

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Métodos de Identificação Fuzzy Para Modelos Autoregressivos
Sazonais Mediante a Função de Autocorrelação Estendida

José Gracildo de Carvalho Júnior

TD : 22 / 2016

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém - Pará - Brasil
2016

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Métodos de Identificação Fuzzy Para Modelos Autoregressivos
Sazonais Mediante a Função de Autocorrelação Estendida

José Gracildo de Carvalho Júnior

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém - Pará - Brasil
2016

Dados Internacionais de Catalogação - na - Publicação (CIP) Sistema de
Bibliotecas da UFPA

Carvalho Júnior, José Gracildo de, 1974-

Métodos de identificação fuzzy para modelos autoregressivos sazonais
mediante a função de autocorrelação estendida / José Gracildo de Carvalho
Júnior.-2016.

Orientador: Carlos Tavares da Costa Júnior

Tese (Doutorado) - Universidade Federal do Pará, Instituto de
Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém,
2016.

1. Lógica fuzzy. 2. Sistemas difusos. 3. Estatística matemática. . I.

Título.

CDD 23. ed. 511.313

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Métodos de Identificação Fuzzy Para Modelos Autoregressivos Sazonais
Mediante a Função de Autocorrelação Estendida

José Gracildo de Carvalho Júnior

Tese de Doutorado Submetida à Avaliação da Banca Examinadora Aprovada Pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Pará e Julgada Adequada Para Obtenção do Grau de Doutor em Engenharia Elétrica na Área de Sistemas de Energia Elétrica.

Aprovada em 13 de Dezembro de 2016

Prof. Evaldo Gonçalves Pelaes, *Dr.*
(Coordenador do Programa de Pós-Graduação em Engenharia Elétrica)

Banca Examinadora

Prof. Carlos Tavares da Costa Júnior, *Dr.*
Universidade Federal do Pará
Orientador acadêmico

Prof. Walter Barra Júnior, *Dr.*
Universidade Federal
Examinador do PPGEE

Prof. Joaquim Carlos Barbosa Queiroz, *Dr.*
Universidade Federal do Pará
Examinador externo ao PPGEE

Prof. Luis Antonio Aguirre, *Dr.*
Universidade Federal de Minas Gerais
Examinador externo ao PPGEE

Prof. Vinícius Duarte Lima, *Dr.*
Universidade Federal do Pará
Examinador externo ao PPGEE

Prof. João Caldas do Lago Neto, *Dr.*
Universidade Federal do Amazonas
Examinador externo ao PPGEE

Dedicatória

*A Meus Pais José Gracildo de Carvalho
e Maria Trindade Corrêa de Carvalho.*

*A Minha esposa Glauce Brandão de Carvalho
e as minhas filhas Giulia Brandão de Carvalho
e Gabrielle Brandão de Carvalho.*

Agradecimentos

Primeiramente à Deus, pela vida, saúde e capacidade para produzir este trabalho.

À minha família, pelo apoio incondicional em todos os momentos de realização do curso de doutorado, em especial nos momentos de grande dificuldade.

Ao Professor Carlos Tavares da Costa Júnior, por ter aceitado me orientar, e também, por me auxiliar em vários momentos durante todo o curso de doutorado.

Ao Professor Joaquim Carlos Barbosa Queiroz, por ter me indicado ao Professor Carlos Tavares da Costa Júnior, como futuro discente do Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da Universidade Federal do Pará (UFPA), e seu provável orientando de doutorado.

Ao professor João Caldas do Lago Neto, pela proposição do tema para a minha tese de doutorado, além das várias orientações, e dos inúmeros esforços e contribuições do mesmo, para que este trabalho fosse concluído com êxito o mais rápido possível.

Ao professor Sandro Dimy Barbosa Bitar, pelas contribuições e incentivos durante a realização do curso de doutorado.

Aos professores Silvia dos Santos de Almeida e Edson Marcos Leal Soares Ramos, pelo apoio, auxílio e suporte incondicional, em todos os momentos de realização do curso de doutorado.

Aos professores Walter Barra Junior e José Augusto Lima Barreiros, pela ajuda e colaboração científica na construção de artigos e conseqüentemente, na elaboração da tese de doutorado.

Aos membros do Laboratório de Sistema de Informação e Georreferenciamento (LASIG) e do Grupo de Estudos e Pesquisas Estatísticas e Computacionais (GEPEC) da Universidade Federal do Pará, pelo suporte técnico computacional em todos os momentos necessários.

À direção do Instituto de Ciências Exatas e Naturais (ICEN) e da Faculdade de Estatística (FAEST) da Universidade Federal do Pará, pela liberação das minhas atividades docentes, em parte do período de realização do curso de doutorado.

Aos amigos da sala de estudos no Laboratório de Engenharia Elétrica: Marcus Ciro, Cleonor Neves, José Rubem Sicchar, Florindo Ayres Junior, Erick Melo, Renan Landau, Anderson de França, Leiliane Cunha, Raimundo Mota, Vanilson Pereira, dentre outros, pelos vários diálogos científicos, e também, pela motivação em vários momentos de extrema dificuldade durante o curso de doutorado.

À Pró-Reitoria de Pesquisa e Pós-Graduação (PROPESP) da Universidade Federal do Pará, pelo financiamento parcial dos artigos produzidos durante o curso de doutorado, mediante o Edital do Programa de Apoio à Publicação Qualificada (PAPQ) para tradução de artigos científicos.

Ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Pará, pela oportunidade de ter realizado este curso de doutorado.

Finalmente, a todos que diretamente ou indiretamente contribuíram de alguma forma, para a realização e conclusão do curso de doutorado em engenharia elétrica na Universidade Federal do Pará.

*“Amo o Senhor Todo Poderoso,
porque ele ouve a minha voz
e as minhas súplicas.
Porque inclinou para mim
os seus ouvidos,
invocá-lo-ei enquanto eu viver ”.*

Salmos CXVI; cap.1,2.

Resumo

CARVALHO JR., José Gracildo de. Métodos de identificação *fuzzy* para modelos autoregressivos sazonais mediante a função de autocorrelação estendida. 2016. Tese (Doutorado em Engenharia Elétrica) - PPGEE/UFPA, Belém - PA, Brasil.

Neste estudo, é proposta uma estratégia baseada na metodologia *fuzzy* para a melhoria do desempenho das previsões de dados mediante um modelo de série temporal. Esta metodologia é concebida para modelagem de processos autoregressivos sazonais de média móvel e pode ser adotada sobre diversas aplicações no mundo real. Por meio da abordagem híbrida, baseada em uma versão da função de autocorrelação *fuzzy*, a interpolação e as capacidades de generalização de sistemas *fuzzy* foram exploradas para se obter uma previsão robusta, mesmo considerando séries de curta ou longa duração. A fim de aumentar a precisão do algoritmo de identificação proposto, vários parâmetros de desempenho foram testados e otimizados por simulações computacionais. Os seguintes parâmetros foram considerados nesse processo: o comprimento de trajetória da série histórica, o número de conjuntos *fuzzy*, e o limite para ativação do suporte dos conjuntos *fuzzy* triangulares. Observou-se que a função de pertinência triangular contribuiu para a melhoria do desempenho no modelo de previsão. Para demonstrar a eficácia da metodologia proposta, foram implementados quatro estudos de caso a partir de dados disponíveis na literatura. Os resultados confirmaram o bom desempenho do algoritmo proposto, permitindo a obtenção de um erro de previsão pequeno, sobretudo, em comparação com metodologias de identificação paramétrica consolidadas na literatura. As projeções produzidas pelo novo método proposto, quando submetidas ao conceito de intervalo de confiança *fuzzy*, demonstraram uma precisão satisfatória.

Palavra-Chave: Função de autocorrelação, modelo autoregressivo sazonal *fuzzy*, previsão *fuzzy*.

Abstract

CARVALHO Jr., José Gracildo de. Methods for identifying fuzzy models autoregressive seasonal by extended autocorrelation function. 2016. Thesis (Doctor in Electrical Engineering) - Postgraduation Program of Electrical Engineering/UFPA, Belém - PA, Brazil.

In this study, a fuzzy-based strategy for improvement of forecasting performance in data time series analysis is proposed. The designed methodology is target to seasonal autoregressive moving average processes modelling and can be applied to an wide range of real world applications. By means of hybrid approach based on a fuzzy version of correlation functions, the interpolating and the generalization capabilities of fuzzy systems are exploited in order to obtain a robust forecasting, even considering series with missing data points. In order to increase the algorithm accuracy, several design parameters were tested and optimized by computational tests. The following parameters are considered in this process: the length of the trajectory of the time series, the number of fuzzy sets, and the limit for activation of the support of the triangular fuzzy sets. It was observed that the membership function of triangular form lead to improved forecasting performance. A simulation to evaluate the accuracy of the forecasting of a fuzzy seasonal autoregressive model is described. To demonstrate the effectiveness of the proposed methodology, four case studies on data from some public data base was carried-out. The results confirm the improved performance of the proposed algorithm, allowing to obtain a reduced forecasting error in comparison to a conventional statistical methodology and fuzzy, for instance. The projections produced by the new method when subjected to fuzzy confidence interval analysis showed satisfactory accuracy.

Keywords: Autocorrelation function, fuzzy seasonal autoregressive model, fuzzy forecasting.

Sumário

Resumo	ix
Abstract	x
Lista de Tabelas	xiv
Lista de Figuras	xxi
1 Introdução	1
1.1 - Contextualização	1
1.2 - Principais Considerações	4
1.3 - Importância e Justificativa da Tese	7
1.4 - A Hipótese Básica da Tese	8
1.5 - Definições dos Objetivos	8
1.5.1 - Objetivo Geral	8
1.5.2 - Objetivos Específicos	8
1.6 - Estrutura do Trabalho	9
2 O Estado da Arte	10
2.1 - Introdução	10
2.2 - Modelos de Previsão Fuzzy	13
2.3 - Conclusão do Capítulo	26
3 Metodologia dos Sistemas Fuzzy	28
3.1 - Introdução	28
3.2 - Teoria dos Conjuntos Fuzzy e Suas Expressões Matemáticas	28
3.2.1 - Definição dos Conjuntos Fuzzy	28
3.2.2 - Conjuntos Fuzzy e Suas Operações Algébricas	31
3.2.3 - Definição de Variável Linguística	32
3.2.4 - Definição das Regras Fuzzy	33
3.2.5 - Definição de Medidas Fuzzy	34
3.2.6 - Definição de Fuzzificador e Defuzzificador	34
3.2.7 - Intervalo de Confiança Fuzzy	37
3.3 - Conclusão do Capítulo	39
4 Série Temporal Fuzzy	40

4.1 - Introdução	40
4.2 - Metodologia Sobre Série Temporal Fuzzy	41
4.2.1 - Modelo Fuzzy Autoregressivo de ordem 1 descrito em “Song <i>et al.</i> (1997)	43
4.3 - Algoritmos de Previsão Para Série Temporal Fuzzy	45
4.3.1 - Algoritmo de Song e Chissom (1993a) e (1993b)	46
4.3.2 - Algoritmo de Chen (1996)	47
4.3.3 - Algoritmo de Chen (2002)	48
4.3.4 - Algoritmo de Liu (2007)	49
4.4 - Alguns Métodos Conhecidos e Disponíveis na Literatura Para Identificação dos Modelos de Série Temporal Fuzzy	50
4.4.1 - Algumas Medidas de Dependência Conhecidas Entre Conjuntos Fuzzy	51
4.5 - Desenvolvimento e Proposição de Novas Ferramentas Para Identificação dos Modelos de Série Temporal Fuzzy	59
4.6 - Um Novo Algoritmo Proposto Para Identificação de Modelos Fuzzy em Série Temporal	61
4.7 - Conclusão do Capítulo	64
5 Simulações Mediante as Ferramentas Desenvolvidas Para a Identificação dos Modelos de Série Temporal Fuzzy	66
5.1 - Simulação de séries temporais fuzzy com várias quantidades de conjuntos <i>fuzzy</i> e combinações de parâmetros ϕ_1 e ϕ_2	66
5.1.1 - Definição do plano experimental	66
5.1.2 - Série temporal <i>fuzzy</i> com função $W_{i,k,j}^{DF_Z}$, $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$, para $T = 24$, $m = 5; 7; 9; 23; 47$, $\phi_1 = 1, 6; 0, 5; 0, 5; 0, 5$ e $\phi_2 = -0, 8; -0, 5; 0, 3; 0, 0$, respectivamente.	67
5.1.3 - Série temporal fuzzy com função $W_{i,k,j}^{DF_Z}$, $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$, para $T = 36$, $m = 5; 7; 9; 23; 47$, $\phi_1 = 1, 6; 0, 5; 0, 5; 0, 5$ e $\phi_2 = -0, 8; -0, 5; 0, 3; 0, 0$, respectivamente.	74
5.1.4 - Série temporal fuzzy com função $W_{i,k,j}^{DF_Z}$, $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$, para $T = 96$, $m = 5; 7; 9; 23; 47$, $\phi_1 = 1, 6; 0, 5; 0, 5; 0, 5$ e $\phi_2 = -0, 8; -0, 5; 0, 3; 0, 0$, respectivamente.	80
5.2 - Conclusão do Capítulo	87
6 Utilização Sobre Dados Reais dos Métodos Propostos Para a Identificação de Modelos de Série Temporal Fuzzy	92
6.1 Estudo de Caso e Comparação dos Resultados Obtidos	92
6.1.1 - Estudo de caso 1: Previsão fuzzy para o preço da energia elétrica em New South Wales, na Austrália.	92
6.1.2 - Estudo de caso 2: Previsão fuzzy para a carga de energia elétrica de curto prazo.	121
6.1.3 - Estudo de caso 3: Previsão fuzzy para a quantidade de matrículas na Universidade do Alabama nos Estados Unidos	143

6.1.4 - Estudo de caso 4: Previsão fuzzy para o índice TAIEX (Taiwan Futures Exchange)	171
6.2 - Conclusão do Capítulo	194
7 Conclusões Finais da Tese e Recomendações Para Trabalhos Futuros	197
7.1 - Conclusões Finais Sobre os Resultados Alcançados Com Este Trabalho . .	197
7.2 - Recomendações Para Trabalhos Futuros	200
A Série Temporal Clássica	202
A.1 - Introdução	202
A.2 - Visão Geral Sobre Série Temporal	202
A.3 - Função de Autocorrelação Amostral e Função de Autocorrelação Parcial Amostral	204
A.3.1 - Formas de Identificação dos Processos Autoregressivos de Médias Móveis Mediante a Função de Autocorrelação Amostral e Parcial . . .	206
A.3.2 - Análise Estatística da Função de Autocorrelação Amostral e Parcial.	206
A.4 - Modelo Sazonal Autoregressivo Integrado de Médias Móveis	210
A.4.1 - Sazonalidade Determinística	210
A.4.2 - Sazonalidade Estocástica	211
A.5 - Adequação do Modelo de Série Temporal	213
A.5.1 - Estacionariedade	213
A.5.2 - Invertibilidade	215
A.5.3 - Ruído Branco	216
A.6 - Função de Autocorrelação Inversa Amostral	217
A.6.1 - Características Para o “Bom Ajuste” de Um Modelo ARIMA	219
A.7 - Função de Autocorrelação Estendida	220
A.7.1 - Características da Função de Autocorrelação Estendida	224
A.8 - Erros de Previsão dos Modelos de Série Temporal Ajustados	226
A.9 - Critérios de Informação Para os Modelos de Série Temporal	229
A.10- Conclusão do Apêndice A	232
Referências Bibliográficas	234

