>	–	0	1,185	0,736	1,186	0,737

Figura A.24: Valores das medidas de precisão obtidos para a STI da Google; 20 passos-à-frente.

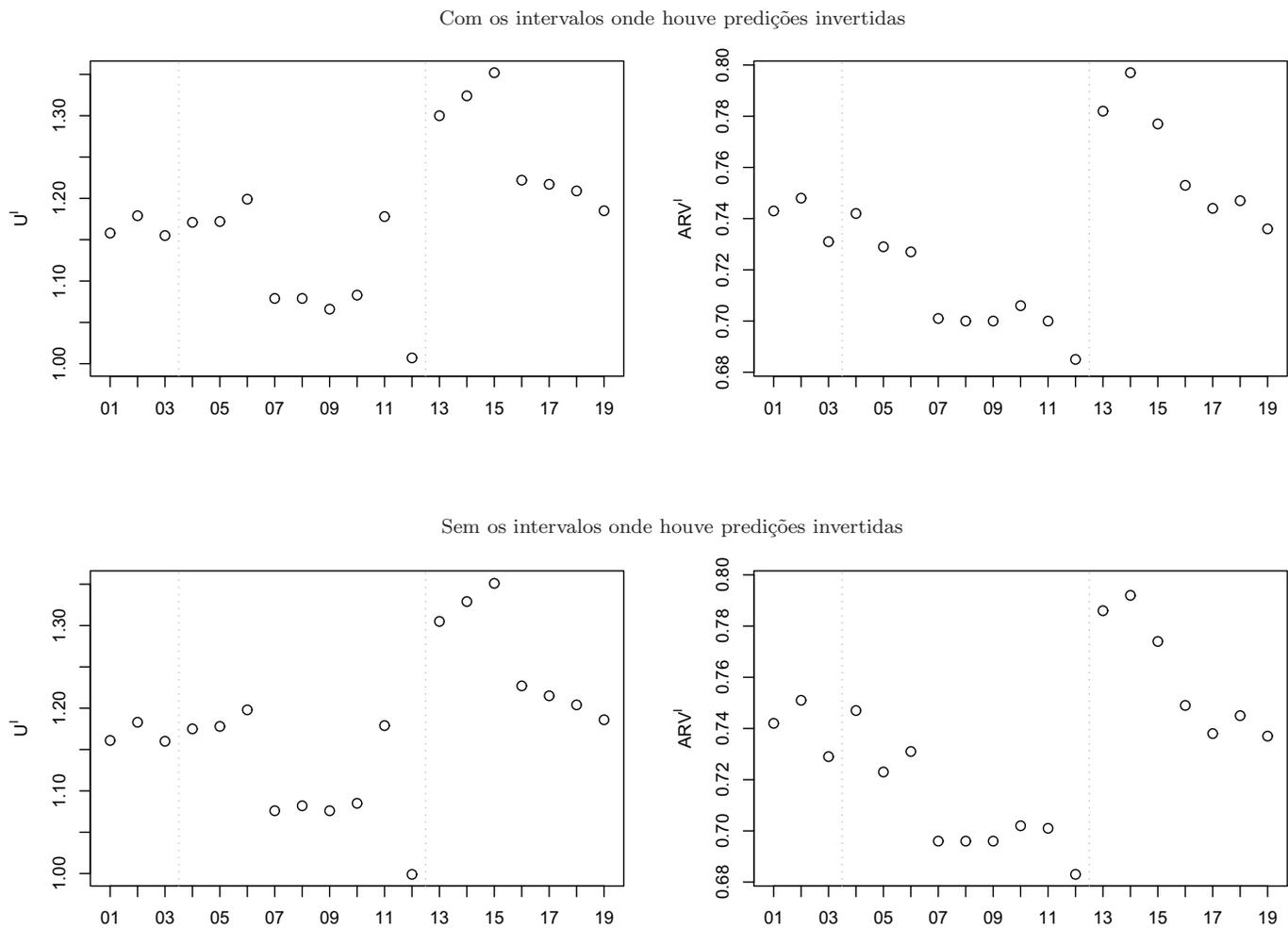


Tabela A.25: Resultados obtidos para a STI de preços de ações da Gol; conjunto de treinamento.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>		0	0,903	0,624	0,919	0,629
MLP <sub>c,r</sub>	•	1	0,922	0,650	0,920	0,646
Híbrido <sub>c,r</sub>		0	0,897	0,613	0,907	0,614
AES <sup>I</sup>	•	21	0,918	0,612	0,920	0,617
AED <sup>I</sup>		11	0,908	0,602	0,905	0,605
HW <sup>I</sup>		9	0,911	0,609	0,913	0,612
iARMA-irrestrito		9	0,873	0,548	0,872	0,543
iARMA-restrito		0	0,868	0,545	0,866	0,537
iARMA-combinado		0	0,866	0,543	0,872	0,543
iARIMA		13	0,884	0,548	0,879	0,549
MLP <sup>I</sup>		11	0,875	0,562	0,867	0,552
Híbrido <sup>I</sup>		0	0,829	0,530	0,830	0,521
AES <sub>s,i</sub>		23	0,997	0,665	0,983	0,664
AED <sub>s,i</sub>		6	0,987	0,659	0,981	0,661
HW <sub>s,i</sub>		4	1,003	0,663	0,994	0,650
ARMA <sub>s,i</sub>	•	19	0,924	0,623	0,933	0,620
ARIMA <sub>s,i</sub>		12	0,917	0,630	0,914	0,621
MLP <sub>s,i</sub>		17	0,906	0,622	0,920	0,633
Híbrido <sub>s,i</sub>		17	0,911	0,617	0,902	0,624

Figura A.25: Valores das medidas de precisão obtidos para a STI da Gol; conjunto de treinamento.