Lista de Tabelas

5.1	- Frequência relativa (%) dos eventos $P W_{i,k,j}^{DFZ} > 2T^{-\frac{1}{2}}, P W_{i,k,j}^{CA}(h; \alpha; b) > 2T^{-\frac{1}{2}}$ e $P \overline{W}_{i,k,j}^{RA}(h) > 2T^{-\frac{1}{2}}$ em 1.000 simulações de processos autoregressivo de ordem 2, com o comprimento da série $T = 24$ ($2T^{-\frac{1}{2}} = 0,4082$) para valores de $m = 5; 7; 9; 23; 47$ (termos linguísticos ou conjuntos fuzzy) e quatro combinações dos parâmetros ϕ_1 e ϕ_2	68
5.2	- Frequência relativa (%) dos eventos $P W_{i,k,j}^{DFZ} > 2T^{-\frac{1}{2}}, P W_{i,k,j}^{CA}(h; \alpha; b) > 2T^{-\frac{1}{2}}$ e $P \overline{W}_{i,k,j}^{RA}(h) > 2T^{-\frac{1}{2}}$ em 1.000 simulações de processos autoregressivo de ordem 2, com o comprimento da série $T = 36$ ($2T^{-\frac{1}{2}} = 0,3333$) para valores de $m = 5; 7; 9; 23; 47$ (termos linguísticos ou conjuntos fuzzy) e quatro combinações dos parâmetros ϕ_1 e ϕ_2	75
5.3	- Frequência relativa (%) dos eventos $P W_{i,k,j}^{DFZ} > 2T^{-\frac{1}{2}}, P W_{i,k,j}^{CA}(h; \alpha; b) > 2T^{-\frac{1}{2}}$ e $P \overline{W}_{i,k,j}^{RA}(h) > 2T^{-\frac{1}{2}}$ em 1.000 Simulações de processos autoregressivo de ordem 2, com o comprimento da série $T = 96$ ($2T^{-\frac{1}{2}} = 0,2041$) para valores de $m = 5; 7; 9; 23; 47$ (termos linguísticos ou conjuntos <i>fuzzy</i>) e quatro combinações dos parâmetros ϕ_1 e ϕ_2	81
6.1	- Resultados da função $W_{i,k,j}^{DFZ}$ para o preço da energia elétrica em New South Wales.	95
6.2	- Probabilidades de significância para os resultados da Tabela (6.1).	96
6.3	- Dados históricos referentes ao preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de <i>fuzzificação</i> realizado a partir de um número <i>fuzzy</i> triangular, relação <i>fuzzy</i> lógica e os preços em dólar da energia elétrica <i>defuzzificados</i>	97
6.4	- Grupos de relações <i>fuzzy</i> lógicas ativadas para os dados da Tabela (6.3).	98
6.5	- Teste da raiz unitária de <i>Dickey-Fuller</i> aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$	98
6.6	- Teste da raiz unitária de <i>Dickey-Fuller</i> aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$	98
6.7	- Teste de significância dos parâmetros do modelo FSARIMA(3; 1; 2)(1; 0; 1) ₁₃	99
6.8	- Teste Q (Ljung-Box) para o modelo FSARIMA(3; 1; 2)(1; 0; 1) ₁₂	99
6.9	- Comparação dos resultados de previsão pelo método proposto (preços <i>defuzzificados</i>) em relação aos três métodos propostos em Dong et al. (2011).101	101

6.10 - Resultados das projeções para o preço em dólar da energia elétrica em New South Wales na Austrália, nos seis instantes de tempo posteriores a série histórica observada.	102
6.11 - Resultados da função $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b)$ para o preço da energia elétrica.104	104
6.12 - Probabilidade de significância para os resultados da Tabela (6.11).	104
6.13 - Processo de <i>fuzzificação</i> realizado a partir de um número <i>fuzzy</i> triangular, relação <i>fuzzy</i> lógica e a série dos preços em dólar da energia elétrica com alfa-nível de 0,25.	105
6.14 - Série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de <i>fuzzificação</i> realizado a partir de um número <i>fuzzy</i> triangular, relação <i>fuzzy</i> lógica e a série dos preços em dólar da energia elétrica com alfa-nível de 0,50.	106
6.15 - Resultados da função $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ para o preço da energia elétrica.107	107
6.16 - Probabilidade de significância para os resultados da Tabela (6.15).	107
6.17 - Resultados da função $W_{i,k,j}^{CA}(h; \alpha = 0, 75; b)$ para o preço da energia elétrica.108	108
6.18 - Probabilidade de significância para os resultados da Tabela (6.17).	108
6.19 - Série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de <i>fuzzificação</i> realizado a partir de um número <i>fuzzy</i> triangular, relação <i>fuzzy</i> lógica e a série dos preços em dólar da energia elétrica com alfa-nível de 0,75.	109
6.20 - Grupos de relações <i>fuzzy</i> lógicas ativadas com alfa-nível de 0,25; 0,50 e 0,75.110	110
6.21 - Teste de significância dos parâmetros do modelo FSARIMA(2; 1; 2)(1; 0; 1) ₁₂	111
6.22 - Teste Q (Ljung-Box) para o modelo FSARIMA(2; 1; 2)(1; 0; 1) ₁₂	111
6.23 - Comparação dos resultados de previsão pelo método proposto (preços com α -nível) em relação aos três métodos propostos em Dong et al. (2011).	113
6.24 - Resultados das projeções para o preço em dólar da energia elétrica em New South Wales na Austrália, correspondendo aos seis instantes de tempo posteriores ao período original da série histórica observada, mediante o método proposto (<i>alfa-nível</i>).	114
6.25 - Grupos de relações <i>fuzzy</i> lógicas ativadas para os dados da Tabela (6.26). 115	115
6.26 - Série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de <i>fuzzificação</i> realizado a partir de um número <i>fuzzy</i> triangular, relação <i>fuzzy</i> lógica e a série das funções de autocorrelação amostrais médias aleatórias.	116
6.27 - Resultados da função $\overline{W}_{i,k,j}^{RA}(h)$ para o preço da energia elétrica.	117
6.28 - Probabilidades de significância para os resultados da Tabela (6.27).	117
6.29 - Teste de significância dos parâmetros do modelo FSARIMA(2; 1; 2)(1; 0; 1) ₂₄	118
6.30 - Teste Q (Ljung-Box) para o modelo FSARIMA(2; 1; 2)(1; 0; 1) ₂₄	118
6.31 - Comparação dos resultados de previsão pelo método proposto (função de autocorrelação média aleatória) em relação aos três métodos propostos em Dong et al. (2011).	119

6.32 - Resultados das projeções para o preço em dólar da energia elétrica em New South Wales na Austrália, correspondendo aos seis instantes de tempo posteriores ao período original da série histórica observada, mediante o método proposto (fac média aleatória).	120
6.33 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de <i>fuzzificação</i> mediante a pertinência máxima pelo número <i>fuzzy</i> triangular, lógica <i>fuzzy</i> simples e agrupada, série dos valores <i>defuzzificados</i> pelo método centroide em cada intervalo u_j associado.	123
6.34 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$	124
6.35 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$	124
6.36 - Resultados da função $W_{i,k,j}^{DFz}$ para a carga de energia elétrica.	124
6.37 - Probabilidade de significância para os resultados apresentados na Tabela 6.36.	125
6.38 - Testes das estimativas dos parâmetros ao modelo FSARIMA(4,1,1)(1,0,1) ₃₆ ajustado a carga de energia elétrica.	125
6.39 - Comparação das previsões pelo método proposto (dados <i>defuzzificados</i>) em relação aos cinco métodos propostos por Sadaei et al. (2014) para a carga de energia elétrica.	126
6.40 - Teste Box-Pierce modificado (Ljung-Box) pela estatística Qui-quadrado ao modelo FSARIMA(4,1,1)(1,0,1) ₃₆ ajustado para carga de energia elétrica.	127
6.41 - Intervalo de confiança <i>fuzzy</i> para as 6 projeções de carga de energia elétrica pelo método proposto (dados <i>defuzzificados</i>).	127
6.42 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de <i>fuzzificação</i> mediante a definição de número <i>fuzzy</i> triangular, lógica <i>fuzzy</i> simples e agrupada, série da carga de energia elétrica com α -nível = 0,25 a partir de cada intervalo u_j associado.	128
6.43 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$ para a carga de energia elétrica.	129
6.44 - Probabilidade de significância dos resultados da Tabela (6.43).	129
6.45 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de <i>fuzzificação</i> mediante a definição de número <i>fuzzy</i> triangular, lógica <i>fuzzy</i> simples e agrupada, série da carga de energia elétrica com α -nível = 0,50 a partir de cada intervalo u_j associado.	130
6.46 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,50; b)$ para a carga de energia elétrica.	131
6.47 - Probabilidade de significância dos resultados da Tabela (6.46).	131
6.48 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de <i>fuzzificação</i> mediante a definição de número <i>fuzzy</i> triangular, lógica <i>fuzzy</i> simples e agrupada, série da carga de energia elétrica com α -nível = 0,75 a partir de cada intervalo u_j associado.	133
6.49 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$ para a carga de energia elétrica.	134
6.50 - Probabilidade de significância dos resultados da Tabela (6.49).	134
6.51 - Testes das estimativas dos parâmetros ao modelo FSARIMA(5,1,1)(1,0,1) ₃₆ ajustado a carga de energia elétrica.	134

6.52	- Teste de <i>Box-Pierce</i> modificado (<i>Ljung-Box</i>) pela estatística Qui-quadrado ao modelo FSARIMA(5,1,1)(1,0,1) ₃₆ ajustado para carga de energia elétrica.	134
6.53	- Comparação dos resultados das previsões mediante o método proposto (α -nível) em relação aos cinco métodos propostos em Sadaei et al. (2014) para a carga de energia elétrica.	136
6.54	- Intervalo de confiança <i>fuzzy</i> para as 6 projeções de carga de energia elétrica pelo método proposto (α -nível).	137
6.55	- Série histórica da carga de energia elétrica em 25/9/2005, com processo de <i>fuzzificação</i> mediante a pertinência máxima do número <i>fuzzy</i> triangular, lógica <i>fuzzy</i> simples e agrupada, série da função de autocorrelação amostral média aleatória em cada intervalo u_j	138
6.56	- Resultados da $\overline{W}_{i,k,j}^{RA}(h; b)$ para a carga de energia elétrica.	139
6.57	- Probabilidade de significância dos resultados da Tabela 6.56.	139
6.58	- Testes das estimativas dos parâmetros ao modelo FSARIMA(3,1,1)(1,0,1) ₃₆ ajustado a carga de energia elétrica.	140
6.59	- Teste de <i>Box-Pierce</i> modificado (<i>Ljung-Box</i>) pela estatística Qui-quadrado ao modelo FSARIMA(3,1,1)(1,0,1) ₃₆ ajustado para carga de energia elétrica.	140
6.60	- Comparação dos resultados das previsões do método proposto (fac média aleatória) com os cinco métodos propostos em Sadaei et al. (2014) para a carga de energia elétrica.	141
6.61	- Intervalo de confiança <i>fuzzy</i> para as 6 projeções de carga de energia elétrica pelo método proposto (fac média aleatória).	142
6.62	- Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de <i>Fuzzificação</i> Realizado a partir de Um Número <i>Fuzzy</i> Triangular, Relação <i>Fuzzy</i> Lógica e os Valores de Matrículas <i>Defuzzificadas</i>	146
6.63	- Grupos de Relações <i>Fuzzy</i> Lógicas Ativadas Para os Dados da Tabela (6.62).	147
6.64	- Resultados da $W_{i,k,j}^{DFZ}$ para as Matrículas na Universidade do Alabama.	148
6.65	- Probabilidades de Significância dos Resultados da $W_{i,k,j}^{DFZ}$ (Tabela 6.64).	148
6.66	- Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$	149
6.67	- Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$	149
6.68	- Teste de Significância dos Parâmetros do Modelo FSARIMA(4; 1; 1)(1; 0; 1) ₁₃	150
6.69	- Teste Q (Ljung-Box) Para o Modelo FSARIMA(4; 1; 1)(1; 0; 1) ₁₃	151
6.70	- Comparação dos resultados de previsões pelo método proposto neste trabalho (função de autocorrelação <i>defuzzificada</i>), em relação a 14 métodos de previsão fuzzy que encontram-se publicados na literatura (Song e Chisson, 1993b - Chen e Tanuwijaya, 2011).	152
6.71	- Comparação dos resultados de previsões pelo método proposto neste trabalho (função de autocorrelação <i>defuzzificada</i>), em relação a 7 métodos de previsão <i>fuzzy</i> que encontram-se publicados na literatura (Gangwar e Kumar, 2012 - Lu et al., 2015).	153

6.72 - Resultados das Projeções Para a Quantidade de Matrículas na Universidade do Alabama nos Estados Unidos Pelo Método Proposto Entre os Anos de 1993 e 1998.	155
6.73 - Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de Fuzzificação Realizado A Partir de Um Número <i>Fuzzy</i> Triangular, Relação <i>Fuzzy</i> Lógica e os Valores de Matrículas Com Alfa-Nível de 0,25.	157
6.74 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b)$ para Matrículas na Universidade do Alabama.	158
6.75 - Probabilidades de Significância da $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b)$ na Tabela (6.74).	158
6.76 - Teste de Significância dos Parâmetros do Modelo FSARIMA(1; 1; 0)(1; 0; 0) ₁₂	159
6.77 - Teste <i>Q</i> <i>Ljung-Box</i> Para o Modelo FSARIMA(1; 1; 0)(1; 0; 0) ₁₂	159
6.78 - Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de Fuzzificação Realizado A Partir de Um Número <i>Fuzzy</i> Triangular, Relação <i>Fuzzy</i> Lógica e os Valores de Matrículas Com Alfa-Nível de 0,50.	160
6.79 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ para Matrículas na Universidade do Alabama.	161
6.80 - Probabilidades de Significância da $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ na Tabela (6.79).	161
6.81 - Teste de Significância dos Parâmetros do Modelo ARIMA(1; 1; 1).	162
6.82 - Teste <i>Q</i> <i>Ljung-Box</i> Para o Modelo ARIMA(1; 1; 1).	162
6.83 - Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de <i>Fuzzificação</i> Realizado A Partir de Um Número <i>Fuzzy</i> Triangular, Relação <i>Fuzzy</i> Lógica e os Valores de Matrículas Com Alfa-Nível de 0,75.	164
6.84 - Grupos de Relações <i>Fuzzy</i> Lógicas Ativadas Com Alfa-Nível de 0,25; 0,50 e 0,75.	164
6.85 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 75; b)$ para Matrículas na Universidade do Alabama.	165
6.86 - Probabilidades de Significância da $W_{i,k,j}^{CA}(h; \alpha = 0, 75; b)$ na Tabela (6.85).	165
6.87 - Teste de Significância dos Parâmetros do Modelo ARIMA(1; 1; 3).	166
6.88 - Grupos de Relações <i>Fuzzy</i> Lógicas Ativadas Para os Dados da Tabela (6.89).	168
6.89 - Dados Históricos da Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Processo de <i>Fuzzificação</i> Realizado Por Um Número <i>Fuzzy</i> Triangular, Relação <i>Fuzzy</i> Lógica e a Série dos Valores da Função de Autocorrelação Amostral Média Aleatória.	169
6.90 - Resultados da $\bar{W}_{i,k,j}^{RA}(h)$ para as Matrículas na Universidade do Alabama.	169
6.91 - Probabilidades de Significância de $\bar{W}_{i,k,j}^{RA}(h)$ na Tabela (6.90).	170
6.92 - Teste de Significância dos Parâmetros do Modelo ARIMA(1; 1; 1).	170
6.93 - Teste <i>Q</i> <i>Ljung-Box</i> Para o Modelo ARIMA(1; 1; 1).	170

6.94 - Série histórica do índice TAIFEX em 3 de Agosto de 1998 a 30 de Setembro de 1998, com processo de <i>fuzzificação</i> utilizando a função de pertinência triangular, relações <i>fuzzy</i> lógicas simples, relações <i>fuzzy</i> lógicas agrupadas e a série temporal <i>defuzzificada</i> do índice TAIFEX.	173
6.95 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$	174
6.96 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$	174
6.97 - Resultados da função $W_{i,k,j}^{DFZ}$ para o índice TAIFEX.	174
6.98 - Probabilidade de significância dos resultados da Tabela 6.97.	175
6.99 - Testes das estimativas dos parâmetros ao modelo FSARIMA(2,1,1)(1,0,1) ₃₆ ajustado para o índice TAIFEX.	175
6.100- Teste <i>Box-Pierce</i> modificado (<i>Ljung-Box</i>) pela estatística Qui-quadrado ao modelo FSARIMA(2,1,1)(1,0,1) ₃₆ ajustado para o índice TAIFEX.	175
6.101- Comparação dos resultados das previsões pelo método proposto (dados <i>defuzzificados</i>) em relação aos cinco métodos propostos em Wang e Chen (2009) para o índice TAIFEX.	177
6.102- Intervalo de confiança <i>fuzzy</i> para as 6 projeções do índice TAIFEX pelo método proposto (dados <i>defuzzificados</i>).	178
6.103- Série histórica do índice TAIFEX de 03/08/1998 a 30/10/1998, com processo de <i>fuzzificação</i> utilizando a função de pertinência triangular, relações <i>fuzzy</i> lógicas simples, relações <i>fuzzy</i> lógicas agrupadas e a série temporal com α -nível = 0,25 para o índice TAIFEX.	179
6.104- Série histórica do índice TAIFEX de 03/08/1998 a 30/10/1998, com processo de <i>fuzzificação</i> utilizando a função de pertinência triangular, relações <i>fuzzy</i> lógicas simples, relações <i>fuzzy</i> lógicas agrupadas e a série temporal com α -nível = 0,50 para o índice TAIFEX.	180
6.105- Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b)$ para o índice TAIFEX.	181
6.106- Probabilidade de significância dos resultados da Tabela (6.105).	181
6.107- Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ para o índice TAIFEX.	182
6.108- Probabilidade de significância dos resultados da Tabela (6.107).	182
6.109- Série histórica do índice TAIFEX de 03/08/1998 a 30/10/1998, com processo de <i>fuzzificação</i> utilizando a função de pertinência triangular, relações <i>fuzzy</i> lógicas simples, relações <i>fuzzy</i> lógicas agrupadas e a série temporal com α -nível = 0,75 para o índice TAIFEX.	183
6.110- Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 75; b)$ para o índice TAIFEX.	184
6.111- Probabilidade de significância dos resultados da Tabela (6.110).	184
6.112- Testes das estimativas dos parâmetros ao modelo FSARIMA(4,1,2)(1,0,1) ₃₆ ajustado para o índice TAIFEX.	184
6.113- Teste <i>Box-Pierce</i> modificado (<i>Ljung-Box</i>) pela estatística Qui-quadrado ao modelo FSARIMA(4,1,2)(1,0,1) ₂₄ ajustado para o índice TAIFEX.	185
6.114- Comparação dos resultados das previsões pelo método proposto (α -nível) em relação aos cinco métodos propostos em Wang e Chen (2009) para o índice TAIFEX.	186

6.115-	Intervalo de confiança <i>fuzzy</i> para as 6 projeções do índice TAIFEX pelo método proposto (α -nível).	187
6.116	Série histórica do índice TAIFEX em 03/08/1998 a 30/10/1998, processo de <i>fuzzificação</i> mediante a pertinência máxima do número <i>fuzzy</i> triangular, lógica <i>fuzzy</i> simples e agrupada, série da função de autocorrelação amostral média aleatória em cada intervalo u_j	188
6.117-	Resultados da função $W_{i,k,j}^{RA}(h; b)$ para o índice TAIFEX.	189
6.118-	Probabilidade de significância dos resultados da Tabela 6.117.	189
6.119-	Testes estatísticos para as estimativas dos parâmetros do modelo FSARIMA(2,1,5)(1,0,1) ₃₆ ajustado para o índice TAIFEX.	190
6.120-	Teste <i>Box-Pierce</i> modificado (<i>Ljung-Box</i>) pela estatística <i>Qui-quadrado</i> ao modelo FSARIMA(2,1,5)(1,0,1) ₃₆ ajustado para o índice TAIFEX.	190
6.121-	Comparação dos resultados das previsões pelo método proposto ($W_{i,k,j}^{RA}(h; b)$) em relação aos cinco métodos propostos em Wang e Chen (2009) para o índice TAIFEX.	191
6.122-	Comparação do MASE obtido por vários métodos para o TAIFEX no período de 3 de agosto de 1998 a 30 de setembro de 1998.	192
6.123-	Comparação do MASE obtido por vários métodos para o TAIFEX no período de 10 de setembro de 1998 a 30 de setembro de 1998.	192
6.124-	Intervalo de confiança <i>fuzzy</i> para as 6 projeções do índice TAIFEX pelo método proposto ($W_{i,k,j}^{RA}(h; b)$).	193
A.1	- Resumo sobre as características observadas nas funções de identificação para modelos de série temporal: FAC, FACP e FACI.	220
A.2	- Exemplo do Processo de Identificação Pela Função de Autocorrelação Estendida Teórica Para Um Modelo ARMA(1;1), tal como, Descrito em Cryer e Chan (2008).	224
A.3	- Resultados da $W_{i,k,j}$ Para o Consumo Total de Energia Elétrica (MW/h).	225
A.4	- Probabilidades de Significância de $W_{i,k,j}$ Para os Resultados da Tabela (A.3).	226

Lista de Figuras

5.1	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = -0, 5$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$	70
5.2	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 1, 6$ e $\phi_2 = -0, 8$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$	71
5.3	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = 0, 0$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$	72
5.4	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = 0, 3$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$	73
5.5	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = -0, 5$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$	76
5.6	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 1, 6$ e $\phi_2 = -0, 8$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$	77
5.7	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = 0, 0$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$	78
5.8	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = 0, 3$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$	79
5.9	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = -0, 5$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$	83
5.10	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 1, 6$ e $\phi_2 = -0, 8$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$	84
5.11	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0, 5$ e $\phi_2 = 0, 0$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$	85

5.12	- Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = 0,3$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$	86
5.13	- Fluxograma para a rotina de identificação em modelos ARMA($p; q$), proposta neste trabalho via simulações realizadas para os valores <i>defuzzificados</i> da série histórica.	89
5.14	- Fluxograma para a rotina de identificação em modelos ARMA($p; q$), proposta neste trabalho via simulações realizadas para os valores da série histórica com <i>alfa-nível</i>	90
5.15	- Fluxograma para a rotina de identificação em modelos ARMA($p; q$), proposta neste trabalho via simulações realizadas para os valores da série fac média aleatória.	91
6.1	- Série temporal diária referente ao preço em dólar da energia elétrica em New South Wales na Austrália, durante 48 instantes de tempo distintos no período das 00:00 AM a 23:30 PM, coletados em intervalos de tempo de meia em meia hora de uma segunda feira.	94
6.2	- Conjuntos <i>fuzzy</i> pela definição de número <i>fuzzy</i> triangular, para o preço em dólar da energia elétrica em New South Wales na Austrália, onde o eixo Y representa a pertinência ao conjunto <i>fuzzy</i> A_i e o eixo X indica o preço em dólar da energia elétrica.	95
6.3	- Série temporal do preço em dólar da energia elétrica em New South Wales na Austrália <i>versus</i> os valores das previsões pelo o método proposto (preços <i>defuzzificados</i>).	103
6.4	- Conjuntos <i>fuzzy</i> construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo Y indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo X o preço em dólar da energia elétrica, com alfa-nível igual a 0,25 para ativar o suporte dos respectivos conjuntos <i>fuzzy</i>	103
6.5	- Conjuntos <i>fuzzy</i> construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo Y indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo X o preço em dólar da energia elétrica, com alfa-nível igual a 0,50 para ativar o suporte dos respectivos conjuntos <i>fuzzy</i>	104
6.6	- Conjuntos <i>fuzzy</i> construídos pela definição de pertinência triangular, onde o eixo Y indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo X o preço em dólar da energia elétrica, com alfa-nível igual a 0,75 para ativar o suporte dos respectivos conjuntos <i>fuzzy</i>	108
6.7	- Série temporal do preço em dólar da energia elétrica em New South Wales na Austrália <i>versus</i> os valores das previsões pelo método proposto (preços com <i>alfa-nível</i>).	114
6.8	- Conjuntos <i>fuzzy</i> pela definição de número <i>fuzzy</i> triangular, onde o eixo Y indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo X o preço em dólar da energia elétrica, para função de autocorrelação amostral média aleatória nos suportes dos conjuntos ativados.	115

6.9	- Série temporal do preço em dólar da energia elétrica em New South Wales na Austrália <i>versus</i> os valores das previsões pelo método proposto (fac média aleatória).	121
6.10	- Série temporal da carga de energia elétrica de curto prazo, durante 48 instantes de tempo no período de 12:30 AM a 11:30 PM coletados a cada meia hora.	121
6.11	- Conjuntos <i>fuzzy</i> para a carga de energia elétrica de curto prazo, durante 48 instantes de tempo no período de 12:30 AM a 11:30 PM coletados a cada meia hora.	122
6.12	- Carga de energia elétrica atual <i>versus</i> o método proposto (dados <i>defuzzificados</i>).	127
6.13	- Conjuntos construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo <i>Y</i> indica a pertinência ao conjunto <i>fuzzy</i> A_i e o eixo <i>X</i> a carga de energia elétrica, com α -nível = 0,25 para ativar o suporte dos conjuntos <i>fuzzy</i> correspondentes.	129
6.14	- Conjuntos construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo <i>Y</i> indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo <i>X</i> a carga de energia elétrica, com α -nível = 0,50 para ativar o suporte dos conjuntos <i>fuzzy</i> correspondentes.	131
6.15	- Conjuntos construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo <i>Y</i> indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo <i>X</i> a carga de energia elétrica, com α -nível = 0,75 para ativar o suporte dos conjuntos <i>fuzzy</i> correspondentes.	132
6.16	- Carga de energia elétrica atual <i>versus</i> o método proposto (α -nível).	137
6.17	- Conjuntos <i>fuzzy</i> para a carga de energia elétrica de curto prazo, durante 48 instantes de tempo no período de 12:30 AM a 11:30 PM coletados a cada meia hora.	139
6.18	- Carga de energia elétrica atual <i>versus</i> o método proposto (fac média aleatória).	142
6.19	- Série Temporal da Quantidade de Matrículas na Universidade do Alabama nos Estados Unidos no Período de 1971 a 1992.	143
6.20	- Conjuntos <i>Fuzzy</i> Pela Definição de Número <i>Fuzzy</i> Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, onde o Eixo <i>Y</i> Representa a Pertinência ao Conjunto <i>Fuzzy</i> A_i e o Eixo <i>X</i> Indica a Quantidade de Matrículas.	146
6.21	- Série Temporal Referente a Quantidade de Matrículas na Universidade do Alabama nos Estados Unidos no Período de 1971 a 1992 <i>Versus</i> a Série Temporal Referente aos Valores das Previsões Mediante o Método Proposto Para o Período de 1971 a 1992.	154
6.22	- Conjuntos <i>Fuzzy</i> Pela Definição de Número <i>Fuzzy</i> Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, Com Alfa-Nível Igual a 0,25, onde o Eixo <i>Y</i> Representa a Pertinência ao Conjunto <i>Fuzzy</i> A_i e o Eixo <i>X</i> Indica a Quantidade de Matrículas.	156

6.23 - Conjuntos <i>Fuzzy</i> Pela Definição de Número <i>Fuzzy</i> Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, Com Alfa-Nível Igual a 0,50, onde o Eixo <i>Y</i> Representa a Pertinência ao Conjunto <i>Fuzzy</i> A_i e o Eixo <i>X</i> Indica a Quantidade de Matrículas.	160
6.24 - Conjuntos <i>Fuzzy</i> Pela Definição de Número <i>Fuzzy</i> Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, Com Alfa-Nível Igual a 0,75, onde o Eixo <i>Y</i> Representa a Pertinência ao Conjunto <i>Fuzzy</i> A_i e o Eixo <i>X</i> Indica a Quantidade de Matrículas.	163
6.25 - Conjuntos <i>Fuzzy</i> Pela Definição de Número <i>Fuzzy</i> Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, onde o Eixo <i>Y</i> Representa a Pertinência ao Conjunto <i>Fuzzy</i> A_j e o Eixo <i>X</i> Indica a Quantidade de Matrículas, Para Auxiliar na Construção da Série dos Valores da Função de Autocorrelação Amostral Média Aleatória.	167
6.26 - Série do índice TAIFEX durante 47 dias no período de 03/08/1998 a 30/09/1998.	171
6.27 - Representação gráfica dos conjuntos <i>fuzzy</i> , onde o eixo <i>X</i> representa os valores do índice TAIFEX no período de 03/08 a 30/09/1998 <i>versus</i> o eixo <i>Y</i> pertinência ao conjunto A_i	174
6.28 - Índice TAIFEX atual <i>versus</i> o método proposto (<i>defuzzificado</i>).	176
6.29 - Conjuntos <i>fuzzy</i> construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo <i>Y</i> indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo <i>X</i> o índice TAIFEX, com α -nível igual a 0,25, para ativar o suporte dos respectivos conjuntos <i>fuzzy</i>	178
6.30 - Conjuntos <i>fuzzy</i> construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo <i>Y</i> indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo <i>X</i> o índice TAIFEX, com α -nível igual a 0,50, para ativar o suporte dos respectivos conjuntos <i>fuzzy</i>	181
6.31 - Conjuntos <i>fuzzy</i> construídos pela definição de número <i>fuzzy</i> triangular, onde o eixo <i>Y</i> indica a pertinência ao conjunto <i>fuzzy</i> A_j e o eixo <i>X</i> o índice TAIFEX, com α -nível igual a 0,75, para ativar o suporte dos respectivos conjuntos <i>fuzzy</i>	182
6.32 - Índice TAIFEX atual <i>versus</i> o método proposto (α -nível).	185
6.33 - Representação gráfica dos conjuntos <i>fuzzy</i> , onde o eixo <i>X</i> representa os valores do índice TAIFEX no período de 03/08 a 30/09/1998 <i>versus</i> o eixo <i>Y</i> pertinência ao conjunto A_j	189
6.34 - Comparação entre os valores do MASE pelos métodos propostos ao índice TAIFEX.	193
6.35 - Índice TAIFEX atual <i>versus</i> o método proposto (fac média aleatória). . .	194
A.1 - Exemplo de Uma Série Temporal Sobre o Consumo de Energia Elétrica no Brasil, entre Janeiro de 2004 e Fevereiro de 2011.	203

A.2 - Exemplo de Uma Série Temporal Sobre o Consumo de Energia Elétrica no Brasil, entre Janeiro de 2004 e Fevereiro de 2011, Na Forma Estacionária (Livre de Tendência).	204
A.3 - Exemplo da FACA Para o Consumo de Energia Elétrica Com 2 <i>lags</i> Significativos.	207
A.4 - Exemplo da FACPA Para o Consumo de Energia Elétrica Com 2 <i>lags</i> Significativos.	207
A.5 - Modelo de Médias Móveis Identificado Pela Função de Autocorrelação Amostral.	208
A.6 - Modelo Autoregressivo Identificado Pela Função de Autocorrelação Parcial.	208

Capítulo 1

Introdução

1.1 - Contextualização

Ferramentas capazes de auxiliar um gestor quanto a medidas à serem adotadas em uma tomada de decisão são de extrema relevância, sobretudo, quando se está diante de um processo em constante evolução, como, por exemplo, o setor de energia elétrica, o qual atualmente dispõe de reservatórios em suas usinas hidrelétricas, em condições abaixo ou próximo dos limites de quinze anos atrás, quando ocorreu a crise do apagão no Brasil.

Neste cenário, medidas emergenciais são necessárias para tentar evitar um colapso no sistema de fornecimento de energia elétrica, tal como, um eventual apagão ou racionamento de energia elétrica no país. A solução encontrada pelos gestores, neste caso, recentemente, foi o acionamento das usinas termelétricas, as quais são mais caras e poluentes, enquanto esperava-se pelas chuvas para normalizar a situação de escassez dos reservatórios. Segundo a avaliação de especialistas da área, esta ação teve como objetivo complementar a geração das hidrelétricas e permitir a recomposição dos reservatórios de água, durante o período mais crítico de fornecimento de energia elétrica no país.