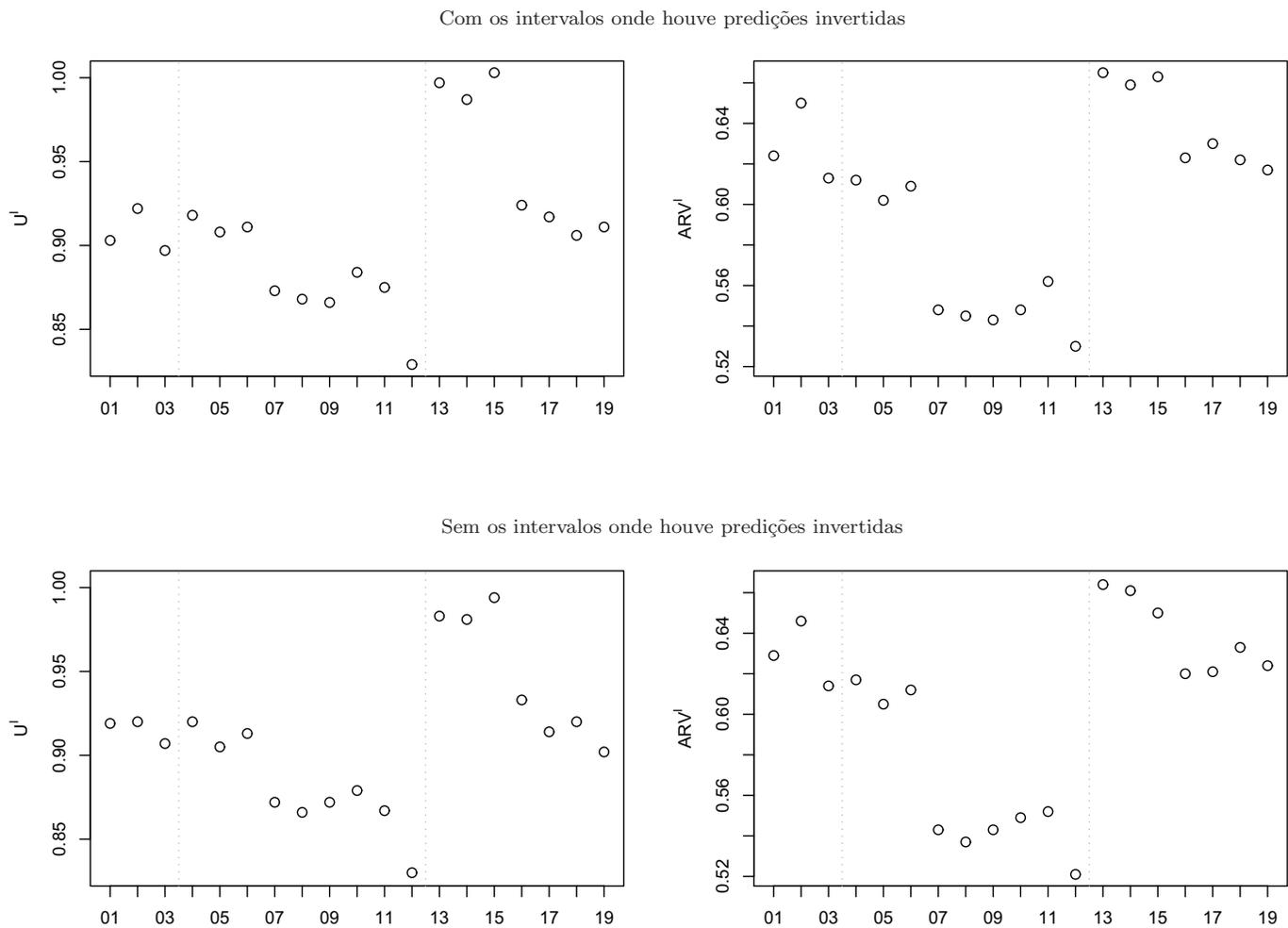


Tabela A.26: Resultados obtidos para a STI de preços de ações da Gol; 5 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,940	0,644	0,941	0,645
MLP <sub>c,r</sub>	–	0	0,944	0,688	0,945	0,689
Híbrido <sub>c,r</sub>	–	0	0,932	0,638	0,932	0,638
AES <sup>I</sup>	–	0	0,938	0,752	0,939	0,753
AED <sup>I</sup>	–	0	0,933	0,631	0,934	0,632
HW <sup>I</sup>	–	1	0,940	0,628	0,943	0,630
iARMA-irrestrito	–	0	0,915	0,594	0,915	0,595
iARMA-restrito	–	0	0,926	0,590	0,927	0,591
iARMA-combinado	–	0	0,915	0,594	0,915	0,595
iARIMA	–	0	0,936	0,621	0,937	0,622
MLP <sup>I</sup>	–	0	0,913	0,627	0,914	0,622
Híbrido <sup>I</sup>	–	0	0,882	0,544	0,883	0,545
AES <sub>s,i</sub>	–	0	1,040	0,800	1,041	0,801
AED <sub>s,i</sub>	–	0	1,021	0,782	1,022	0,783
HW <sub>s,i</sub>	–	0	1,047	0,792	1,047	0,793
ARMA <sub>s,i</sub>	–	0	0,958	0,699	0,958	0,700
ARIMA <sub>s,i</sub>	–	0	0,961	0,670	0,962	0,671
MLP <sub>s,i</sub>	–	0	0,948	0,689	0,949	0,690
Híbrido <sub>s,i</sub>	–	0	0,930	0,671	0,931	0,672

Figura A.26: Valores das medidas de precisão obtidos para a STI da Gol; 5 passos-à-frente.

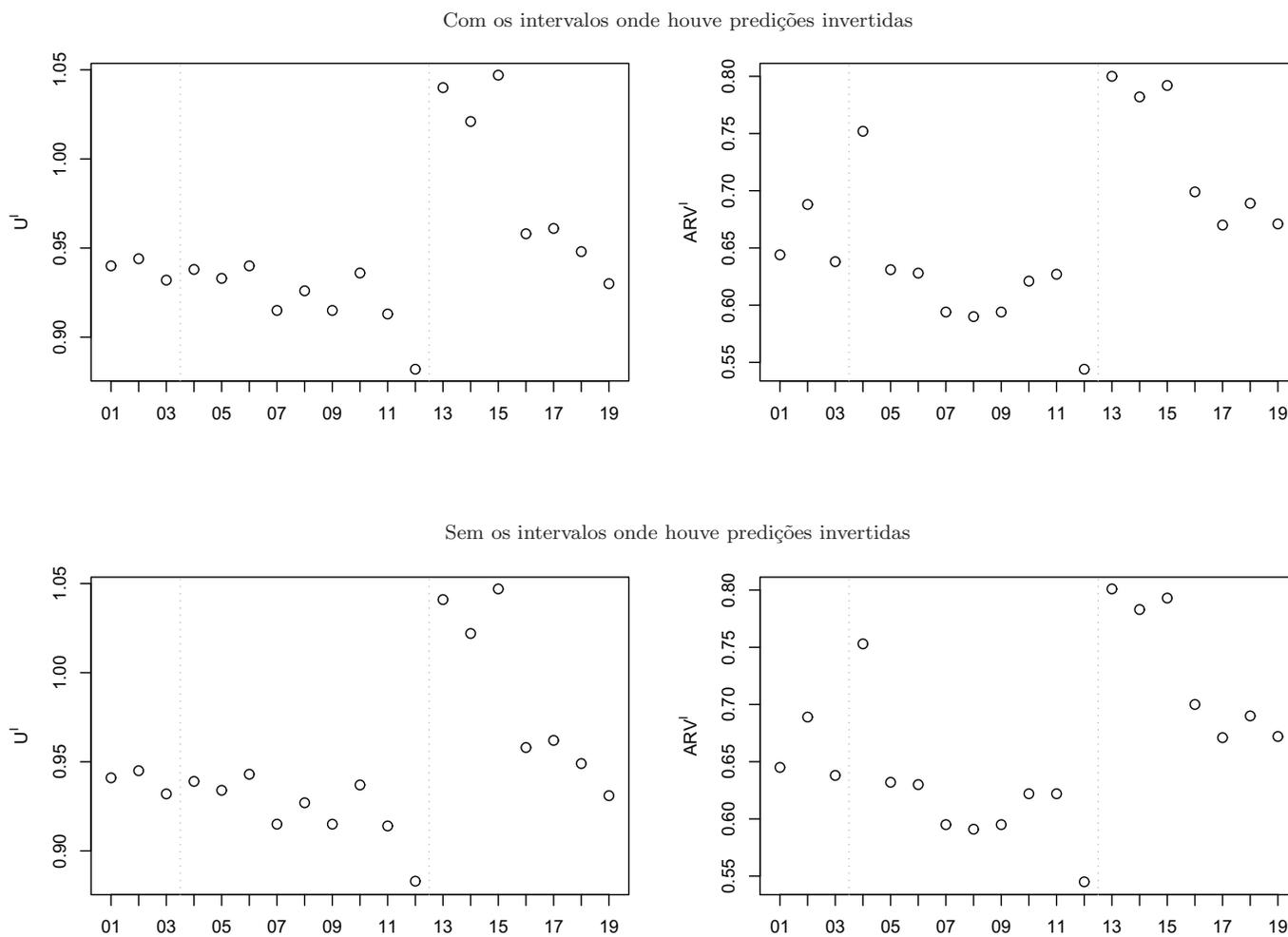


Tabela A.27: Resultados obtidos para a STI de preços de ações da Gol; 20 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	U <sup>I</sup>	ARV <sup>I</sup>	U <sup>I</sup>	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	1,160	0,698	1,160	0,696
MLP <sub>c,r</sub>	–	0	1,142	0,704	1,144	0,707
Híbrido <sub>c,r</sub>	–	0	1,121	0,689	1,123	0,685
AES <sup>I</sup>	–	1	1,111	0,708	1,116	0,703
AED <sup>I</sup>	–	0	1,112	0,676	1,115	0,680
HW <sup>I</sup>	–	1	1,129	0,690	1,128	0,695
iARMA-irrestrito	–	0	1,011	0,586	1,015	0,582
iARMA-restrito	–	0	1,021	0,589	1,023	0,579
iARMA-combinado	–	0	1,011	0,586	1,015	0,582
iARIMA	–	2	1,045	0,604	1,046	0,600
MLP <sup>I</sup>	–	0	1,068	0,608	1,068	0,611
Híbrido <sup>I</sup>	–	0	1,007	0,548	1,001	0,547
AES <sub>s,i</sub>	–	0	1,300	0,781	1,302	0,776
AED <sub>s,i</sub>	–	0	1,324	0,773	1,326	0,768
HW <sub>s,i</sub>	–	0	1,312	0,762	1,310	0,759
ARMA <sub>s,i</sub>	–	1	1,130	0,727	1,133	0,724
ARIMA <sub>s,i</sub>	–	0	1,129	0,697	1,131	0,692
MLP <sub>s,i</sub>	–	1	1,189	0,692	1,189	0,690
Híbrido <sub>s,i</sub>	–	0	1,099	0,653	1,099	0,657

Figura A.27: Valores das medidas de precisão obtidos para a STI da Gol; 20 passos-à-frente.

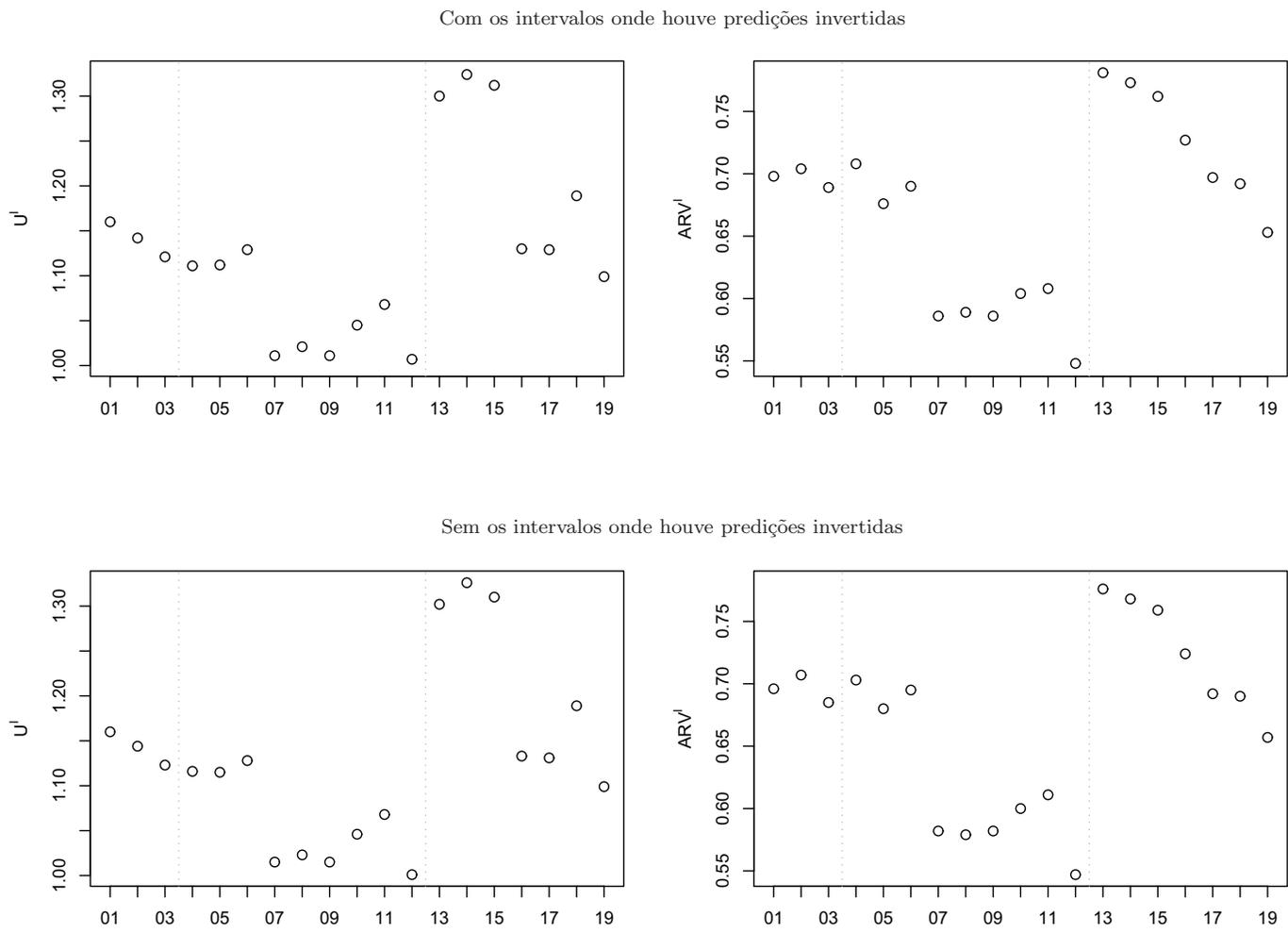


Tabela A.28: Resultados obtidos para a STI de temperaturas no Rio de Janeiro; conjunto de treinamento.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$		0	0,701	0,435	0,703	0,436
$MLP_{c,r}$		0	0,709	0,448	0,712	0,451
Híbrido $_{c,r}$		0	0,698	0,426	0,698	0,427
$AES^I$	•	0	0,682	0,422	0,685	0,425
$AED^I$		0	0,677	0,417	0,679	0,418
$HW^I$		0	0,660	0,414	0,663	0,417
iARMA-irrestrito		6	0,619	0,329	0,621	0,331
iARMA-restrito		0	0,615	0,327	0,617	0,328
iARMA-combinado		0	0,614	0,322	0,621	0,331
iARIMA		0	0,623	0,350	0,627	0,355
$MLP^I$		0	0,628	0,337	0,632	0,343
Híbrido $^I$		0	0,609	0,322	0,612	0,325
$AES_{s,i}$		0	0,706	0,500	0,700	0,493
$AED_{s,i}$		0	0,701	0,505	0,702	0,505
$HW_{s,i}$		0	0,692	0,473	0,693	0,474
$ARMA_{s,i}$		3	0,683	0,423	0,683	0,424
$ARIMA_{s,i}$		1	0,690	0,432	0,691	0,433
$MLP_{s,i}$		0	0,688	0,422	0,689	0,423
Híbrido $_{s,i}$		0	0,679	0,419	0,675	0,418

Figura A.28: Valores das medidas de precisão obtidos para a STI do Rio de Janeiro; conjunto de treinamento.

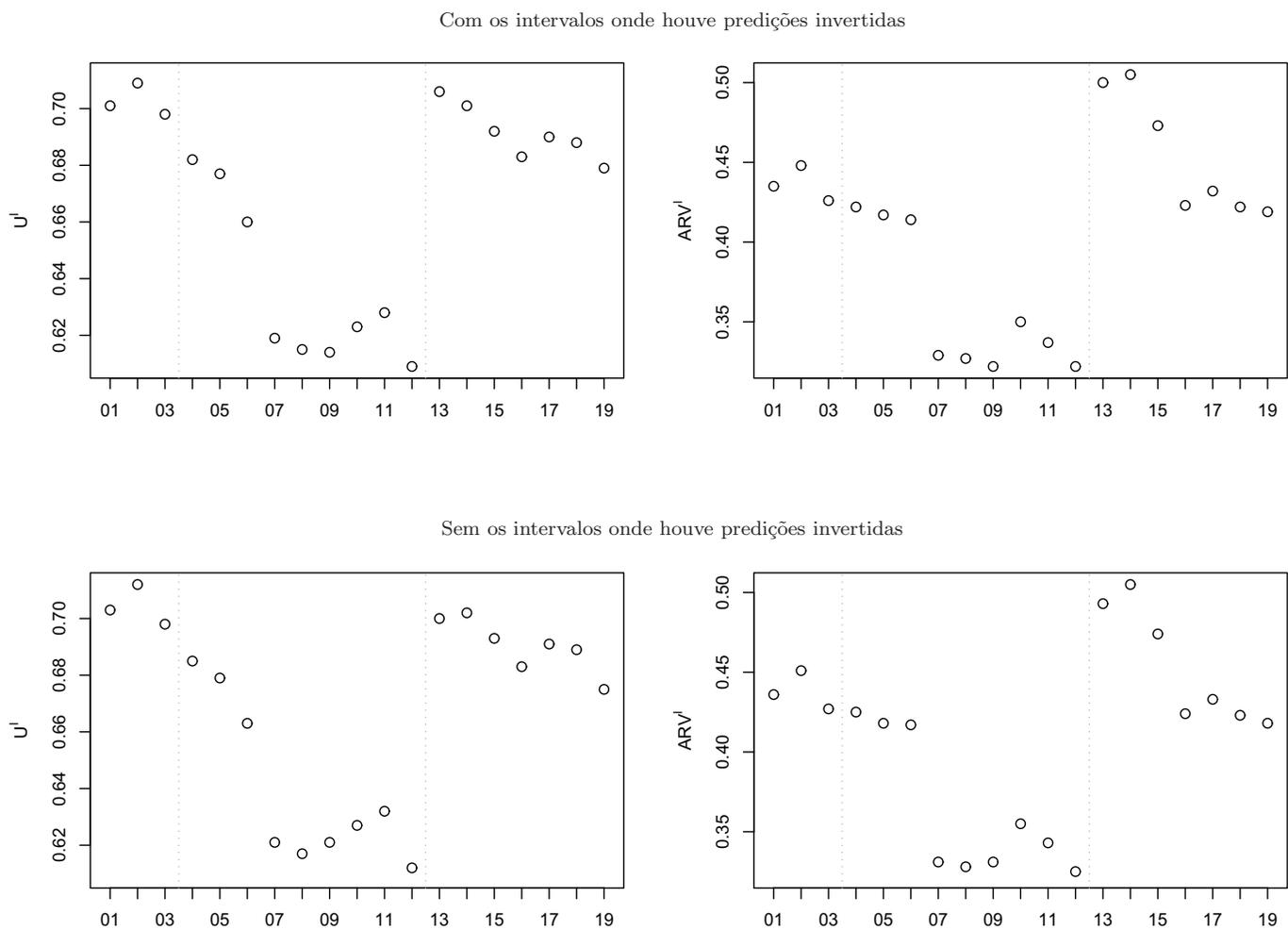


Tabela A.29: Resultados obtidos para a STI de temperaturas no Rio de Janeiro; 5 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,723	0,567	0,723	0,567
MLP <sub>c,r</sub>	–	0	0,736	0,594	0,736	0,594
Híbrido <sub>c,r</sub>	–	0	0,717	0,573	0,717	0,573
AES <sup>I</sup>	–	0	0,711	0,473	0,711	0,473
AED <sup>I</sup>	–	0	0,704	0,468	0,704	0,468
HW <sup>I</sup>	–	0	0,679	0,445	0,679	0,445
iARMA-irrestrito	–	0	0,653	0,426	0,653	0,426
iARMA-restrito	–	0	0,657	0,423	0,657	0,423
iARMA-combinado	–	0	0,653	0,426	0,653	0,426
iARIMA	–	0	0,668	0,483	0,668	0,483
MLP <sup>I</sup>	–	0	0,643	0,407	0,643	0,407
Híbrido <sup>I</sup>	–	0	0,636	0,402	0,636	0,402
AES <sub>s,i</sub>	–	0	0,803	0,657	0,803	0,657
AED <sub>s,i</sub>	–	0	0,800	0,697	0,800	0,697
HW <sub>s,i</sub>	–	0	0,770	0,643	0,770	0,643
ARMA <sub>s,i</sub>	–	0	0,734	0,541	0,734	0,541
ARIMA <sub>s,i</sub>	–	0	0,747	0,609	0,747	0,609
MLP <sub>s,i</sub>	–	0	0,713	0,545	0,713	0,545
Híbrido <sub>s,i</sub>	–	0	0,698	0,553	0,698	0,553

Figura A.29: Valores das medidas de precisão obtidos para a STI do Rio de Janeiro; 5 passos-à-frente.