Especialistas do setor de energia defendem a teoria de que, os riscos para o cenário energético brasileiro nos próximos anos são reais, quanto à incapacidade de suprir a necessidade da demanda existente no país. Segundo eles, a ocorrência de aumento dos custos relativos à geração e, conseqüentemente, para a transmissão de energia, aliados à pressão exercida por um setor específico do governo, em favor da redução das tarifas exigidas para a renovação dos contratos das empresas do setor, podem inviabilizar futuros investimentos na área, o que neste caso, seria uma medida direta para contornar a curto e médio prazo o problema de fornecimento de energia elétrica no país, dada a diminuição da rentabilidade que as concessionárias passarão a ter.

Neste contexto, se faz necessário à utilização de mecanismos potencialmente capazes

de prever panoramas indesejáveis, dentro do cenário energético brasileiro, tais como, uma elevada demanda no consumo de energia elétrica e uma redução excessiva nos níveis dos reservatórios das hidrelétricas, dentre outras situações, por conseguinte, impactando diretamente sobre o planejamento energético e podendo causar riscos excessivos para produção, distribuição e consumo de energia elétrica.

Quando se está diante de um cenário de incertezas, no qual se busca descobrir ou pelo menos estimar a ocorrência de certo evento inerente a um processo em análise, ferramentas com capacidade significativa para auxiliar a tomada de decisão se tornam essenciais. Neste sentido, há possibilidade de identificar, por vezes modelar e ainda, monitorar o comportamento de certa característica, com o decorrer do tempo de observação de um processo específico.

Ao se tratar de séries históricas referentes, sobretudo à área de energia, como por exemplo, a oferta, a demanda ou consumo de energia elétrica em uma região, fica evidenciada a contribuição que as ferramentas pertencentes à metodologia de série temporal proporcionam ao ambiente energético em questão.

Tem-se então uma melhor interpretação, análise, monitoramento e previsão das características observáveis, tal como, ocorre em séries temporais de outras áreas do conhecimento, em que os dados coletados pertencem a um intervalo de tempo definido, como se observa desde o surgimento desta metodologia na década de 70 (setenta), na literatura relacionada à série temporal.

Contudo, apesar da metodologia de série temporal ser aplicada em vários problemas de previsão, esta técnica torna-se ineficiente, caso haja inexistência de séries históricas com trajetória mínima (pelo menos vinte e quatro instantes de tempo), como também, se os valores históricos disponíveis forem incorretamente registrados; caso haja ausência de observação(ões) durante o intervalo de tempo da série estudada, ou ainda, se os registros são observados em termos linguísticos (dados não numéricos); etc.

A teoria de conjuntos *fuzzy* proposta por Zadeh (1965) vem ao encontro de uma necessidade inerente dos processos monitorados em um intervalo de tempo estabelecido, onde a metodologia clássica de série temporal é comprometida em termos de precisão e confiabilidade dos resultados obtidos. Na ausência de informações temporais padronizadas, a metodologia *fuzzy* se caracteriza como uma ferramenta eficiente, com grande capacidade

de capturar a forma com que a série de dados se desenvolve no decorrer do tempo, sem desconsiderar a variabilidade das informações agregadas ao processo em análise.

Neste sentido, o processo de modelagem e monitoramento, se torna capaz de produzir resultados de previsão para uma determinada característica de interesse, a partir do modelo *fuzzy* associado com uma série de dados, estes por sua vez, obtidos sistematicamente em instantes de tempo onde o processo se desenvolve, o que suscita de forma teórica, uma análise mais abrangente e precisa, ao que se alcança com a modelagem clássica de série temporal, em relação aos dados monitorados em espaços de tempo contínuos.

Ao se utilizar a metodologia de série temporal, para a análise, interpretação e previsão de dados observados, a grande questão é identificar de maneira eficiente, o modelo que capture (represente) as variações e características gerais presentes nestes dados. Portanto, existe uma necessidade inerente junto ao processo de modelagem, que passa pela identificação do modelo de série temporal, mediante a determinação (estimação) dos parâmetros, que no caso de séries não estacionárias, se tem um modelo auto regressivo integrado de médias móveis (ARIMA), com os parâmetros p , d e q .

Assim, se deve obter estimativas preliminares, onde o parâmetro p , representa o componente autoregressivo (AR); o parâmetro d , o número de diferenças necessárias na série, para tornar a mesma estacionária, ou seja, livre de tendência; e o parâmetro q , representando o componente de médias móveis (MA). Este tipo de metodologia suscita os chamados métodos não-determinísticos (estocásticos), assim como, os métodos paramétricos em termos de modelagem estatística, por considerar os erros associados aos dados observados, além de estimar os parâmetros do modelo de série temporal relacionados aos dados de interesse, tal como, se discute em Aguirre (2007).

De forma equivalente à metodologia clássica de série temporal, na abordagem *fuzzy* para o caso de processos autoregressivos, também se deve seguir uma rotina para a identificação do modelo. Porém, para dados *fuzzy*, se tem os parâmetros “ p ” e “ m ” em modelos autoregressivos *fuzzy* (FAR), o que implica diretamente na determinação da ordem do modelo *fuzzy*, mediante uma função de autocorrelação *fuzzy*, onde, a série de dados originais passa a ser *fuzzificada*, ou seja, agrupada em “ m ” conjuntos *fuzzy*, de acordo com o grau de pertinência do valor observado na série, em relação a um dos “ m ” conjuntos *fuzzy*, os quais são previamente construídos por um especialista.

Neste contexto, a estrutura de correlação instituída entre as observações da série, para uma variável linguística X_i , aliada a quantidade “ m ” de conjuntos *fuzzy*, implicam de forma direta na estimação paramétrica do modelo autoregressivo *fuzzy* e conseqüentemente, em um critério para indicar ou determinar a ordem do parâmetro “ p ” autoregressivo no modelo de série temporal *fuzzy*. Neste caso, a escolha da função de autocorrelação *fuzzy* é extremamente relevante para o processo de identificação do modelo, visto que, dependendo da função de autocorrelação que seja escolhida, a eficiência da rotina de identificação pode ser comprometida, ou seja, uma possível ordem do parâmetro “ p ”, em um processo autoregressivo de uma série temporal *fuzzy*, pode vir a ser mal determinada, comprometendo desta forma as previsões do modelo identificado.

1.2 - Principais Considerações

Dentre os procedimentos de identificação dos modelos de séries temporais tem-se:

1. Avaliar a possível necessidade de transformação da série original, com o intuito de estabilizar a variância, ou seja, tornar os dados mais simétricos em relação ao que se apresentavam anteriormente, se aproximando desta forma de uma distribuição de probabilidade normal.
2. Aplicar uma ou mais diferenças na série original, para assim torná-la estacionária resultando com isto, em reduzir um processo ARIMA à ARMA (auto regressivo de médias móveis), logo com um número menor de parâmetros (apenas p e q), a ser estimado, o que atende de maneira eficaz ao critério da parcimônia, isto é, um modelo mais simples possível, e que represente de forma eficiente o comportamento padrão dos dados analisados.
3. Mediante a análise da função de autocorrelação parcial amostral (FACP) e função de autocorrelação amostral (FAC), as quais são estimadas para identificar o componente autoregressivo e de média móvel, respectivamente, em um processo ARMA (p, q), devem-se obter funções que precisam demonstrar comportamentos similares à suas equivalentes quantidades teóricas.

A escolha da ferramenta adequada, a partir de um modelo de previsão, tal como, o proposto pela metodologia desenvolvida por *Box e Jenkins*, com o intuito de capturar

toda a variabilidade inerente aos dados mediante um modelo ARIMA, consiste em uma tomada de decisão correta com vistas a obter êxito, em um planejamento operacional de qualquer sistema, que se esteja monitorando em um intervalo de tempo contínuo, além de gerar previsões consistentes, isto é, previsões de maior precisão e com a capacidade potencializada de antecipação à cenários futuros.

Na busca de ferramentas capazes de monitorar processos, de maneira cada vez mais precisa e menos suscetível às variações inerentes aos dados observados, sem conseqüentemente, desconsiderar as características particulares de cada variável abordada nos fenômenos em análise, tem-se à teoria dos conjuntos *fuzzy* proposta por *Lotfi Zadeh*, a qual se caracteriza por uma abrangência teórica, a partir de métodos matemáticos, com o auxílio da implementação de critérios subjetivos.

Neste contexto, esta teoria possibilita uma avaliação investigativa e eficiente, sobretudo quando incorporada aos limites alcançados pelos preceitos da metodologia de série temporal, originando uma ferramenta *híbrida* que aborda duas técnicas simultaneamente, o método clássico de previsão mediante série temporal e a teoria *fuzzy*, como observado em Song e Chissom (1993a, 1993b, 1994), Tseng *et al.* (2001), Chen e Chen (2014), Egrioglu (2014), Lu *et al.* (2014), etc.

Song e Chissom (1993a e 1994) apresentaram como proposta de previsão uma série de dados históricos, onde esta série sofre uma abordagem a partir de equações relacionais *fuzzy*, para com isso, conseguir informações suficientes capazes de determinar e avaliar os modelos de série temporal *fuzzy*.

Em contra partida, Tseng e Tzeng (2001) propõem um modelo de previsão mediante o método de regressão linear *fuzzy*, com o intuito de *fuzzificar* os parâmetros *Crisp* obtidos, ou seja, parâmetros estimados mediante a metodologia clássica do modelo SARIMA (sazonal autoregressivo integrado de médias móveis), o que suscita um modelo de previsão sazonal autoregressivo de médias móveis *fuzzy* (SARIMA *fuzzy*).

Conseqüentemente, surgem destas linhas metodológicas descritas anteriormente, duas expressivas correntes de pesquisas acadêmicas, as quais são responsáveis por uma proposta de sinergia, entre os métodos de previsão clássicos e a teoria dos conjuntos *fuzzy*, que são: Tseng e Tzeng (2001, 2002) que propõem uma regressão *fuzzy* via otimização linear, em

contra partida, Song e Chissom (1993a, 1994), Chen (1996), Hwang, *et. al.* (1998), Huarng (2001), Song (2003) defendem a utilização das equações relacionais *fuzzy*.

Segundo Lago Neto (2011) se pode garantir que, Wang e Zhang (1992), tal como, Song e Chissom (1993a, 1994), são os pioneiros quanto ao desenvolvimento das equações relacionais *fuzzy*, mediante pesquisas lideradas pelos mesmos, cujas definições e teoremas estabelecidos na literatura, permitiram o conhecimento teórico de um algoritmo para a identificação de modelos. A partir da metodologia desenvolvida pelos autores supracitados, também se tornou possível a construção de modelos considerando o uso de equações relacionais *fuzzy*, combinadas com raciocínio aproximativo. Lago Neto (2011) afirma ainda, que Song e Chissom (1993a, 1994) suscitaram três casos que motivam a pesquisa científica quanto à abordagem no ambiente *fuzzy*:

Caso 1: Modelar os valores da série temporal $Z(t) \forall t \in T$, onde $T = \{-\infty; \infty\}$, sob a forma particionada mediante conjuntos *fuzzy*, mesmo quando coletados de forma determinística.

Caso 2: Utilizar o conceito de processos estocásticos *fuzzy*, teorizado por Wang e Zhang (1992), que considera os valores da série temporal $Z(t)$ como um conjunto *fuzzy*, mas utiliza o conceito clássico de probabilidade para avaliar a ocorrência de cada um dos valores observados.

Caso 3: Após a aquisição de uma série temporal $Z(t)$, assumir que todos os valores desta série, assim como, as respectivas probabilidades de ocorrência associadas a cada um dos valores desta mesma série $Z(t)$, impreterivelmente em sua plenitude constituem conjuntos *fuzzy*.

Neste sentido, uma questão que desperta real interesse científico, passa por determinar de forma eficiente como avaliar e monitorar os dados de uma série histórica, mediante um modelo estocástico que represente de forma fidedigna o comportamento desta série, a partir da proposta metodológica quanto às equações relacionais *fuzzy*. Por conseguinte, a teoria *fuzzy* inevitavelmente surge como candidata natural, a ser considerada na rotina de modelagem à série de dados históricos de interesse, desde que, os valores envolvidos sejam admitidos como conjuntos *fuzzy*.

Para o caso das observações históricas de $Z(t)$ constituírem uma série temporal, e ainda,

ao se estabelecer duas séries, devido à separação das observações *fuzzy* em relação às probabilidades *fuzzy*, conseqüentemente, a metodologia sugerida por Song e Chissom (1993b) e explorada por Chen (1996), Hwang et. al. (1998) e Huarng (2001), poderá ser adotada sem nenhuma restrição.

Modelar (identificar) um processo no ambiente *fuzzy*, em linhas gerais se torna um grande desafio quando da aquisição de uma série de dados históricos, pois se estará buscando estabelecer algum tipo de associação entre eventos *fuzzy* e probabilidades *fuzzy*, os quais são inerentes ao algoritmo de modelagem dos dados. Toda via, há necessidade de estabelecer uma ferramenta pertencente à metodologia de série temporal, que por sua vez será combinada com a teoria dos conjuntos *fuzzy*. O objetivo é auxiliar a identificação do modelo de série temporal *fuzzy*, com o qual será plenamente possível realizar previsões e monitorar processos observados em intervalos de tempo contínuos, mediante os limites alcançados pela metodologia de série temporal *fuzzy*.

1.3 - Importância e Justificativa da Tese

A importância deste trabalho é caracterizada, em virtude da grande necessidade de determinação de forma eficiente da ordem dos parâmetros das séries temporais não estacionárias, sobretudo, as que são oriundas de processos autoregressivos, tal que, se represente de forma satisfatória o comportamento dos processos em análise, gerados a partir de medições monitoradas, por uma ferramenta capaz de capturar a dependência necessária entre os dados observados.

Neste contexto, este trabalho irá considerar o problema relacionado às séries históricas, que apresentem comportamento autoregressivo sazonal e, assim, com a necessidade de monitoramento por uma ferramenta suficientemente capaz de identificar o modelo de série temporal equivalente, auxiliando desta forma na análise e interpretação dos dados históricos.

Acredita-se que, tal resultado poderá ser alcançado, mediante a combinação da metodologia clássica de série temporal (Box e Jenkins, 1976), com a teoria dos conjuntos *fuzzy* (Zadeh, 1965), proposta inicialmente por Song e Chissom (1993b), seguida por Chen (1996), Huarng (2001), dentre outros.

1.4 - A Hipótese Básica da Tese

Parte-se da hipótese que a função de autocorrelação estendida, quando combinada à função de pertinência *fuzzy* triangular, apresenta uma grande precisão na determinação da ordem dos modelos de série temporal. E ainda, supõe-se que esta função contribui fortemente, para escolha correta de um modelo capaz de gerar previsões com baixíssimo erro e, portanto, de forma mais precisa, em comparação a outras funções de determinação paramétrica disponíveis na literatura.

1.5 - Definições dos Objetivos

1.5.1 - Objetivo Geral

Realizar a identificação dos modelos de séries temporais via definição de número *fuzzy* triangular, incorporada à construção de conjuntos *fuzzy* e pela função de autocorrelação estendida.

1.5.2 - Objetivos Específicos

Os objetivos específicos desta tese são:

1. Estabelecer um modelo de previsão para série temporal *fuzzy*, composto por “*m*” conjuntos *fuzzy* definidos a priori e concebidos mediante a definição de número *fuzzy* triangular;
2. Desenvolver a metodologia conceitual para função de autocorrelação estendida *fuzzy* e adotá-la como ferramenta de determinação da ordem para modelos de série temporal *fuzzy*;
3. Testar a eficiência da função de autocorrelação estendida *fuzzy*, quanto à determinação paramétrica sobre modelos de série temporal *fuzzy*, avaliando o seu desempenho mediante simulações realizadas em vários cenários, isto é, diversos comprimentos de trajetória da série temporal simulada, várias quantidades de conjuntos *fuzzy* durante o processo de *fuzzificação* dos dados, além de variadas combinações de parâmetros nos modelos que forem simulados;
4. Medir o desempenho da função de autocorrelação estendida *fuzzy*, a partir da utilização de métricas para avaliação do erro de previsão, baseado na comparação dos

resultados obtidos com as previsões pelo modelo identificado, em relação aos resultados de previsões disponíveis na literatura, mediante outras ferramentas de identificação consolidadas.