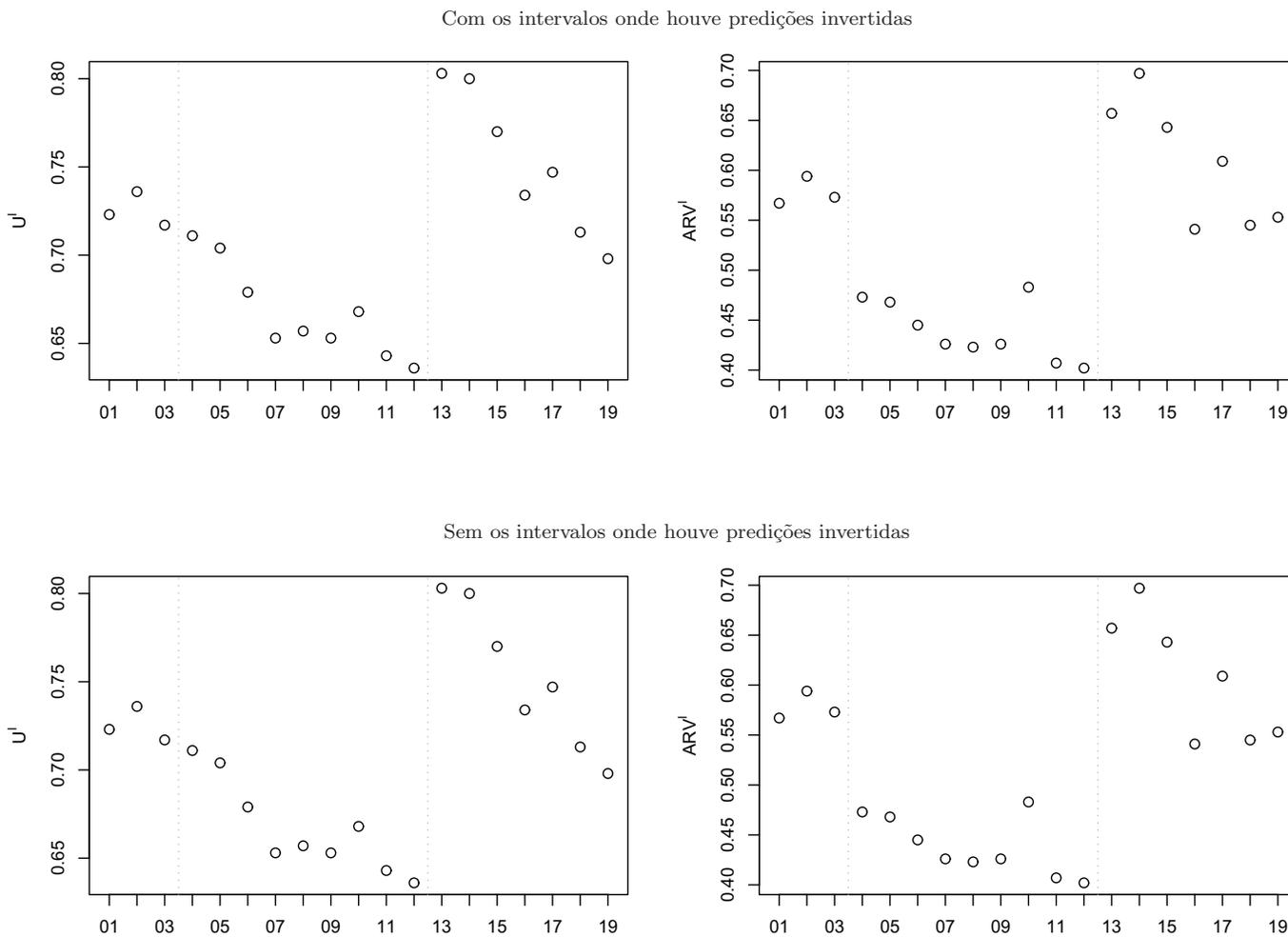


Tabela A.30: Resultados obtidos para a STI de temperaturas no Rio de Janeiro; 20 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$	–	0	0,840	0,686	0,839	0,684
$MLP_{c,r}$	–	0	0,843	0,715	0,842	0,715
Híbrido $_{c,r}$	–	0	0,836	0,677	0,837	0,674
$AES^I$	–	0	0,807	0,544	0,806	0,544
$AED^I$	–	0	0,803	0,542	0,803	0,542
$HW^I$	–	0	0,794	0,512	0,794	0,512
iARMA-irrestrito	–	1	0,777	0,507	0,772	0,505
iARMA-restrito	–	0	0,774	0,499	0,770	0,498
iARMA-combinado	–	0	0,775	0,507	0,774	0,505
iARIMA	–	0	0,768	0,537	0,769	0,538
$MLP^I$	–	0	0,760	0,477	0,762	0,478
Híbrido $^I$	–	0	0,761	0,467	0,766	0,468
$AES_{s,i}$	–	0	0,826	0,693	0,822	0,682
$AED_{s,i}$	–	0	0,838	0,727	0,835	0,724
$HW_{s,i}$	–	0	0,819	0,681	0,817	0,678
$ARMA_{s,i}$	–	1	0,828	0,630	0,825	0,628
$ARIMA_{s,i}$	–	0	0,825	0,676	0,822	0,673
$MLP_{s,i}$	–	0	0,810	0,616	0,811	0,614
Híbrido $_{s,i}$	–	0	0,814	0,620	0,814	0,621

Figura A.30: Valores das medidas de precisão obtidos para a STI do Rio de Janeiro; 20 passos-à-frente.

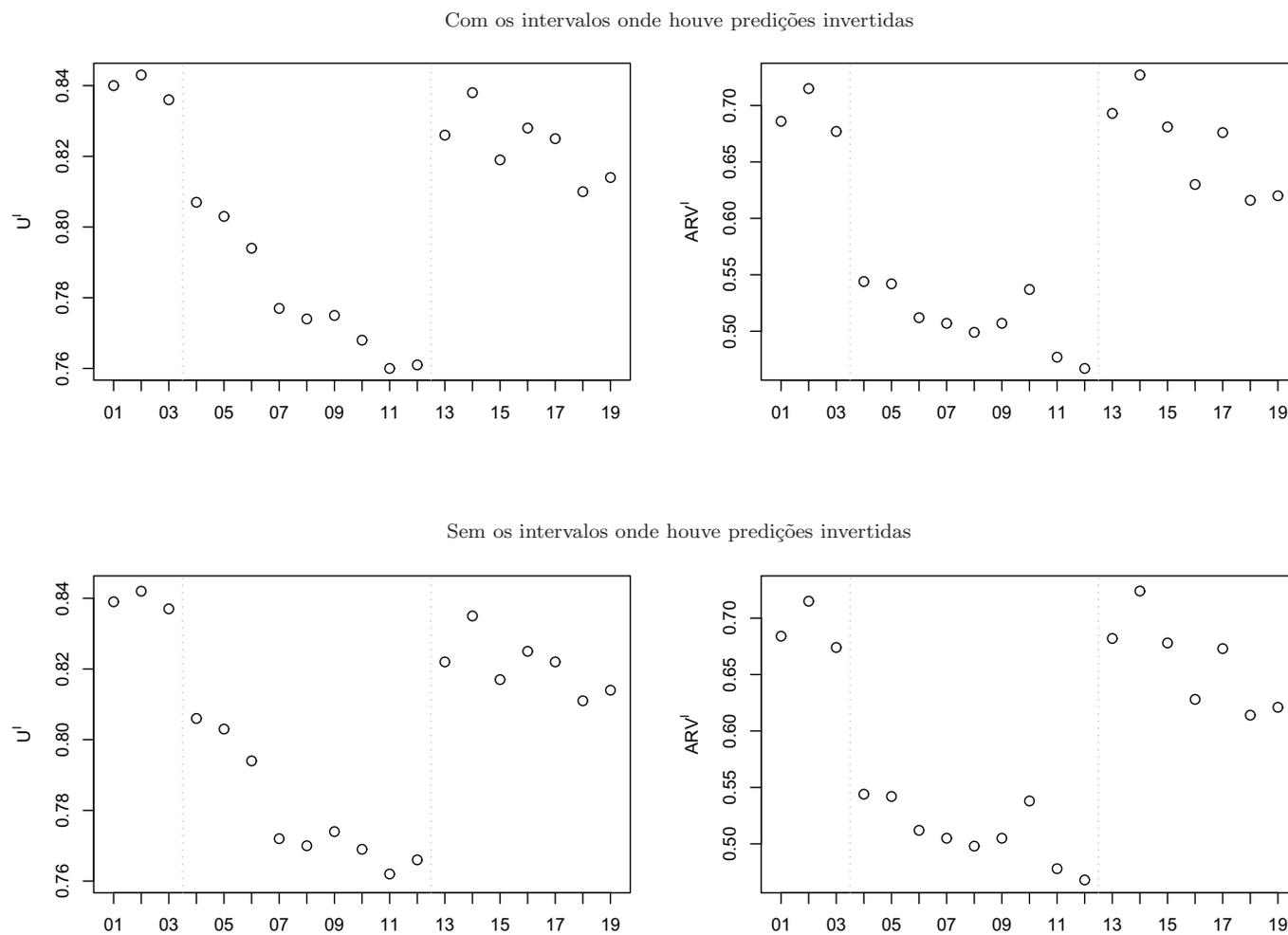


Tabela A.31: Resultados obtidos para a STI de temperaturas em Nova Iorque; conjunto de treinamento.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>		0	0,529	0,319	0,523	0,322
MLP <sub>c,r</sub>		0	0,530	0,296	0,535	0,301
Híbrido <sub>c,r</sub>		0	0,520	0,290	0,523	0,295
AES <sup>I</sup>	•	0	0,534	0,282	0,538	0,281
AED <sup>I</sup>		7	0,531	0,277	0,536	0,279
HW <sup>I</sup>		0	0,511	0,217	0,512	0,216
iARMA-irrestrito		12	0,517	0,218	0,521	0,223
iARMA-restrito		0	0,514	0,220	0,516	0,222
iARMA-combinado		0	0,514	0,213	0,521	0,223
iARIMA		8	0,513	0,224	0,519	0,231
MLP <sup>I</sup>		0	0,515	0,230	0,521	0,237
Híbrido <sup>I</sup>		0	0,509	0,208	0,510	0,213
AES <sub>s,i</sub>	•	0	0,544	0,354	0,543	0,349
AED <sub>s,i</sub>		5	0,547	0,359	0,546	0,360
HW <sub>s,i</sub>		0	0,531	0,320	0,535	0,323
ARMA <sub>s,i</sub>		4	0,528	0,295	0,527	0,298
ARIMA <sub>s,i</sub>		3	0,523	0,289	0,521	0,281
MLP <sub>s,i</sub>		0	0,516	0,286	0,519	0,290
Híbrido <sub>s,i</sub>		0	0,518	0,270	0,514	0,275

Figura A.31: Valores das medidas de precisão obtidos para a STI de Nova Iorque; conjunto de treinamento.

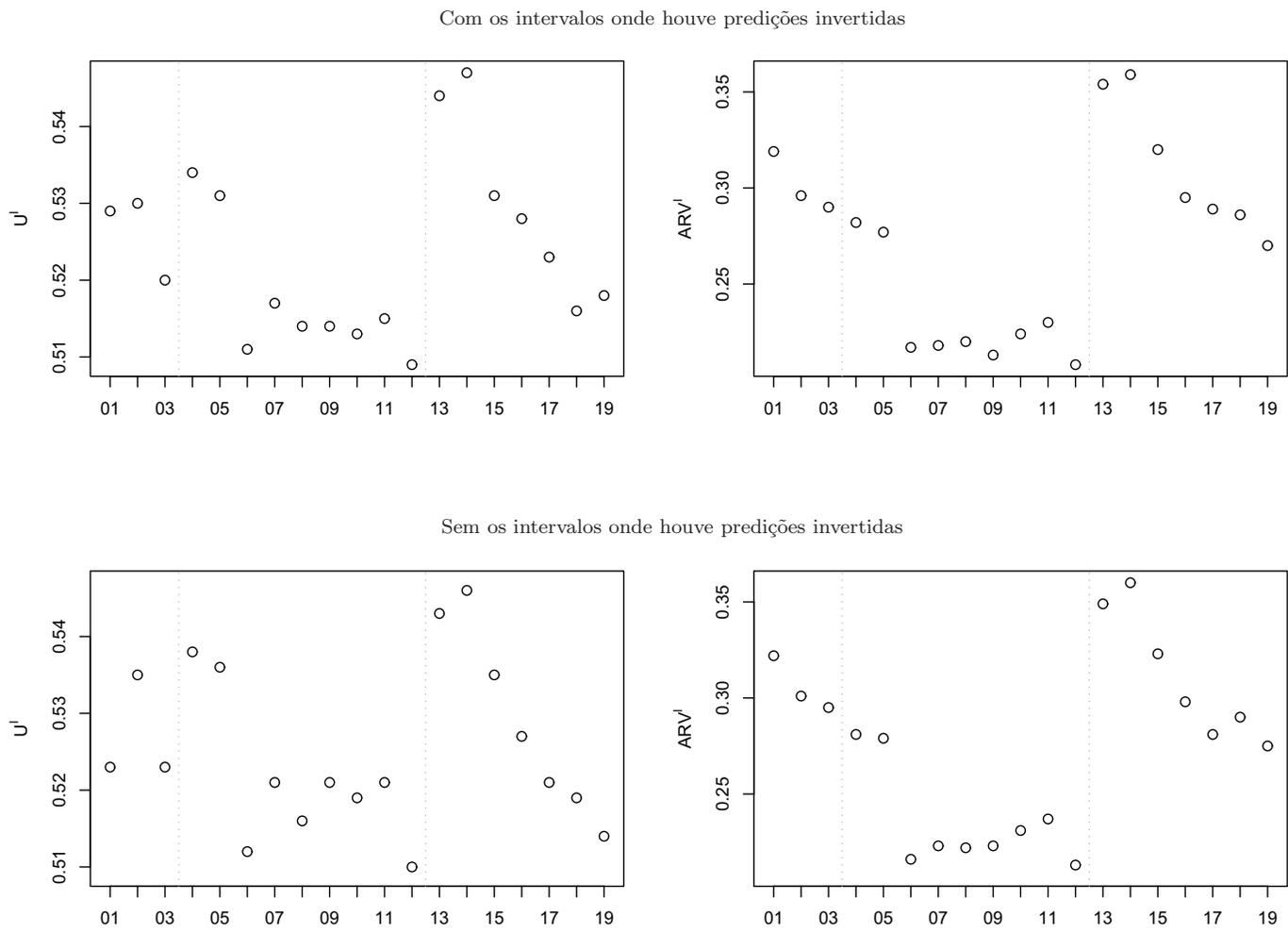


Tabela A.32: Resultados obtidos para a STI de temperaturas em Nova Iorque; 5 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	ARV <sup>I</sup>	$U^I$	ARV <sup>I</sup>
ARMA <sub>c,r</sub>	–	0	0,595	0,375	0,595	0,375
MLP <sub>c,r</sub>	–	0	0,592	0,372	0,592	0,372
Híbrido <sub>c,r</sub>	–	0	0,584	0,368	0,584	0,368
AES <sup>I</sup>	–	0	0,575	0,357	0,575	0,357
AED <sup>I</sup>	–	0	0,577	0,361	0,577	0,361
HW <sup>I</sup>	–	0	0,559	0,342	0,559	0,342
iARMA-irrestrito	–	0	0,549	0,316	0,549	0,316
iARMA-restrito	–	0	0,552	0,317	0,552	0,317
iARMA-combinado	–	0	0,549	0,316	0,549	0,316
iARIMA	–	0	0,558	0,322	0,558	0,322
MLP <sup>I</sup>	–	0	0,538	0,311	0,538	0,311
Híbrido <sup>I</sup>	–	0	0,541	0,313	0,541	0,313
AES <sub>s,i</sub>	–	0	0,617	0,378	0,617	0,378
AED <sub>s,i</sub>	–	0	0,620	0,386	0,620	0,386
HW <sub>s,i</sub>	–	0	0,592	0,360	0,592	0,360
ARMA <sub>s,i</sub>	–	0	0,589	0,351	0,589	0,351
ARIMA <sub>s,i</sub>	–	0	0,571	0,335	0,571	0,335
MLP <sub>s,i</sub>	–	0	0,569	0,342	0,569	0,342
Híbrido <sub>s,i</sub>	–	0	0,563	0,328	0,563	0,328

Figura A.32: Valores das medidas de precisão obtidos para a STI de Nova Iorque; 5 passos-à-frente.