1.6 - Estrutura do Trabalho

Esta tese encontra-se dividida em sete capítulos, a saber:

- Capítulo 1: Apresenta a introdução, importância e justificativa da tese, objetivo geral e específicos, além das limitações do trabalho;
- Capítulo 2: Apresenta o estado da arte sobre modelos de previsão *fuzzy*;
- Capítulo 3: Apresenta a metodologia de sistemas *fuzzy*;
- Capítulo 4: Trata do estudo teórico conceitual contemplado pela série temporal *fuzzy*;
- Capítulo 5: Apresenta as simulações sobre as ferramentas desenvolvidas neste trabalho, quanto a capacidade das mesmas em determinar corretamente a ordem de uma série temporal *fuzzy*;
- Capítulo 6: Aplicação das ferramentas propostas neste estudo para a identificação dos modelos de série temporal *fuzzy*, mediante quatro estudos de caso sobre dados disponíveis na literatura;
- Capítulo 7: Apresenta as conclusões finais da tese e sugestões para trabalhos futuros.

Capítulo 2

O Estado da Arte

Nesta seção discute-se o marco conceitual e a metodologia de análise de série temporal *fuzzy*, desde o seu surgimento até os dias atuais, com o intuito de identificar a fronteira do conhecimento sobre o tema citado acima, mediante os conceitos, técnicas e aplicações que serão considerados como referencial teórico neste estudo. Assim, faz-se um comentário crítico da estrutura dos modelos de previsão, amplamente difundidos na literatura especializada. Posteriormente, trata-se dos conceitos gerais sobre o modelo de previsão *fuzzy*, chegando a propostas de modelagem em série temporal *fuzzy* recentes, para nortear os desdobramentos que serão alcançados no decorrer deste trabalho, a partir das ferramentas que serão propostas para identificar e conseqüentemente auxiliar na melhoria da acurácia das previsões *fuzzy*, a serem realizadas pelo modelo sugerido.

2.1 - Introdução

A modelagem de sistemas dinâmicos tem grande importância para as áreas de economia, engenharia, estatística e física. O termo modelagem não se restringe somente ao aspecto estrutural utilizado no modelo, mas também aos procedimentos de identificação, ajustes e diagnósticos de tais modelos. Nesse contexto, o problema de investigar a associação de diferentes categorias de modelos de monitoramento e estratégia de controle discreto, mediante equações lineares possui uma longa história, tendo sido investigado por Barnard (1959), Whittle (1969), Astrom (1970), Box & Jenkins e MacGregor (1974); Urni (1978, 1979), Darbellay e Slama (2000), Kim *et. al.* (2002) e Conejo *et. al.* (2005).

Em conseqüência, Barnard (1959), sugeriu um modelo interessante para descrever as perturbações em um processo, supondo que o nível local da média do processo μ , está sujeito à ocorrência de saltos aleatórios em intervalos, cujos comprimentos são distribuídos segundo a distribuição Poisson com média θ . A magnitude b de cada salto, é admitida ser normalmente distribuída com média zero e desvio padrão σ_b , permitindo o monitoramento

estatístico do processo, mediante à aplicação de gráfico de controle para sinalizar possíveis ocorrências de causas especiais (que não é fruto da variação natural de um processo).

Contudo, Box e Jenkins (1974), demonstraram que o modelo sugerido por Barnard (1959), constitui um caso particular do modelo ARIMA (0,1,1), enfatizando que o uso de gráficos de controle no monitoramento estatístico do processo, representa uma excelente estratégia para reduzir a variabilidade, quando este apresenta média estável. Caso a média do processo exiba algum tipo de tendência, vários autores, entre eles Montgomery e Friedman (1989), recomendam a utilização de técnicas de teoria de controle.

Especificamente na área de energia elétrica, atribui-se o pioneirismo na utilização de modelos estocásticos a Noel D. Uri, que em 1978, propôs um modelo híbrido combinando análise econométrica e a metodologia de séries temporais, descrita por Box e Jenkins (1974), para prever a ocorrência mensal de picos de energia elétrica, mediante o comportamento de alterações nas variáveis econômicas e climáticas.

Urie Flanagan (1979), utilizando a metodologia de *Box-Jenkins*, identificaram um modelo de previsão para a produção de petróleo e gás natural nos Estados Unidos. A década de 90 foi marcada, principalmente pela sofisticação dos modelos econométricos e das técnicas de previsão de uso final de energia. Hammond e Mackay (1993) se propõem a discutir a estratégia, para identificação de um modelo de previsão com o intuito de projetar a demanda de óleo e gás na Inglaterra, tendo como horizonte o ano de 2010.

Em decorrência das linhas de pesquisa anteriores, Gonzáles Chavez et. al. (1999), utilizou um modelo ARIMA, para construir um modelo de previsão que descrevesse o comportamento da produção e do consumo de energia, na região da Astúrias localizada no norte da Espanha. Por sua vez, Darbellay e Slama (2000), considerando o problema de não linearidade na previsão do consumo de energia elétrica no curto prazo, apresentaram estudos de simulação propondo uma medida de dependência estatística não linear.

A análise de correlação a partir dessa medida permite escolher entre os modelos lineares (ARIMA) e não linear (rede neural), o que melhor representa a evolução da previsão da demanda de eletricidade de curto prazo. Seguindo também a linha de modelagem de sistemas dinâmicos, Cuaresma et. al. (2004), realizou estudos comparativos entre modelos univariados, para estimar o preço diário no mercado de energia elétrica.

Os modelos especificados incluíram modelos autoregressivos, modelos autoregressivos de média móvel e modelos de componentes não observados. Em consequência, Conejo et. al. (2005), realizou estudos comparativos de diferentes técnicas de previsão, para estimar o preço diário no mercado espanhol de energia elétrica. A previsão do preço, tanto no curto, como no longo prazo, são de interesse das empresas e dos consumidores para determinar suas respectivas estratégias de formulação de contratos.

Por sua vez, Lago Neto et. al. (2006), utilizou o modelo autoregressivo sazonal integrável de média móvel - SARIMA, para descrever a estrutura dinâmica da demanda de energia elétrica e do consumo de óleo diesel, em um parque termelétrico existente no Estado do Amazonas, avaliando desta forma, o impacto do custo da aquisição de combustível, na receita operacional da concessionária CEAM projetada para o período 2005 a 2008.

Não por acaso, uma quantidade expressiva de propostas considerando modelos de previsão híbridos, a partir de um modelo de série temporal, passaram a ser propostos em um passado recente na literatura científica. Este fato é justificado em sua maioria, devido à necessidade de aumentar a capacidade preditiva dos modelos de previsão desenvolvidos. Há de se destacar sobretudo, que os modelos de série temporal, seja no ambiente clássico ou não, necessariamente devem atender a algumas suposições básicas (Hyndman e Koehler 2006, Hylleberg 1995).

Neste contexto, Reston Filho et al. 2014, propôs um modelo de previsão via série temporal, sob uma forma híbrida, pois contempla a metodologia pertencente à inteligência artificial, mais precisamente, redes neurais artificiais combinadas com uma série temporal clássica. Observa-se uma proposta de previsão para o preço de energia elétrica de curto prazo, mediante a combinação de dois modelos, inicialmente um autoregressivo de médias móveis, agregado a um modelo de redes neurais artificiais (Reston Filho et al. 2014).

Esta proposta considera dentre outros aspectos, a projeção do preço da energia elétrica para um horizonte de doze passos futuros, referente ao mercado Brasileiro de energia, que segundo o autor supracitado, apresenta um comportamento bem particular, isto é, sem semelhanças com outros mercados consumidores de energia elétrica no mundo. Segundo ainda Reston Filho et al. (2014), os resultados obtidos foram considerados satisfatórios, pois demonstraram que o método proposto possui capacidade suficiente de prever os pontos de sazonalidade, associados a série temporal analisada, além de precisão razoável do

modelo estimado, mediante a medida de desempenho adotada no trabalho, em especial, quando o método de previsão foi comparado a outros modelos de previsão temporal.

2.2 - Modelos de Previsão Fuzzy

Como descrito na introdução, à análise de série temporal constitui um método de previsão amplamente utilizado. Embora possa ser aplicada, a vários problemas de previsão, tal metodologia de predição apresenta inconsistência (imprecisão dos estimadores dos parâmetros ajustados ao modelo de série temporal), quando existe a possibilidade dos valores históricos, serem registrados de forma imprecisa ou quando, os dados são registrados a partir de termos linguísticos.

Uma forma de superar esta limitação consiste em, utilizar o modelo de regressão linear *fuzzy* proposto por Tanaka e Ishibuchi (1992). O modelo constitui a base da estratégia proposta por Tseng e Tzeng (2002), na modelagem da família de modelos autoregressivos integrados de média móvel sazonal *fuzzy*, denotado por $FSARIMA(p; d; q)(P; D; Q)_s$.

Uma outra proposta em discussão no meio acadêmico, destaca que, para modelar uma série temporal, quando as observações históricas são imprecisas ou estão representadas na forma linguística, consiste em utilizar as equações relacionais *fuzzy*. Neste contexto, o pioneirismo deve-se à pesquisa desenvolvida por Wang e Zhang (1992) e por Song e Chissom (1993a, 1994), cujas definições e teoremas, permitiram o desenvolvimento teórico de um algoritmo para a identificação e construção de modelos, mediante o uso de equações relacionais *fuzzy* e raciocínio aproximativo.

Com o mesmo intuito de contornar os problemas de imprecisão dos dados, ou a forma linguística de expressar estes, Song e Chissom (1993a), propuseram um algoritmo de previsão *fuzzy*, composto basicamente por quatro passos ou etapas a serem adotadas:

Etapa 1: Divisão do universo de discurso (U) em intervalos de mesmo tamanho;

Etapa 2: Definição dos conjuntos *fuzzy* no universo de discurso e fuzzificação da série temporal, além da derivação das relações lógicas *fuzzy*, existentes nas informações fuzzificadas;

Etapa 3: Previsão;

Etapa 4: Defuzzificação dos valores preditos.

Song e Chissom defendem o uso de equações relacionais *fuzzy*, para solucionar o pro-

blema de previsão *fuzzy*. A partir dessa proposta, um número crescente de pesquisadores entre eles Sullivan e Woodall (1994), Chen (1996), Hwang *et al.* (1998), Huarng (2001) e Song (2003), vêm aperfeiçoando o algoritmo proposto por Song e Chissom. No mesmo sentido dos pesquisadores citados anteriormente, ou seja, com o objetivo de desenvolver uma alternativa metodológica, para driblar as possíveis imprecisões dos dados observados em uma série temporal, Sullivan e Woodall (1994) propuseram um modelo markoviano, com o intuito de reduzir o esforço computacional inerente ao processo, na etapa correspondente à derivação das relações lógicas *fuzzy*, existentes nas informações fuzzificadas. Chen (1996) revisou o algoritmo de Song e Chissom, propondo um procedimento para simplificar as operações de composição *max-mim*, envolvidas na obtenção da matriz relacional, resultando assim em uma maior exatidão das previsões *fuzzy* realizadas.

Lee e Wang (1996), por sua vez, propôs a utilização de conjuntos *fuzzy* triangulares, para fuzzificar os valores históricos da série temporal. Em seus estudos, Lee destaca que o tamanho do universo de discurso, a quantidade de intervalos, tal como, o tipo de operador utilizado no modelo de projeção *fuzzy*, podem influenciar o resultado das projeções geradas pelos modelos *fuzzy*, com amplitude intervalar constante e arbitrariamente escolhido, desenvolvidos tanto por Song e Chissom (1993a, 1994), como também por Chen (1996).

Contudo, em 1997, Song, em conjunto com outros pesquisadores, desenvolveu estudos teóricos que permitiram a extensão do conceito de série temporal *fuzzy*, para caracterizar um processo dinâmico, quando as observações são necessariamente conjuntos *fuzzy* e as probabilidades, com que se assume a ocorrência de tais valores, também, são valores *fuzzy*.

Este processo foi então denominado, série temporal estocástica *fuzzy* e todos os resultados dessas pesquisas, tem se mostrado encorajadores, à medida que comprovam o bom desempenho dos modelos de série temporal *fuzzy*, em relação aos modelos tradicionais (clássicos) quando há presença de dados imprecisos na série. Porém, um dos principais desafios dessa metodologia, consiste em melhorar significativamente a exatidão dos valores preditos.

Desde Song e Chissom (1993a), vários procedimentos operacionais têm sido propostos para superar o problema quanto à exatidão dos valores preditos, considerando questões relativas a quantidade e a amplitude adequada dos intervalos, passando também pela

condição de ponderação no cálculo das relações *fuzzy*, aos métodos de *fuzzificação* e *defuzzificação* das informações, e à introdução de heurísticas melhorando as projeções *fuzzy*.

Alguns anos depois, Huarng (2001) demonstrou mediante estudos de simulação, que a escolha da amplitude dos intervalos, para caracterizar uma partição do universo de discurso, afeta de forma substancial as previsões *fuzzy*. A partir dessas simulações, este autor propôs dois algoritmos estruturados, inicialmente o primeiro baseado na média e posteriormente o segundo na distribuição do comprimento do valor absoluto, da primeira diferença da série temporal, para determinar a amplitude dos conjuntos *fuzzy* a serem construídos. Huarng, também, introduziu o conceito heurístico de tendência crescente e decrescente, na composição de grupos *fuzzy* lógicos para melhorar a exatidão das previsões.

Corroborando com uma grande corrente de pesquisadores preocupados, com a questão da precisão dos modelos de previsão *fuzzy*, Cardoso (2003) propôs três métodos de previsão para a logística de distribuição de jornais em regiões específicas, onde o primeiro método considerou a metodologia de série temporal clássica, já o segundo método levou em consideração redes neurais artificiais e finalmente, o terceiro procedimento metodológico, abordou a previsão da distribuição dos jornais considerando agrupamentos e regras *fuzzy*.

De acordo com este pesquisador, os dois primeiros métodos (algoritmos) não conseguiram alcançar resultados satisfatórios, diferentemente do terceiro método desenvolvido, que devido as suas características estruturais tornarem possível agregar ao problema de previsão *fuzzy*. O conhecimento do especialista quanto à concepção dos efeitos resultantes do processo em questão, como por exemplo, o dia da semana onde a demanda de jornais é maior ou menor, conseqüentemente este procedimento contribuiu segundo Cardoso (2003), para atender as reais necessidades da logística de distribuição de jornais, mediante um modelo de previsão *fuzzy*.

Um estudo experimental na forma de simulações foi realizado por Vargens et al. (2003), onde o objetivo deste estudo consistiu em desenvolver uma ferramenta capaz, de gerar previsões sobre a produção de cacau de forma precisa e com um custo baixo, com vistas a fornecer informações suficientes a um tomador de decisão, no caso o produtor da cultura em questão, quanto aos procedimentos a serem adotados pelo mesmo, buscando sempre auxiliá-lo como uma ferramenta de apoio para utilização de recursos financeiros na produção de cacau.

Os autores supracitados desenvolveram um sistema fundamentado na lógica *fuzzy*, capaz de calcular as variações da produção de cacau mediante informações que podem interferir diretamente no processo em estudo. Com o diferencial segundo Vargens et al. (2003) de ser um procedimento de baixo custo, em comparação a outros métodos existentes, o especialista, no caso o produtor, não necessariamente precisa possuir conhecimento pleno sobre a técnica desenvolvida, pois segundo os autores o método desenvolvido deve atuar como uma ferramenta de auxílio à tomada de decisão do produtor, algo que foi considerado satisfatório mediante os resultados das simulações realizadas.

Ainda no ano de 2003, ocorreu o surgimento de uma metodologia de previsão *fuzzy* direcionada ao preço de curto prazo da energia elétrica no Brasil, a qual foi baseada mais especificamente em sistemas *neuro-fuzzy* e nos programas de planejamento da operação do sistema elétrico Brasileiro. De acordo com Medeiros (2003), o diferencial e ponto favorável desta metodologia proposta, caracteriza-se pelo fato da mesma obter distribuições estimadas do preço de energia em um curto prazo com uma variabilidade menor, em comparação a procedimentos que considerem apenas o planejamento operacional.