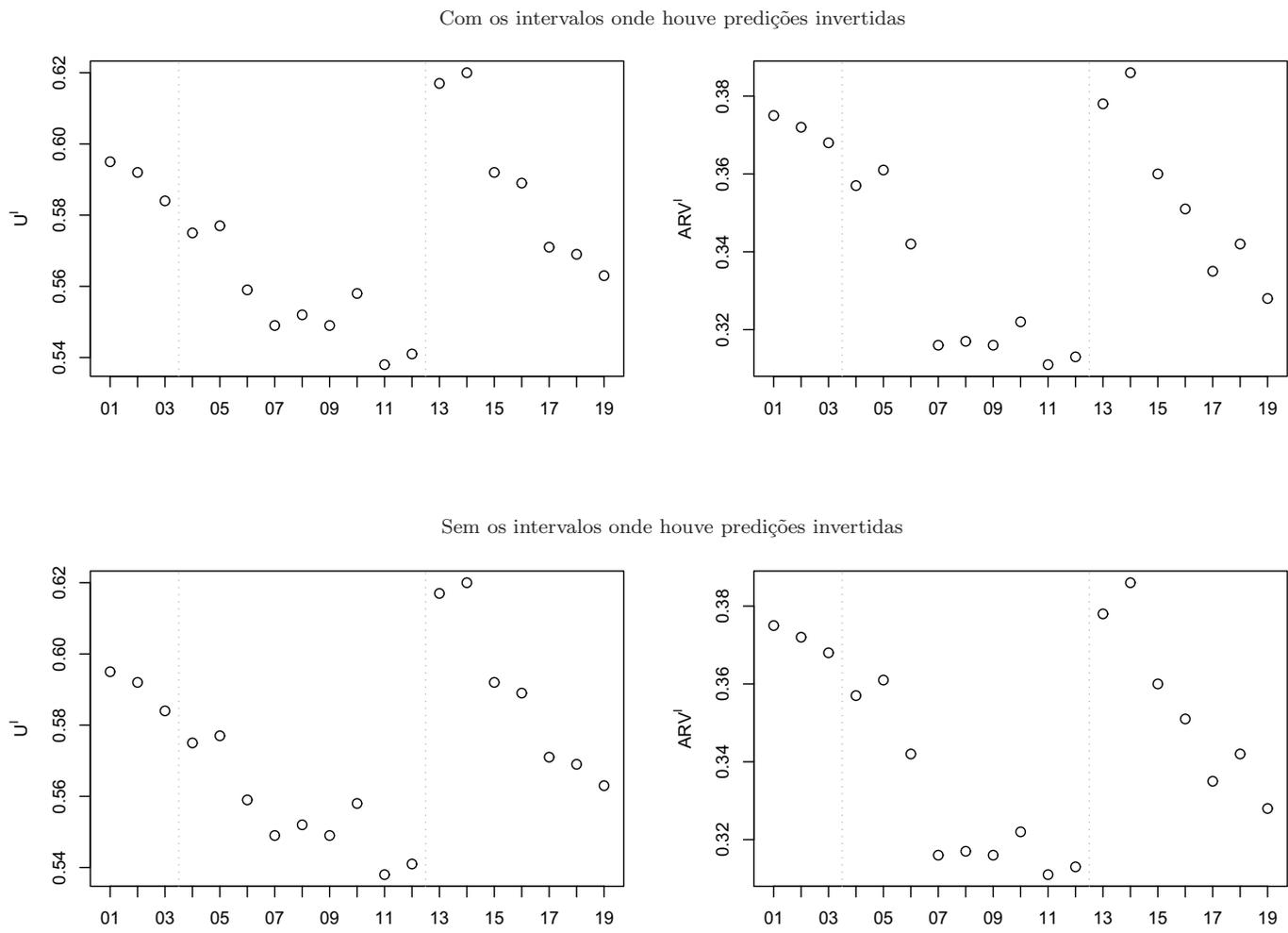
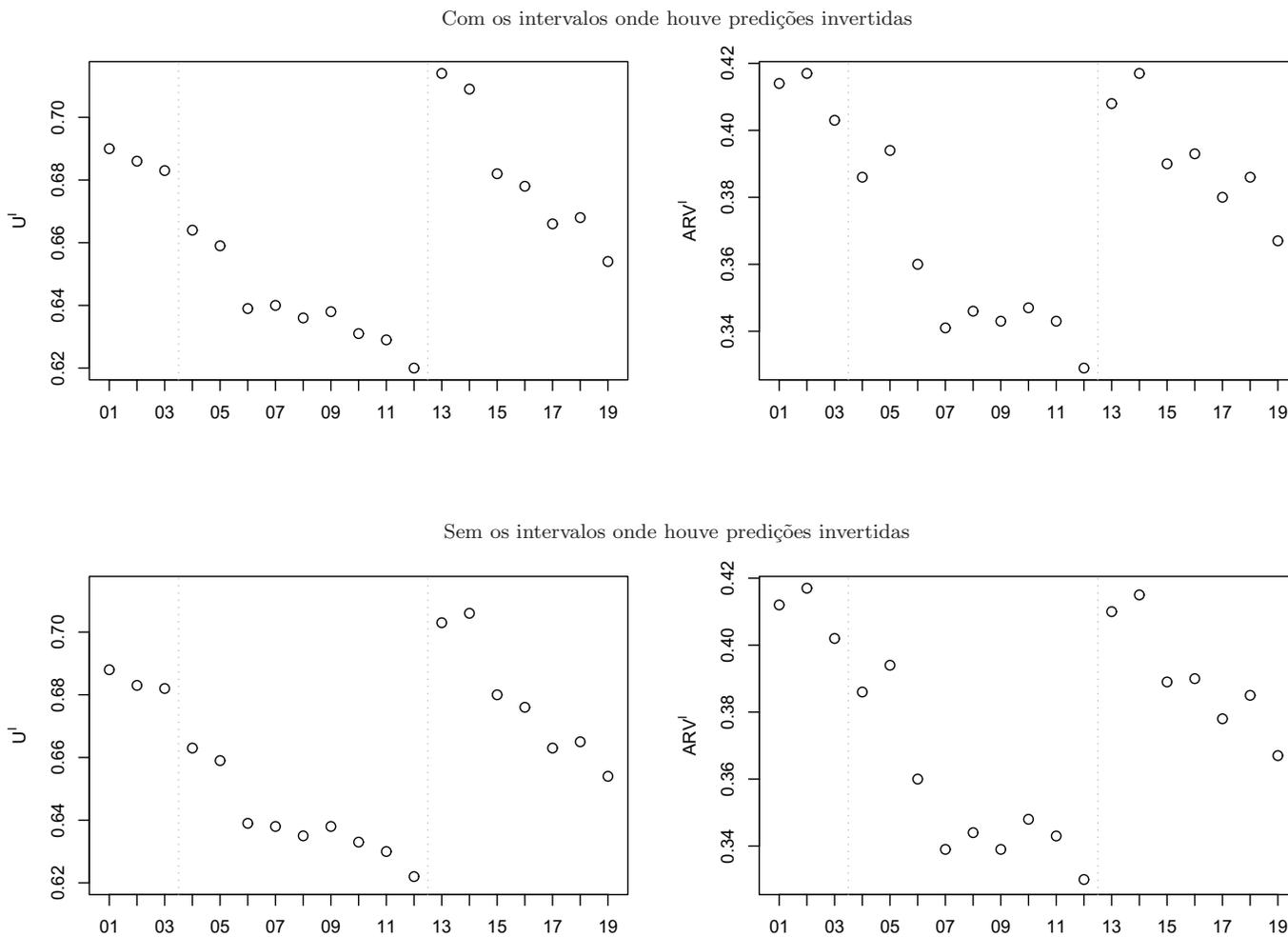


Tabela A.33: Resultados obtidos para a STI de temperaturas em Nova Iorque; 20 passos-à-frente.

Modelo	Rejeita $H_0$	Com predições invertidas			Sem predições invertidas	
		Pred. inv.	$U^I$	$ARV^I$	$U^I$	$ARV^I$
$ARMA_{c,r}$	–	0	0,690	0,414	0,688	0,412
$MLP_{c,r}$	–	0	0,686	0,417	0,683	0,417
Híbrido $_{c,r}$	–	0	0,683	0,403	0,682	0,402
$AES^I$	–	0	0,664	0,386	0,663	0,386
$AED^I$	–	0	0,659	0,394	0,659	0,394
$HW^I$	–	0	0,639	0,360	0,639	0,360
iARMA-irrestrito	–	2	0,640	0,341	0,638	0,339
iARMA-restrito	–	0	0,636	0,346	0,635	0,344
iARMA-combinado	–	0	0,638	0,343	0,638	0,339
iARIMA	–	0	0,631	0,347	0,633	0,348
$MLP^I$	–	0	0,629	0,343	0,630	0,343
Híbrido $^I$	–	0	0,620	0,329	0,622	0,330
$AES_{s,i}$	–	0	0,714	0,408	0,703	0,410
$AED_{s,i}$	–	0	0,709	0,417	0,706	0,415
$HW_{s,i}$	–	2	0,682	0,390	0,680	0,389
$ARMA_{s,i}$	–	1	0,678	0,393	0,676	0,390
$ARIMA_{s,i}$	–	0	0,666	0,380	0,663	0,378
$MLP_{s,i}$	–	0	0,668	0,386	0,665	0,385
Híbrido $_{s,i}$	–	0	0,654	0,367	0,654	0,367

Figura A.33: Valores das medidas de precisão obtidos para a STI de Nova Iorque; 20 passos-à-frente.



# Apêndice B

## Artigos publicados e submetidos

Esta seção apresenta a lista de artigos publicados e submetidos durante o período do doutorado e relacionados ao tema de pesquisa desta tese, i.e., unindo a análise de séries temporais com a análise de dados simbólicos.

1. “ARMA-based modeling framework to interval-valued time series forecasting”. Submetido em 2010 [com Francisco A. T. De Carvalho].
2. “Combining exponential smoothing and neural networks for interval-valued time series modelling”. Submetido em 2010 [com Francisco A. T. De Carvalho].
3. “Holt’s exponential smoothing and neural network models for forecasting interval-valued time series”. Aceito para publicação na revista *International Journal of Forecasting* em 2010 [com Francisco A. T. De Carvalho].
4. “Forecasting models for interval-valued time series”. Publicado na *Neurocomputing*, 71:3344–3352, 2008 [com Francisco A. T. De Carvalho e Teresa B. Ludermir].
5. “Neural networks and exponential smoothing models for symbolic interval time series processing – applications in stock market”. Publicado no *Proceedings of the Eighth International Conference on Hybrid Intelligent Systems, HIS’2008*, 326–331, 2008 [com Francisco A. T. De Carvalho].
6. “Fitting a least absolute deviation regression model on interval-valued data”. Publicado no *Advances in Artificial Intelligence, SBIA’2008*, 5249:207–216, 2008 [com Francisco A. T. De Carvalho].

7. “A hybrid model for symbolic interval time series forecasting”. Publicado na *Lecture Notes in Computer Science – ICONIP’2006*, 4233:934–941, 2006 [com Francisco A. T. De Carvalho e Teresa B. Ludermir].
8. “Hybrid model with dynamic architecture for forecasting time series”. Publicado no *Proceedings of the International Joint Conference on Neural Networks, 2006, IJCNN’06*, 3742–3747, 2006 [com Gecynalda S. S. Gomes, Francisco A. T. De Carvalho, Teresa B. Ludermir e Aluisio Araujo].
9. “Symbolic interval time series forecasting using a hybrid model”. Publicado no *Proceedings of the Ninth Brazilian Symposium on Neural Networks, SBRN ’06*, IEEE Computer Society, 202–207, 2006 [com Francisco A. T. De Carvalho e Teresa B. Ludermir].

# Referências Bibliográficas

- [1] Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.
- [2] Arroyo, J. e Maté, C. Introducing interval time series: accuracy measures. In *COMPSTAT 2006: 17th Conference of Int. Association for Statistical Computing and International Federation of Classification Societies 2006 Conference: Data Science and Classification*, pages 1139–1146. Roma, Italy, 2006.
- [3] Arroyo J. e Maté, C. Forecasting histogram time series with k-nearest neighbours methods. *International Journal of Forecasting*, 25:192–207, 2009.
- [4] Arroyo, J, Roque, A., Maté, C. e Sarabia, A. Exponential smoothing methods for interval time series. In *Proceedings of the First European Symposium on Time Prediction*, pages 231–240. Otaniemi, Finland, 2007.
- [5] Barrodale, I. e Roberts F. D. K. Solution of an overdetermined system of equation in the  $\ell_1$  norm. *Communications of the Association for Computing Machinery*, 17:319–320, 1974.
- [6] Billard, L. Dependencies in bivariate interval-valued symbolic data. In *D. Banks, L. House, F.R. McMorris, P. Arabie e W. Gaul (eds), Classification, Clustering, and Data Mining Applications*, pages 319–354. Berlin: Springer-Verlag, 2004.
- [7] Billard, L. e Diday, E. Regression analysis for interval-valued data. In *Proceedings of the Seventh Conference of the International Federation of Classification Societies (IFCS'00)*, pages 369–374. Springer-Verlag, Belgium, 2000.
- [8] Billard, L. e Diday, E. Symbolic regression analysis. In *K. Jajuga, A. Sokolowski e H.-H. Bock (Eds.), Classification, Clustering, and Data Analysis: Recent Advances and Applications*, pages 281–288. Berlin: Springer-Verlag, 2002.
- [9] Billard, L. e Diday, E. From the statistics of data to the statistics of knowledge: symbolic data analysis. *Journal of American Statistical Association*, 98(462):470–487, 2003.
- [10] Billard, L. e Diday, E. *Symbolic Data Analysis: Conceptual Statistics and Data Mining*. Chichester: Wiley, 2006.
- [11] Bock, H. H. e Diday, E. *Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data*. Springer, Berlin Heidelberg, 2000.

- [12] Box, G. E. P. e Jenkins, G. M. *Time Series Analysis, Forecasting and Control*, Edição revisada. San Francisco: Holden Day, 1976.
- [13] Brito, P. Hierarchical and pyramidal clustering for symbolic data. *Journal of the Japanese Society of Computational Statistics*, 15(2):231–244, 2002.
- [14] Brockwell, P. J. e Davis, R. A. *Introduction to Time Series and Forecasting* (2ª edição). Springer-Verlag, New York, 2002.
- [15] Brown, R. G. *Statistical forecasting for inventory control*. New York: McGraw-Hill, 1959.
- [16] Broyden, C. G. The convergence of a class of double-rank minimization algorithms 2. *Journal of the Institute of Mathematics and its Applications*, 6:222–231, 1970.
- [17] Buchinsky, M. Recent advances in quantile regression: a practical guideline for empirical research. *Journal of Human Resources*, 33:88–126, 1997.
- [18] Byrd, R. H., Lu, P., Nocedal, J. e Zhu, C. A limited memory algorithm for bound constrained optimization. *SIAM Journal Scientific Computing*, 16:1190–1208, 1995.
- [19] Cazes, P., Chouakria, A., Diday, E. e Schektman, Y. Extension de l'analyse en composantes principales à des données de type intervalle. *Revue de Statistique Appliquée*, 14:5–24, 1997.
- [20] Chambers, J. *The R Foundation for Statistical Computing - Version 2.2.0*. Bell Laboratories (Lucent Technologies), 2005.
- [21] Chouakria, A., Diday, E. e Cazes, P. An improved factorial representation of symbolic objects. In *Studies and Research, Proceedings of the Conference on Knowledge Extraction and Symbolic Data Analysis*, pages 276–289. Luxembourg: Office for Official Publications of the European Communities, 1998.
- [22] Clemen, R. T. Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5:559–583, 1989.
- [23] Cybenko, G. Approximation by superposition of sigmoidal functions. *Mathematics of Control, Signals and Systems*, 2:303–314, 1989.
- [24] Darbellay, G. A. e Slama, M. Forecasting the short-term demand for electricity: Do neural networks stand a better chance? *International Journal of Forecasting*, 16:71–83, 2000.
- [25] De Carvalho, F. A. T., Brito, P. e Bock, H-H. Dynamic clustering for interval data based on  $L_2$  distance. *Computational Statistics*, 21:231–250, 2006b.
- [26] De Carvalho, F. A. T., Csernel, M. e Lechevallier, Y. Clustering constrained symbolic data. *Pattern Recognition Letters*, 30:1037–1045, 2009a.

- [27] De Carvalho, F. A. T., De Souza, R., Chavent, M. e Lechevallier, Y. Adaptive Hausdorff distances and dynamic clustering of symbolic interval data. *Pattern Recognition Letters*, 27:167–179, 2006a.
- [28] De Souza, R. M. C. R. e De Carvalho F. A. T. Clustering of interval data based on city-block distances. *Pattern Recognition Letters*, 25:353–365, 2004.
- [29] Diday, E. The symbolic approach in clustering and related methods of data analysis. In *H.-H. Bock (Ed.), Classification and Related Methods of Data Analysis*, pages 526–537. Amsterdam: North-Holland, 1987a.
- [30] Diday, E. Inequality restrictions in regression analysis. *Première Journées Symbolique - Numérique, Université Paris IX Dauphine, December*, 1987b.
- [31] Diday, E. Probabilist, possibilist and belief objects for knowledge analysis. *Annals of Operations Research*, 55:227–276, 1995.
- [32] Diday, E. Categorization in symbolic data analysis. In *H. Cohen e C. Lefebvre (Eds.), Handbook of Categorization in Cognitive Science*. Amsterdam: Elsevier, 2005.
- [33] Diday, E. e Esposito, F. An introduction to symbolic data analysis and the SODAS software. *Intelligent Data Analysis*, 7:583–601, 2003.
- [34] Diday, E. e Noirhomme-Fraiture, M. *Symbolic Data Analysis and the SODAS Software*. New York: Wiley-Interscience, 2008.
- [35] Diebol, F. X. *Elements of forecasting*. Cincinnati: South-Western College Publishing, 1998.
- [36] Dorffner, G., Leitgeb, E. e Koller, H. Toward improving exercise ECG for detecting ischemic heart disease with recurrent and feedforward neural nets. *Neural Networks for Signal Processing IV, IEEE*, pages 499–508, 1994.
- [37] Ehlers, R. S. *Análise de Séries Temporais*. Departamento de Estatística, UFPR, 2005. Disponível em <http://www.est.ufpr.br/~ehlers/notas> (Acesso em julho de 2009).
- [38] Fletcher, R. A new approach to variable metric methods. *Computer Journal*, 13:317–322, 1970.
- [39] García-Ascanio, C. e Maté, C. Electric power demand forecasting using interval time series: A comparison between VAR and iMLP. *Energy Policy*, 38:715–725, 2010.
- [40] Gardner, E. S., Jr. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4:1–28, 1985.
- [41] Gardner, E. S., Jr. Exponential smoothing: The state of the art – Part II. *International Journal of Forecasting*, 22:637–666, 2006.