O fato de este autor trabalhar com uma das ferramentas pertencentes aos sistemas *fuzzy*, no caso os sistemas *neuro-fuzzy*, possibilitou incorporar ao sistema, variáveis que influenciam de forma direta o preço da energia elétrica no Brasil, tais como, a energia natural afluyente, energia armazenada, a demanda existente, dentre outras variáveis inerentes ao processo de cálculo do preço da energia elétrica, onde a eficiência do método desenvolvido pelo autor citado acima, foi avaliada em comparação com modelos de redes neurais.

Assim como Cardoso (2003), ou seja, também utilizando agrupamentos e regras *fuzzy*, porém, para o planejamento de sistemas hidroelétricos na previsão de vazões naturais médias mensais, Magalhães (2004) utilizou um algoritmo de agrupamento *fuzzy c-means*, com o intuito de capturar as características estruturais, a variabilidade e reconhecimento de padrões, quanto à similaridade inerente aos dados históricos de interesse, presentes na série temporal com tendência positiva ou negativa. Segundo Magalhães (2004), os resultados das previsões obtidas pelo modelo com agrupamentos *fuzzy*, se mostraram superiores ao serem comparados com modelos de previsão via série temporal clássica apenas, esta última, geralmente é muito utilizada para previsão de vazões no setor hidroelétrico, como se constata na literatura especializada sobre este tema.

Ainda no ano de 2004, um modelo de previsão de carga ativa por barramento foi proposto por Salgado (2004), com o objetivo de auxiliar o planejamento do sistema elétrico quanto à operacionalização deste sistema. Esta proposta consiste em duas metodologias de previsão, onde o autor classificou como previsão individual e previsão agregada, sendo que a primeira considera a série temporal de carga para cada barramento isoladamente, logo, se construiu um modelo de previsão via série temporal para cada um dos barramentos, o que aumenta consideravelmente o esforço computacional desta rotina.

Para a segunda metodologia de previsão, esta é dividida em três etapas, onde a primeira consiste no agrupamento dos barramentos mediante suas similaridades, onde foi utilizado um algoritmo de agrupamento *fuzzy c-means*, já as demais etapas que consistem na previsão de carga, são uma decorrência direta dos resultados da primeira etapa. De acordo com Salgado (2004), o processo de previsão de cargas por barramento foi melhorado com o auxílio do algoritmo de agrupamento, pois mediante os resultados se tornou possível identificar barramentos com necessidade de tratamento específico, algo que não é satisfatório mediante a previsão individual de carga pela primeira metodologia proposta.

Na Busca de implementar sistemas especialistas em controles inteligentes, Bilobrovec (2004) utilizou na rotina de automação de processos a teoria de lógica *fuzzy*, para com isto ser capaz de obter um modelo computacional quanto a previsão *fuzzy*, referente ao comportamento de um motor elétrico em função do nível de tensão que o mesmo foi submetido. Este estudo se baseou na implementação das rotinas computacionais construídas, a partir da agregação das técnicas pertencentes à inteligência artificial, mais especificamente a lógica *fuzzy* em conjunto com o sistema operacional utilizado pelo autor citado acima, tornando possível que os métodos para implementação do sistema especialista fosse alcançado.

A metodologia proposta por Bilobrovec (2004) foi validada segundo o autor, mediante a utilização de dados relacionados a testes de bancada em motores elétricos, onde foi possível garantir total aplicabilidade do controlador *fuzzy* proposto, sobretudo, em sistemas que não possuem uma metodologia de controle inteligente do seu processo.

Tal como, os pesquisadores citados anteriormente indicaram, ou seja, manifestando preocupação com relação ao problema de recorrência e ponderação, nos métodos de previsão de série temporal *fuzzy*, Yu (2005) recomenda a utilização, de pesos distintos para as

relações *fuzzy*, ou ainda, grupos lógicos que possuam maior frequência, refletindo desta forma, a importância individual do grupo ao longo do processo de previsão.

Toda via, Tsaur et al. (2005) destaca a carência de estudos que avaliem, o efeito gerado pela matriz relacional *fuzzy*, sob as projeções, propondo um método analítico, para determinar uma matriz relacional *fuzzy* estável, que melhore a identificação das relações *fuzzy* no procedimento de previsão.

Por sua vez, Pucciarelli (2005) investigou fundamentos teóricos no que diz respeito, a análise, desenvolvimento e implementação de algoritmos para tornar possível modelar dados, referentes a sistemas dinâmicos e para gerar previsões a partir de um modelo de série temporal, combinado com a metodologia *fuzzy* de *Takagi-Sugeno*, de maneira que os conjuntos *fuzzy* do antecedente e os parâmetros do consequente foram estimados, mediante métodos de agrupamentos e identificação paramétrica, logo, como os consequentes são sistemas dinâmicos, resulta em uma menor quantidade de regras *se-então*, o que facilita a identificação do modelo estimado.

Pucciarelli (2005) comparou métodos de agrupamento em três estudos de caso, de onde o mesmo concluiu que o desempenho dos modelos sofreu influência do método utilizado para o agrupamento, assim como os parâmetros do consequente foram influenciados pelo método de agrupamento adotado. Com os resultados obtidos, para Pucciarelli (2005) ficou evidente, que os modelos de série temporal *fuzzy* do tipo *Takagi-Sugeno*, representam ferramentas com resultados satisfatórios, no que diz respeito à concepção de um modelo de série temporal.

No ano seguinte, Leone Filho (2006) com o intuito de prever a carga de energia elétrica de curto prazo, propôs uma rotina de previsão baseada nos conceitos alcançados pela metodologia de série temporal, com o objetivo também de estabelecer um reconhecimento de padrão no fenômeno estudado. Este autor procedeu de forma a realizar várias previsões, para uma série temporal específica, no entanto, a partir de várias ferramentas de previsão combinadas a técnica de série temporal, o que tornou possível a este autor combinar as diversas soluções obtidas, para então se chegar a uma solução com resultado superior, ao encontrado mediante uma única ferramenta.

Dentre as técnicas combinadas a série temporal para gerar previsões, Leone Filho (2006) utilizou redes neurais *fuzzy*, semelhante ao sistema de inferência *fuzzy* não adaptativo. Os

resultados alcançados com a metodologia proposta por este autor se mostraram relevantes, no que diz respeito à confiabilidade e precisão das previsões geradas, pois reduziram os erros de previsão, o que suscitou um modelo de previsão menos suscetível a variabilidade inerente aos dados observados, quando a rede neural *fuzzy* foi incorporada ao modelo de previsão via série temporal.

A justifica segundo Leone Filho (2006), deve-se ao fato da metodologia clássica de série temporal ser adequada a dados lineares, algo que não condiz com a natureza real de uma série referente à carga de energia elétrica, pois esta representa um sistema não-linear na maioria dos casos observados, tal como, já foi comprovado na literatura especializada sobre este tema.

Contudo, Liu (2007) destacou que o desafio da metodologia de previsão *fuzzy*, reside no fato de que suas projeções constituem valores pontuais. Para contornar essa semelhança com a metodologia de *Box-Jenkins* (método clássico de previsão), Liu propôs representar os termos linguísticos de uma variável linguística, a partir da definição de número *fuzzy* trapezoidal, permitindo assim, o uso de operações aritméticas de adição e multiplicação entre os números *fuzzy*, para a obtenção dos valores preditos.

Esta sugestão permite a inserção do conceito de um alfa nível, o que segundo este autor possibilita gerar projeções intervalares. Neste mesmo ano, Singh (2007b) apresentou um novo método de previsão em série temporal *fuzzy*, que considera a diferença sobre os parâmetros do modelo estimado, o qual segundo este autor tem como característica principal a simplicidade de seus procedimentos.

Todavia, para Singh (2007b) o maior problema quanto as previsões em série temporal *fuzzy*, diz respeito a precisão dos valores preditos, que segundo esta proposta de Singh (2007b) pode ser superada por uma aproximação computacional simplificada para as previsões realizadas, tal como, este autor demonstrou ao utilizar dados reais disponíveis na literatura, que serviram para validar a sua proposta e comparar os seus resultados, com de outros pesquisadores que também realizaram previsões mediante modelo de série temporal *fuzzy*.

Posteriormente ao ano de 2007, Cheng et al. (2008b) apresentou a comunidade científica, a proposta de um modelo de expectativa adaptativo incorporado a metodologia de série temporal *fuzzy*, utilizando dados de treinamento referentes ao índice do mercado de Taiwan

(TAIEX), com o objetivo de reduzir os recentes erros de previsão, que estavam assolando investidores quanto a cotação deste índice no mercado financeiro da época.

Neste sentido, o método desenvolvido por este autor incorporou a série temporal *fuzzy*, um modelo de expectativa adaptativo no processo de previsão, para com isto modificar os erros de previsão inerentes ao modelo de série temporal *fuzzy* utilizado, o que se mostrou relevante para o processo de previsão *fuzzy*, via série temporal em comparação a outros métodos apresentados na literatura, tanto para a série temporal do TAIEX, como também para a série da quantidade de matrículas na Universidade do Alabama, no período utilizado pelo autor.

Buscando desenvolver métodos de previsão eficientes em longo prazo referente ao consumo de energia elétrica, Campos (2008) optou por utilizar um grupo de métodos para previsão já consolidados na literatura, como o modelo autoregressivo integrado de médias móveis (linear), modelo autoregressivo integrado de médias móveis (não-linear), modelo de redes neurais artificiais e modelo de redes neurais *fuzzy*.

Desta forma, previsões do consumo de energia elétrica para 60 passos, após o período considerado para estimar os modelos, foram realizadas por este pesquisador, o qual comparou a eficiência e precisão dos modelos a partir de três medidas de acurácia, utilizadas em grande escala na metodologia de série temporal.

Para Campos (2008), a utilização dos quatro modelos citados anteriormente, proporcionou um ganho substancial quanto ao planejamento e tomada de decisão por um possível especialista no setor elétrico, mediante os resultados dos modelos estimados, pois estes se mostraram satisfatórios, quanto à previsão do consumo de energia elétrica, especificamente em um horizonte de 60 meses, tal como, este autor constatou para a cidade de New England e para o Estado de Minas Gerais.

Ainda no ano de 2008, surgiu uma proposta metodológica de previsão que contempla redes neurais *fuzzy*, onde os modelos de sistemas dinâmicos não-lineares complexos, foram obtidos por procedimentos sistemáticos. De acordo com Hell (2008), caso a rede neural *fuzzy* utilize este novo algoritmo de treinamento proposto, os resultados serão mais robustos, desde que, os dados de treinamento representem comportamentos anômalos ou erros gerados na obtenção dos dados.

A metodologia proposta por Hell (2008) auxiliou a construção de modelos de previsão em série temporal e modelagem térmica de transformadores de potência. A robustez dos modelos propostos por Hell (2008) foram ratificadas segundo o autor, assim como, um desempenho superior em termos de precisão das previsões, e também um menor esforço computacional, ao terem seus resultados comparados a outras metodologias disponíveis na literatura sobre o tema em questão.

No ano seguinte, Kai e Wen-Gang (2009) apresentaram um modelo heurístico de série temporal *fuzzy*, com o intuito de realizar previsões, com diferenças de ordem “ n ” na série original, e ainda, o modelo de série temporal *fuzzy* proposto, se baseia na sequência de *fibonacci*, garantindo com isso, que as previsões geradas por este modelo apresentam uma precisão maior do que, se esta intervenção não for realizada.

Devido a série ser não estacionária, segundo Kai e Wen-Gang (2009), isso reduz a carga computacional do processo de modelagem, para posterior análise e monitoramento da série. Mesmo que se trate de uma série de dados históricos não-linear, segundo estes autores, a precisão das previsões é maximizada em relação a não aplicação desta metodologia defendida pelos mesmos.

Chi, et. al. (2010), assim como vários outros pesquisadores, também apresentaram um modelo heurístico de série temporal *fuzzy* para realizar previsões, porém com diferenças de segunda ordem na série original baseada sobre otimização por enxame de partículas. Os resultados empíricos do modelo de série temporal *fuzzy* ajustado apresentaram uma maior eficiência da matriz relacional *fuzzy*, e ainda, uma precisão maior para as previsões do índice de matrículas na Universidade do Alabama, como também forneceu uma plataforma global, em relação aos índices de ações disponibilizados no mercado financeiro, superando assim os modelos disponíveis na literatura.

Por conseguinte, buscando ferramentas para análise, interpretação e monitoramento de dados históricos com a maior precisão possível, Yan-Hua e Li-Xia (2010) apresentam métodos de aplicação de uma série temporal *fuzzy*, investigando as propriedades e métodos de uma série temporal *fuzzy*, tal como, a apresentada por Song e Chissom (1993a, 1994).

A partir da comparação entre a série temporal clássica e a série temporal *fuzzy* mediante suas estruturas teóricas, cálculo, procedimento de previsão e aplicação da série temporal *fuzzy*, Yan-Hua e Li-Xia (2010) estabeleceram parâmetros de comparação entre as duas

metodologias para poder caracterizar quando um processo apresenta melhores resultados em relação ao outro, onde os mesmos concluíram que, se a série de dados históricos for composta por um numero reduzido de observações ou se a série for descrita através de dados linguísticos (variável qualitativa), a série temporal *fuzzy* é aplicável sem restrições ou perda de precisão, ao contrário do que ocorre com o procedimento clássico de série temporal que torna-se impreciso e não aplicável, respectivamente, o que implica na não utilização da metodologia clássica em dados com estas características.

No ano seguinte, Bai et al. (2011) apresenta uma proposta de previsão para bolsa de futuros de Taiwan (TAIFEX) e numero de matrículas na Universidade do Alabama via série temporal *fuzzy*, onde estes autores estabelecem um modelo heurístico com tempo invariante em uma série temporal *fuzzy*, classificada como de grande simplicidade ao utilizar a precisão das previsões do modelo de observações, para o treinamento do preditor da tendência com o intuito de gerar os valores de previsão nas fases de testes do procedimento.

Bai et al. (2011) ressalta que com esta metodologia ocorreu um aumento, na precisão das previsões que foram realizadas, ao comparar os resultados obtidos com outros métodos de previsão por série temporal *fuzzy* já conhecidos, para as duas bases de dados utilizadas por estes pesquisadores, o que levou Bai et al. (2011) a concluir que o aumento da precisão das previsões alcançadas, mediante os métodos que os mesmos propuseram foi significativo, ao se trabalhar com a metodologia de série temporal *fuzzy*, sobretudo, caso o tempo de observação dos dados seja invariante durante o período analisado.

Em um passado recente, Carvalho Jr., et al. (2012) apresentou uma forma de identificação para modelos autoregressivos sazonais, tal que, os parâmetros destes modelos são *fuzzificados*, o que resultou em um modelo de série temporal sazonal *fuzzy*. O autor aplicou “*n*” diferenças nos dados históricos *defuzzificados* das matrículas da Universidade do Alabama nos EUA, entre os anos de 1971 e 1992, como também utilizou a função de autocorrelação inversa para concluir o processo de identificação do modelo. Com isto, foi possível gerar previsões para o número de matrículas, de forma mais consistente, isto é, menos suscetível a variabilidade inerente aos dados, pois segundo esta publicação, uma identificação incorreta compromete todos os resultados decorrentes do modelo de série temporal.

Neste contexto, o autor propôs uma metodologia de identificação do modelo de série

temporal *fuzzy*, com base na estabilização da série (série livre de tendência), a partir de diferenciações na série temporal original com tendência, porém, neste caso combinada à metodologia de conjuntos *fuzzy*. Com a mesma intenção de Carvalho Jr., et al. (2012), isto é, a previsão do número de matrículas na Universidade do Alabama nos EUA, Jasim, et al. (2012), desenvolveu um novo algoritmo para previsão de matrículas baseado em série temporal *fuzzy*. A proposta apresentada por este autor, descreve um método de primeira ordem e com tempo variante do processo, porém, os dados utilizados neste procedimento retratam as matrículas desta universidade, entre os anos de 1948 e 2009. A eficiência do procedimento proposto pelo autor foi avaliada, a partir de previsões realizadas para esta série de dados, em comparação com outras propostas de previsão, que encontram-se disponíveis na literatura especializada.

Com o intuito de propor uma metodologia eficaz quanto a previsões em série temporal *fuzzy*, Egrioglu et al. (2013a), apresentou uma metodologia híbrida que considera redes neurais artificiais e algoritmo genético, incorporados na série temporal *fuzzy* para aumentar a precisão das previsões produzidas.

Desta maneira, um agrupamento *fuzzy* e a metodologia de algoritmo genético foram implementados no processo de *fuzzificação* da série, assim como, as relações *fuzzy* foram estabelecidas com o auxílio de redes neurais artificiais. Neste sentido, segundo Egrioglu et al. (2013a), ao utilizar no processo de *fuzzificação* da série temporal o método de agrupamento e a definição das relações *fuzzy*, com o emprego de redes neurais artificiais, tal como, ao aperfeiçoar o processo de melhoria das previsões com o auxílio de algoritmo genético, se tornou possível observar o aumento significativo na qualidade das previsões.

Assim, as previsões obtidas com a série temporal *fuzzy* resultante deste modelo híbrido, aplicando as técnicas propostas neste trabalho para a série da quantidade de matrículas na Universidade do Alabama nos EUA, que tiveram seus resultados comparados a outras propostas disponíveis na literatura, e que também utilizaram série temporal *fuzzy* para previsão destes mesmos dados entre os anos de 1971 e 1992.

Para Uslu et al. (2014), a metodologia de série temporal *fuzzy* é concebida geralmente em três etapas, no caso a *fuzzificação*, a identificação das relações *fuzzy* e por fim, a *defuzzificação*, as quais indiscutivelmente desempenham um papel relevante na rotina de

previsão *fuzzy*, pois caso pelo menos uma destas etapas seja desempenhada incorretamente, a capacidade preditiva do modelo resultante do processo *fuzzy* será comprometida.

No entanto, segundo estes autores a etapa de *defuzzificação* não desperta expressivo interesse ou pelo menos não possui tantos estudos publicados, em relação às outras duas etapas citadas, motivando Uslu et al. (2014), a considerar a frequência de recorrências das relações *fuzzy* na etapa de *defuzzificação*, por acreditarem que esta característica reflete o comportamento da série temporal. Conseqüentemente, os autores sugerem que se utilize a informação obtida mediante as relações *fuzzy* na etapa de *defuzzificação*. A validação deste método foi realizada com o auxílio de dados reais disponíveis na literatura.

Sadaei et al. (2014), objetivando realizar previsões de curta duração em um ambiente necessariamente *fuzzy*, considerou um algoritmo híbrido em sua concepção estrutural, devido tal estrutura ser composta por um algoritmo *fuzzy* exponencialmente ponderado, combinado com um algoritmo de busca harmônica reforçada, onde inicialmente esta proposta considera o reconhecimento da ordem das relações *fuzzy* lógicas, com o intuito de medir a autocorrelação entre as mesmas.

De acordo com Sadaei et al. (2014), em sua segunda etapa, o método busca determinar os intervalos ideais, assim como, os coeficientes que devem ser utilizados no modelo mediante um conjunto de dados para treinamento, e finalmente, este método utiliza as informações adquiridas nas etapas anteriores para gerar previsões baseado em um horizonte de instante $t = 48$, para previsões de curto prazo.

Esta última etapa segundo Sadaei et al. (2014), se mostrou fundamental para o método atingir êxito, algo que de acordo com estes autores pôde ser comprovado com aplicação dos métodos propostos, em oito estudos de caso reais encontrados na literatura, pois a análise dos erros obtidos no algoritmo *fuzzy*, demonstraram resultados satisfatórios em termos de precisão das previsões.

Também com uma proposta de estrutura híbrida para o modelo de previsão *fuzzy*, Chen e Chen (2014), atentos a etapa de *fuzzificação*, no caso, a partição do universo de discurso, além dos métodos de *defuzzificação* para os dados discretos resultantes, etapas que segundo estes autores influenciam diretamente na acurácia das previsões, assim, uma combinação de discretização por entropia com transformada rápida de Fourier, foram implementadas para contornar estas etapas citadas.

A partir de então, o algoritmo com um novo modelo *fuzzy* de previsão foi estabelecido, para a metodologia de série temporal *fuzzy* segundo Chen e Chen (2014). O método proposto foi validado mediante dados de instituições financeiras disponíveis na literatura, onde foi possível concluir o bom desempenho dos métodos propostos, devido estes terem apresentado resultados superiores aos obtidos por outros métodos amplamente conhecidos, no contexto de previsão via série temporal *fuzzy*.

Foi possível constatar ainda, segundo Chen e Chen (2014), que devido à implementação do método proposto utilizar uma rotina de inicialização, este atualiza de forma incremental a capacidade de previsão do modelo, além de que, o mecanismo de aprendizagem incremental do modelo proposto, se mostrou eficiente mesmo quando foi submetido a grandes conjuntos de dados financeiros.

No ano de 2015, a quantidade de métodos para identificação dos modelos de previsão *fuzzy* continua sendo bastante expressiva na literatura especializada. Abdollahzade et al. (2015), por exemplo, propõem um método híbrido para previsão de séries temporais caóticas e não-lineares, baseado em um modelo *neuro-fuzzy* linear local (NFL), combinado com a análise do espectro singular otimizado (AESO).

Abdollahzade et al. (2015) utilizou dados reais para validação do método proposto, que tratam de séries temporais disponíveis na literatura como: atividade solar e acidentes fatais registrados nos Estados Unidos, dentre outros. Devido às séries caóticas e não lineares possuírem dinâmica e comportamento complexo, a obtenção de previsões por esses modelos geralmente são complicadas, especialmente em relação a um horizonte de previsão muito longo. Segundo Abdollahzade et al. (2015), a análise do espectro singular processa os dados, o que gera uma eliminação dos componentes de ruídos e implica num melhoramento de desempenho das previsões. Assim, mediante um algoritmo de otimização por enxame de partículas, os parâmetros são aperfeiçoados, dando origem em seguida ao modelo de previsão *neuro-fuzzy* linear local.

Com o intuito de prever a carga de energia elétrica em regiões específicas, Efendi et al. (2015), Lee e Hong (2015) e Wang et al. (2015), propõem modelos distintos, porém, com uma característica particular em comum, a utilização de métodos oriundos da chamada inteligência artificial, integrados na estrutura híbrida de seus algoritmos de previsão.

Efendi et al. (2015), sugere uma nova maneira de determinação dos termos linguísticos concebidos sobre a série temporal, para tanto, este utiliza índices numéricos na abordagem linguística. Para compensar a presença do erro de previsão agregado ao processo, este autor considera uma ponderação das relações *fuzzy* lógicas. Os dados da carga diária de energia elétrica da Malásia foram utilizados para validar o método proposto Efendi et al. (2015), segundo o qual, os erros retornados pelo modelo de previsão, foram inferiores aos procedimentos vistos na literatura.

Evidenciando as causas e consequentes efeitos de uma partição equivocada do universo de discurso, onde uma série temporal *fuzzy* está definida, Chen e Chen (2015), além de, Lu et al. (2015), estabelecem estruturas de modelagem híbridas resultantes de métodos de previsão *fuzzy*, que consideram uma forma de partição do universo de discurso, mediante informações granulares nos intervalos *fuzzy*, visto que, segundo estes autores uma partição incorreta compromete invariavelmente o desempenho do modelo de previsão.

Chen e Chen (2015) realizaram previamente uma partição do universo de discurso em “ n ” intervalos, de acordo com uma quantidade de partições pré-estabelecida, posteriormente, as informações granulares são construídas, de acordo com a amplitude dos intervalos observados, estes por sua vez, preservando a tendência (padrão) observada nos valores da série. Em seguida, os intervalos são otimizados de acordo com a amplitude de cada um deles, para que as informações granulares possam ser obtidas, e associadas de forma contínua aos intervalos equivalentes tornando-os informativos.

Com o intuito de contornar a falta de exatidão das funções que determinam a ordem de modelos de séries temporais, e ainda, buscando reduzir o erro de previsão do modelo identificado, Carvalho Jr. e Costa Jr. (2017), propuseram um método de identificação *fuzzy*, que combina a função de autocorrelação estendida à um sistema fuzzy de pertinências triangulares. A estrutura híbrida do modelo aliada à otimização na estimação dos parâmetros do modelo de previsão, contribuiu de forma significativa para a redução dos erros e aumento da precisão das previsões realizadas.

2.3 - Conclusão do Capítulo

Neste capítulo, procurou-se apresentar boa parte do suporte teórico e de aplicações descritos por diversos autores, com pesquisas significativas e com relevância reconhecida

pela comunidade científica, no que diz respeito à questão que envolve a previsão *fuzzy*, ou seja, técnicas de previsão mediante a metodologia de sistemas *fuzzy*, seja ela combinada com as ferramentas de série temporal clássica ou de outras técnicas pertencentes ao contexto alcançado pela inteligência artificial, como por exemplo, as redes neurais e os algoritmos genéticos, além de considerar também, os procedimentos computacionais especialistas, como técnicas de agrupamento e similaridade, suavização de partículas, otimização de sistemas, etc., com o intuito, de tentar contornar o problema de estimação paramétrica inerente aos modelos de previsão *fuzzy*. Assim, se tornou possível perceber uma quantidade relevante de trabalhos direcionados ao tema da previsão *fuzzy*, sobretudo, nos últimos vinte e um anos, quando Song e Chisson (1993a) desencadearam uma busca contínua por métodos e ferramentas, capazes de contornar os erros de previsão presentes em todos os modelos construídos, com o objetivo de prever uma determinada característica de interesse, que neste caso específico é descrita no ambiente *fuzzy*.

A precisão dos modelos de previsão aumentou com o passar dos anos, muito pelo descobrimento de novas técnicas de modelagem, como também pela evolução constante que os sistemas operacionais e *softwares* foram submetidos nos últimos anos. Tudo isso, veio a colaborar com propostas metodológicas cada vez mais precisas e menos vulneráveis, a questão da variabilidade presente em toda base de dados coletada durante um estudo, variabilidade esta que interfere diretamente na precisão das projeções realizadas, como se pode constatar na metodologia de série temporal consolidada na literatura especializada. Portanto, todas as propostas apresentadas neste capítulo, indicam uma forte e expressiva linha de pesquisa voltada para o estudo de metodologias, capazes de se antever aos fatos, no sentido de apresentar ferramentas distintas que combinam diversas técnicas e procedimentos disponíveis no conhecimento científico, no entanto, com um único objetivo que é o de gerar projeções confiáveis, sem desconsiderar as particularidades de cada processo analisado, e ainda, ao combinar as mais diversas ferramentas de previsão com a teoria *fuzzy*, ser possível agregar o conhecimento do especialista ao processo em questão, para com isto, auxiliar de forma significativa uma possível tomada de decisão por um gestor. Neste sentido, mediante as discussões realizadas neste capítulo, se descreverá os capítulos subsequentes, como o da metodologia dos sistemas *fuzzy* que é apresentado no capítulo a seguir.

Capítulo 3

Metodologia dos Sistemas Fuzzy

3.1 - Introdução

Nesta seção serão apresentadas definições sobre a teoria de conjuntos *fuzzy*, além de teoremas e resultados importantes que subsidiam a metodologia *fuzzy*, sobretudo, relacionada aos fundamentos de série temporal que nortearão os desenvolvimentos teóricos, que se discutirá neste trabalho. Partes dos conceitos abordados nesta seção terão, como fonte de pesquisa autores como: Zadeh (1965), Kwakernaak (1978), Dubois e Prade (1980), Klement (1980), Bradshaw (1983), Wang (1997), dentre outros. Os teoremas apresentados nesta seção terão suas provas desenvolvidas, de forma detalhada ou simplesmente citada as referências para obtê-las, ratificando desta forma o conteúdo científico abordado neste trabalho.

3.2 - Teoria dos Conjuntos Fuzzy e Suas Expressões Matemáticas

3.2.1 - Definição dos Conjuntos Fuzzy

Definição 3.1. Seja U um conjunto arbitrário que representa o universo de discurso, tal que, este conjunto é não vazio. Um subconjunto de U , por exemplo, A , logo, o subconjunto *fuzzy* A de U se caracteriza por uma função dada por, $\varphi_A : U \rightarrow \{0; 1\}$, a qual é definida pela Equação (3.1).

$$\varphi_A(x) = \begin{cases} 1, & \text{se } x \in A; \\ 0, & \text{se } x \notin A. \end{cases} \quad (3.1)$$

Portanto, a função $\varphi_A(x)$ é conhecida como função de pertinência do subconjunto *fuzzy* A . Em termos de notação matemática, se pode representar o subconjunto A , a partir da Equação (3.2).

$$A = \{[x, \varphi_A(x); x \in U]\}. \quad (3.2)$$

Observação 3.1. A partir de uma função característica $I_A(x)$, onde, $I_A : U \rightarrow \{0; 1\}$, do conjunto A , tal que, $I_A(x) = 1$, caso $x \in A$ e $I_A(x) = 0$, caso $x \notin A$. Neste sentido, para

a função característica $I_A(x)$ equivalente à função de pertinência $\varphi_A(x)$, ou seja $\varphi_A(x) \rightarrow I_A(x)$, o conjunto A é considerado como um conjunto clássico.

Definição 3.2. Seja U o universo de discurso e A um subconjunto *fuzzy* de U . Portanto, A pode ser definido por uma função de pertinência no domínio de U e contra-domínio no intervalo $[0; 1]$. O suporte do subconjunto A é definido pelo conjunto clássico descrito na Equação (3.3).

$$\text{supp } A = \{X \in U : \varphi_A(x) > 0\}, \quad (3.3)$$

onde, $\text{supp } A$ representar o suporte do subconjunto *fuzzy* A , como se observa em Wang, (1997).

Definição 3.3. Seja U o universo de discurso e A um subconjunto *fuzzy* de U , onde para cada α -nível, tal que, $\alpha \in [0; 1]$, então A representa um conjunto clássico definido pela Equação (3.4).

$$[A]^\alpha = \{x \in U : \varphi_A(x) \geq \alpha\}, \quad (3.4)$$

se $\alpha > 0$ e $[A]^0$ é o fecho do suporte de A ($\text{supp}A$), ou seja, $[A]^0 = \overline{\text{supp } A}$ (ver, Bitar 2009).

Observação 3.2. “O fecho do conjunto A é o conjunto de todos os pontos aderentes à A , ou seja, o conjunto de todos os limites de seqüências de pontos de A ” (Bitar, 2009, pg. 92).

O conjunto *fuzzy* construído sob uma restrição ou limite estabelecido ao domínio do conjunto levando em consideração um valor de α , pode ser admitido como um α -nível. Neste contexto, todos os elementos do domínio de uma função, que possuem valor, ou seja, grau de pertinência $\varphi_A(x)$ no mínimo igual ou superior ao valor de α , indicam o conjunto *fuzzy* resultante do α -nível.

Definição 3.4. Seja A um subconjunto *fuzzy*, este será denominado de número *fuzzy* quando o universo de discurso (conjunto universo) no qual $\varphi_A(x)$ está definida, representar o conjunto dos números reais, desde que, as condições descritas a seguir sejam atendidas.

- (a) $[A]^\alpha = \emptyset$, para $0 \leq \alpha \leq 1$;
- (b) $[A]^\alpha$ é um intervalo fechado em \mathbb{R} (reta real), para $0 < \alpha \leq 1$;

(c) O subconjunto *fuzzy* A deve possuir suporte limitado.

O processo de escolha (determinação) da função de pertinência $\varphi_A(x)$, se torna um dos procedimentos mais importantes dentro da metodologia *fuzzy*, pois, a partir da identificação da mesma os conceitos alcançados por esta ferramenta, remete a teoria matemática para conjuntos *fuzzy* implementada por Zadeh (1965). A escolha da função de pertinência não depende exclusivamente de uma regra que associe uma base de dados, ou uma característica numérica qualquer (variável alatória), a uma função específica, mas sim, ao conhecimento de um especialista quanto ao comportamento dos dados, ou ainda, pode ser escolhida uma função em virtude da complexidade computacional para implementação das suas rotinas algébricas. Dentre as funções de pertinência mais utilizadas na literatura, se pode destacar a função triangular, trapezoidal e gaussiana, como também, a função trivalente, dentre outras.