- [42] Gioia, F. e Lauro, N. C. Dependence and interdependence analysis for interval-valued variables. In *V. Batagelj, H.-H. Bock, A. Ferligoj e A. Ziberna (Eds.), Data Science and Classification*. Berlin: Springer-Verlag, 2006a.
- [43] Gioia, F. e Lauro, N. C. Principal component analysis on interval data. *Computational Statistics*, 21:343–363, 1987b.
- [44] Gioia, F. e Lauro, N. C. Basic statistical methods for interval data. *Italian Journal of Applied Statistics*, 17:75–104, 2005.
- [45] Goldfarb, D. A family of variable metric methods derived by variational means. *Mathematics of Computation*, 24:23–26, 1970.
- [46] Guo, J., Li, W. e Gao, F. Descriptive statistics of non-uniform interval symbolic data. In *Proceedings of the First ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, pages 831–834. Shanghai, China, 2009.
- [47] Han, A., Hong, Y., Lai, K. K. e Wang, S. Interval time series analysis with an application to the sterling-dollar exchange rate. *Jrl. Syst. Sci. & Complexity*, 21:558–573, 2008.
- [48] Hannan, E. J. e Kavalieris, L. A method for autoregressive-moving average estimation. *Biometrika*, 71:273–280, 1984.
- [49] Hannan, E. J. e Rissanen, J. Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, 69:81–94, 1982.
- [50] Hansen, J. V. e Nelson, R. D. Neural networks and traditional time series methods: a synergistic combination in state economic forecasts. *IEEE Transactions on Neural Networks 1997*, 8:863–873, 1997.
- [51] Haykin, S. *Neural networks: a comprehensive foundation* (2ª edição). Prentice-Hall, New Jersey, 1999.
- [52] Haykin, S. *Redes neurais: princípios e prática* (2ª edição). Bookman, Porto Alegre, 2001.
- [53] Holt, C. C. Forecasting seasonals and trends by exponentially weighted averages. *International Journal of Forecasting*, 20:5–13, 2004.
- [54] Iripino, A. 'spaghetti' PCA analysis: An extension of principal components analysis to time dependent interval data. *Pattern Recognition Letters*, 27:504–513, 2006.
- [55] Judge, G. G. e Takayama, T. Inequality restrictions in regression analysis. *Journal of the American Statistical Association*, 61:166–181, 1966.
- [56] Kaastra, I. e Boyd, M. Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 10:215–236, 1996.
- [57] Kendall, M. G., Stuart, A. e Ord, J. K. *Advanced theory of statistics* (4ª edição). Volume 3. Griffin: London, 1983.

- [58] Koreisha, S. e Pukkila, T. A generalized least squares approach for estimation of autoregressive moving average models. *Journal of Time Series Analysis*, 11:139–151, 1990.
- [59] Kotsialos, A., Papageorgiou, M. e Poulimenos, A. Long-term sales forecasting using Holt-Winters and neural network methods. *Journal of Forecasting*, 24:353–368, 2005.
- [60] Lagarias, J. C., Reeds, A., Wright, M. H. e Wright, P. E. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9:112–147, 1998.
- [61] Lawson, C. L. e Hanson, R. J. *Solving Least Squares Problem*. Prentice-Hall, New York, 1974.
- [62] Lima Neto, E. A., De Carvalho, F. A. T. e Tenório, C. P. Univariate and multivariate linear regression methods to predict interval-valued features. In *G. I. Webb e X. Yu (Eds.). AI 2004, Lectures Notes on Artificial Intelligence 3339*, pages 526–537. Berlin: Springer-Verlag, 2004.
- [63] Lima Neto, E. A. e De Carvalho, F. A. T. Centre and range method for fitting a linear regression model to symbolic interval data. *Computational Statistics and Data Analysis*, 52:1500–1515, 2008.
- [64] Lima Neto, E. A. e De Carvalho, F. A. T. Constrained linear regression models for symbolic interval-valued variables. *Computational Statistics and Data Analysis*, 54(2):333–347, 2010.
- [65] Liu, D. C. e Nocedal, J. On the limited-memory BFGS method for large scale optimization. *Mathematical Programming*, 45:503–528, 1989.
- [66] Maia, A. L. S., De Carvalho, F. A. T. e Ludermir, T. B. A hybrid model for symbolic interval time series forecasting. *Lecture Notes in Computer Science – ICONIP’2006*, 4233:934–941, 2006.
- [67] Maia, A. L. S., De Carvalho, F. A. T. e Ludermir, T. B. Forecasting models for interval-valued time series. *Neurocomputing*, 71:3344–3352, 2008.
- [68] Maia, A. L. S. e De Carvalho, F. A. T. Holt’s exponential smoothing and neural network models for forecasting interval-valued time series. *International Journal of Forecasting*, (A a aparecer).
- [69] Maia, A. L. S. e De Carvalho, F. A. T. Fitting a least absolute deviation regression model on interval-valued data. In *Proceedings of the 19th Brazilian Symposium on Artificial Intelligence, SBIA’08*, volume 5249, pages 207–216. Springer, Berlin Heidelberg, 2008.
- [70] Maia, A. L. S. e De Carvalho, F. A. T. Neural networks and exponential smoothing models for symbolic interval time series processing – applications in stock market. In *Proceedings of the Eighth International Conference on Hybrid Intelligent Systems*, pages 326–331. Montreal, Canada, 2008.

- [71] Makridakis, S., Wheelwright, S. C. e Hyndman, R. J. *Forecasting, Methods and Applications* (3ª edição). Wiley, 1998.
- [72] Mills, T. C. *The econometric modelling of financial time series*. Cambridge: Cambridge University Press, 2003.
- [73] Moore, R. E. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1966.
- [74] Morettin, P. A. e Toloi, C. M. C. *Análise de séries temporais*. São Paulo: Edgard Blücher, 2006.
- [75] Navidi, W. *Statistics for engineers and scientists*. McGraw-Hill, New York, 2006.
- [76] Nocedal, J. e Wright, S. J. *Numerical Optimization*. Springer-Verlag, 1999.
- [77] Pretchelt, L. A set of neural network benchmark problems and benchmarking rules. Technical report, University Karlsruhe, Germany, 21/94, 1994.
- [78] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2006. <http://www.R-project.org>.
- [79] Roque, A., Maté, C., Arroyo, J e Sarabia, A. iMLP: Applying multi-layer perceptrons to interval-valued data. *Neural Processing Letters*, 25:157–169, 2007.
- [80] Rumelhart, D. E., Hinton, G. E. e Williams, R. J. Learning internal representations by error propagation. *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: foundations*. MIT Press, Cambridge, MA, USA, 1:318–362, 1986.
- [81] Russell, S. e Norvig, P. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, New Jersey, 1995.
- [82] Schwarz, G. Estimating the dimension of a model. *The Annals of Statistics*, 6:461–464, 1978.
- [83] Shanno, D. F. Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation*, 24:647–657, 1970.
- [84] Sharda, R. e Patil, R. B. Neural networks as forecasting experts: an empirical test. In *Proceedings of the International Joint Conference on Neural Networks*, volume 2, pages 491–494. Washington, D.C., 1990.
- [85] Smith, R.W. Visual hypothesis testing with confidence intervals. In *Proceedings of the Twenty-second Annual SAS User Group International Conference*, pages 1252–1257. San Diego, California, 1997.
- [86] Su, C. T., Tong, L. I. e Leou, C. M. Combination of time series and neural network for reliability forecasting modeling. *J. Chin. Inst. Ind. Eng.*, 14(4):419–429, 1997.

- [87] Taskaya-Temizel, T. e Ahmad, K. Are ARIMA neural network hybrids better than single models? In *Proceedings of the International Joint Conference on Neural Networks 2005*. Montreal, Canada, 2005.
- [88] Tian, J., Juhola, M. e Gronfors, T. AR parameter estimation by a feedback neural network. *Computational Statistics and Data Analysis*, 25:17–24, 1997.
- [89] Tseng, F. M., Yu, H. C. e Tzeng, G. H. Combining neural network model with seasonal time series ARIMA model. *Technological Forecasting and Social Change*, 69:71–87, 2002.
- [90] Venables, W. N. e Ripley, B. D. *Modern Applied Statistics with S* (4ª edição). New York: Springer-Verlag, 2002.
- [91] Wedding, D. K. II e Cios, K. J. Time series forecasting by combining RBF networks, certainty factors, and the Box–Jenkins model. *Neurocomputing*, 10(3):149–168, 1996.
- [92] Winters, P. R. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6:324–342, 1960.
- [93] Zadeh, L. A. Fuzzy sets. *Information and Control*, 8:338–353, 1978.
- [94] Zhang, G. Time series forecasting using a hybrid ARIMA and neural network model. *Journal of Neurocomputing*, 50:159–175, 2003.