Definição 3.5. Seja U o universo de discurso da variável linguística $Z_i : i = 1; 2; \dots; n$ ($n \leq T$) e $\varphi(Z_i)$ um número *fuzzy* que representa o conjunto *fuzzy* triangular associado com Z_i . Portanto, $\varphi(Z_i) : U \rightarrow [0; 1]$ será admitido como um conjunto *fuzzy* triangular, caso este seja definido sobre um intervalo $u_j \subset U (j = 1; 2; \dots; m)$, de acordo com a função dada pela Equação (3.5).

$$\varphi(Z_i) = \begin{cases} 0, & \text{se } Z_i \leq a_j; \\ \frac{Z_i - a_j}{b_j - a_j}, & \text{se } a_j < Z_i \leq b_j; \\ \frac{c_j - Z_i}{c_j - b_j}, & \text{se } b_j < Z_i \leq c_j. \\ 0, & \text{se } Z_i > c_j. \end{cases} \quad (3.5)$$

onde, a_j , c_j e b_j são parâmetros escalares que indicam a posição dos vértices e ponto médio (valor da pertinência máxima) do triângulo isóceles, respectivamente. Estes parâmetros estão sobre as extremidades e o ponto central de u_j , respectivamente, e u_j representa o suporte de cada um dos m conjuntos *fuzzy* triangulares associados com Z_i (veja também em Lago Neto et al. 2011). De forma equivalente, se pode resumir a Equação (3.5), como se apresenta na Equação (3.6).

$$f(Z_i; a_j; b_j; c_j) = \max \left[\min \left(\frac{Z_i - a_j}{b_j - a_j}, \frac{c_j - Z_i}{c_j - b_j} \right), 0 \right], \quad (3.6)$$

onde, os parâmetros a_j e c_j são a base do triângulo, tal que, $a_j < c_j$ e o parâmetro b representa o pico do triângulo, onde $a_j < b_j < c_j$.

Observação 3.3. Segundo Bitar (2009), um número *fuzzy* triangular A pode ser representado usualmente mediante a notação $A = (a_j; b_j; c_j)$, sendo a_j ; b_j e c_j as respectivas abscissas dos vértices do triângulo gerado pelo gráfico da função de pertinência dada pela Equação (3.5). Neste sentido, se pode mostrar que $[A]^\alpha = [\alpha \times (b_j - a_j) + a_j; \alpha \times (b_j - c_j) + a_j]$.

Em consequência, para um número *fuzzy* triangular se torna possível agregar a definição de α -nível, que foi definida anteriormente. Portanto, seja $[A]^\alpha$ um intervalo *crisp* (não-*fuzzy*) de um número *fuzzy* triangular, o qual foi obtido mediante a utilização de um α -nível, tal que, para todo α exista $\alpha \in [0; 1]$. Neste sentido, a partir da Equação (3.5), a Equação (3.7) é verdadeira.

$$\frac{[a_j]^\alpha - a_j}{b_j - a_j} = \alpha; \quad \frac{c_j - [c_j]^\alpha}{c_j - b_j} = \alpha. \quad (3.7)$$

Neste sentido,

$$[a_j]^\alpha = [b_j - a_j]\alpha + a_j, \quad \text{como também,} \quad [c_j]^\alpha = -[c_j - b_j]\alpha + c_j.$$

Desta forma, a Equação (3.8) representa o intervalo de α -nível ou intervalo possibilístico.

$$[A]^\alpha = \{[a_j]^\alpha; [c_j]^\alpha\} = \{[b_j - a_j]\alpha + a_j; -[c_j - b_j]\alpha + c_j\}. \quad (3.8)$$

3.2.2 - Conjuntos Fuzzy e Suas Operações Algébricas

Na teoria clássica de conjuntos se tem operações usuais entre os mesmos, tais como, a união, interseção e complementar. Estas operações podem ser aplicadas sobre os conjuntos *fuzzy*, porém, para que isto ocorra se tem uma extensão destas representações algébricas, de maneira que, para a união se tem como extensão o conceito de operador *s-norma*, para a operação de interseção a extensão é o operador *t-norma* e para o complemento a extensão em termos de conjuntos *fuzzy* é o operador complemento, como pode ser verificado em Lago Neto (2011), pg. 16. Neste sentido, se apresenta as definições abaixo.

Definição 3.6. Sejam dois conjuntos *fuzzy* A e B , se pode afirmar que $A = B$, se somente se, $\mu_A(x) = \mu_B(x)$, para todo $x \in U$, onde $\mu_A(x)$ é a função de pertinência do conjunto

fuzzy A e $\mu_B(x)$ é a função de pertinência do conjunto *fuzzy* B , tal como, se observa em Wang (1997).

Definição 3.7. Seja um conjunto *fuzzy* A e $\mu_A(x)$ a função de pertinência deste conjunto. O conjunto *fuzzy* complementar de A em U (universo de discurso) é dado por \bar{A} , tal que, $\mu_{\bar{A}}(x)$ representa a função de pertinência do conjunto *fuzzy* \bar{A} . Portanto, a função de pertinência $\mu_{\bar{A}}(x)$ pode ser obtida a partir da Equação (3.9).

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x). \quad (3.9)$$

Definição 3.8. Sejam os conjuntos *fuzzy* A e B , assim como, as suas funções de pertinência $\mu_A(x)$ e $\mu_B(x)$, respectivamente. O conjunto *fuzzy* união entre A e B em U (universo de discurso) é dado por $A \cup B$. A função de pertinência $\mu_{A \cup B}(x)$ é expressa pela Equação (3.10).

$$\mu_{A \cup B}(x) = \max[\mu_A(x); \mu_B(x)]. \quad (3.10)$$

Definição 3.9. Sejam os conjuntos *fuzzy* A e B , assim como, as suas funções de pertinência $\mu_A(x)$ e $\mu_B(x)$, respectivamente. O conjunto *fuzzy* interseção entre A e B em U (universo de discurso) é dado por $A \cap B$. A função de pertinência $\mu_{A \cap B}(x)$ é dada pela Equação (3.11).

$$\mu_{A \cap B}(x) = \min[\mu_A(x); \mu_B(x)]. \quad (3.11)$$

3.2.3 - Definição de Variável Linguística

Na literatura clássica de conjuntos (*Crisp*), se tem estes como grupo de resultados de uma variável quantitativa (representada necessariamente por números) ou qualitativa (representada por rótulos, categorias ou termos lingüísticos). Uma variável linguística *fuzzy* é necessariamente uma extensão do caso clássico de conjuntos para uma variável qualitativa, com o diferencial de que esta possui os seus rótulos ou categorias descrevendo conjuntos *fuzzy*, contidos em um universo de discurso U e que desta forma, estão relacionados a partir de uma função de pertinência aos possíveis valores que uma variável X , por exemplo, possa assumir neste mesmo conjunto universo U .

Definição 3.10. Se uma variável originalmente numérica possui termos lingüísticos para representar seus valores possíveis, esta variável é classificada como variável lingüística, onde cada palavra, categoria ou rótulo representa um conjunto *fuzzy*, definido no mesmo universo de discurso que a variável em questão.

Definição 3.11. Uma variável lingüística é caracterizada por quatro termos (X , T , U e M), tal como, se descreve abaixo.

Termo 1: X é o nome ou rótulo da variável lingüística, como por exemplo, a intensidade de chuva, a temperatura ambiente, a velocidade de um carro, etc.

Termo 2: T Representa o conjunto de valores (termos) lingüísticos que X pode assumir, como por exemplo, $T = \{\text{muito baixa, um pouco baixa, média, um pouco alta, muito alta}\}$.

Termo 3: U descreve o conjunto universo ou domínio físico que a variável lingüística X , possui seus valores *Crisp* na forma quantitativa, como por exemplo, $U = \{0; \dots; X_{(máximo)}\}$.

Termo 4: M representa a regra semântica para cada categoria de valor lingüístico em T relacionada com um conjunto *fuzzy* em U , como por exemplo, M pode representar uma intensidade de chuva “muito baixa”, “um pouco baixa”, “média”, “um pouco alta”, “muito alta”, com a função de pertinência associada a cada uma destas categorias, a partir do valor que a variável X assume em U .

3.2.4 - Definição das Regras Fuzzy

A partir de uma associação com o raciocínio humano e raciocínio lógico, se torna possível estabelecer regras para os sistemas *fuzzy*, tais como, “se” uma quantidade X de água é aquecida até atingir a temperatura de cem graus célsius, em condições normais de temperatura e pressão, “então” a água entrará em ebulição. Neste sentido, o conceito de proposição *fuzzy* se estabelece dando origem as chamadas regras *fuzzy* “se-então”, como está representado na Equação (3.12).

$$SE < \text{proposição fuzzy} > ENTÃO < \text{proposição fuzzy} > . \quad (3.12)$$

Assim, compreender as regras “se-então” depende do entendimento sobre proposição *fuzzy*.

Definição 3.12. Uma proposição *fuzzy* pode ser caracterizada como “simples”, caso esta expresse uma frase do tipo: x é A , tal que, x é uma variável lingüística e A retrata um valor lingüístico de x , onde, A é um conjunto *fuzzy* definido no domínio físico de x .

3.2.5 - Definição de Medidas Fuzzy

Segundo Ortega (2001), o conceito de medida *fuzzy* foi estabelecido por Sugeno (1974). Este conceito consiste em uma forma natural para mensurar incertezas, sobretudo quando existe uma subjetividade inerente aos valores utilizados, durante o processo de medição das quantidades de interesse, consideradas por um especialista.

Definição 3.13. Sejam Ω o espaço amostral (conjunto não-vazio), \mathfrak{S} uma σ - álgebra e μ uma função, onde, $\mu : \mathfrak{S} \rightarrow [0; 1]$ é uma medida *fuzzy*, desde que, satisfaça as condições a seguir.

Condição 1: $\mu(\emptyset) = 0$ e $\mu(\Omega) = 1$;

Condição 2: $\mu(A) \leq \mu(B)$, se $A \subseteq B$, para todo, $A \in \mathfrak{S}$ e $B \in \mathfrak{S}$;

Condição 3: Para qualquer sequência crescente de conjuntos mensuráveis, dada por A_n , tal que, $n = 1; 2; \dots$, se pode garantir que a Equação (3.13) é atendida.

$$\lim_{n \rightarrow \infty} \mu(A_n) = \mu \left(\lim_{n \rightarrow \infty} A_n \right). \quad (3.13)$$

Portanto, a terna $(\Omega, \mathfrak{S}, \mu)$ é denominada frequentemente de espaço de medida *fuzzy*.

3.2.6 - Definição de Fuzzificador e Defuzzificador

Durante parte do processo de utilização da metodologia *fuzzy*, se tem um sistema composto de uma máquina de inferência *fuzzy*, combinada com regras pré-estabelecidas dentro de uma base de regras *fuzzy*, enquanto se realiza a rotina de procedimentos do sistema mediante um conjunto *fuzzy* X em U , até um conjunto *fuzzy* Y em V . Tem-se assim um sistema *fuzzy*, com a máquina de inferência *fuzzy* condicionada à base de regras *fuzzy*, atuando como um “conectivo”, entre a *fuzzificação* (variável de entrada X) e a *defuzzificação* (variável de saída Y), para que a rotina do processo de implementação deste sistema seja concluída de forma satisfatória.

Definição 3.14. Seja $\mathbf{x}^* \in U \subset \mathbb{R}^n$ um valor real pontual de um conjunto *fuzzy* A' em U , então o *fuzzificador* é definido como um mapeamento sobre \mathbf{x}^* .

Segundo Wang (1997), o processo de *fuzzificação* é regido por três etapas apresentadas a seguir.

Etapa 1:

O processo de *fuzzificação* considera o fato que a variável de entrada é um ponto \mathbf{x}^* *Crisp*, onde o conjunto *fuzzy* A' apresenta um alto valor de pertinência para \mathbf{x}^* ;

Etapa 2:

Se a variável de entrada para o sistema *fuzzy* é influenciada por um sinal de ruído (erro), então é desejável que o *fuzzificador* apresente capacidade suficiente de ajudar a suprimir este ruído, ao longo do desenvolvimento do sistema *fuzzy*;

Etapa 3:

O *fuzzificador* deve contribuir de maneira a simplificar os cálculos envolvidos na máquina de inferência *fuzzy* e, por conseguinte reduzir o esforço computacional envolvido na implementação do sistema *fuzzy*.

Fuzzificador Triangular:

O *fuzzificador* triangular desempenha um mapeamento em $\mathbf{x}^* \in U$, a partir de um conjunto *fuzzy* A' em U , o qual apresenta função de pertinência dada pela Equação (3.14).

$$\mu_{A'}(\mathbf{x}) = \begin{cases} \left(1 - \frac{|\mathbf{x}_1 - \mathbf{x}_1^*|}{b_1}\right) \star \cdots \star \left(1 - \frac{|\mathbf{x}_n - \mathbf{x}_n^*|}{b_n}\right), & \text{se } |\mathbf{x}_i - \mathbf{x}_i^*| \leq b_i, i = 1; 2; \cdots; n. \\ 0, & \text{caso contrário.} \end{cases} \quad (3.14)$$

onde, b_i representa parâmetros positivos envolvidos no processo e “ \star ” descreve um operador *t-norma*, que é geralmente utilizado como produto algébrico ou mínimo.

Ortega (2001) defende que a *defuzzificação* é um procedimento, que permite interpretar a distribuição de possibilidades da saída de um modelo lingüístico *fuzzy* de forma quantitativa, ou seja, a *defuzzificação* fornece um valor numérico representativo que captura o significado essencial dessa distribuição de possibilidades.

Definição 3.15. A *defuzzificação* é definida como um mapeamento de um conjunto *fuzzy* B' sobre $V \subset R$, o que representa a saída da máquina de inferência *fuzzy*, para um ponto *Crisp* $y^* \in V$.

A rotina de *defuzzificação* de um número caracteriza, um ponto específico em que V melhor representa o conjunto *fuzzy* B' , segundo defende Wang (1997) e ainda, segundo este autor, isso equivale ao valor médio de uma variável aleatória. Ao se optar por uma rotina de *defuzzificação*, dentre as disponíveis na literatura, três critérios citados a seguir devem ser considerados.

Critério 1: Plausibilidade. O ponto y^* deve representar o conjunto *fuzzy* B' , de um

ponto de vista intuitivo, como por exemplo, o ponto y^* pode estar localizado aproximadamente no ponto médio do suporte do conjunto *fuzzy* B' , ou tendo um alto grau de pertinência em relação ao conjunto *fuzzy* B' ;

Critério 2: Simplicidade Computacional. Este critério é particularmente importante em controle *fuzzy*, devido os controladores *fuzzy* operarem em tempo real;

Critério 3: Continuidade. Uma pequena mudança no conjunto *fuzzy* B' , não deve resultar em uma grande mudança em y^* .

A seguir, é apresentado o tipo de *defuzzificação* utilizado neste trabalho.

Defuzzificação Centro de Gravidade:

A *defuzzificação* centro de gravidade, também denominada de centróide, especifica o y^* , como o centro de uma área coberta por uma função de pertinência do conjunto *fuzzy* B' , tal que, para o caso da variável Y ser discreta ou contínua, y^* é dado pela Equação (3.15) e (3.16), respectivamente.

$$y^* = \frac{\sum_y y \mu_{B'}(y)}{\sum_y \mu_{B'}(y)} \quad (3.15)$$

e

$$y^* = \frac{\int_V y \mu_{B'}(y) dy}{\int_V \mu_{B'}(y) dy}, \quad (3.16)$$

sendo que, \int_V representa uma integral convencional.

Caso se admita $\mu_{B'}(y)$, como sendo uma função densidade de probabilidade de uma variável aleatória, então a *defuzzificação* centro de gravidade resulta em um valor médio da variável aleatória. Há momentos em que se deseja eliminar a relação $y \in V$, desde que, o valor de pertinência ao conjunto *fuzzy* B' seja muito pequeno, o que leva a uma *defuzzificação* centro de gravidade indexada, como dado na Equação (3.17).

$$y^* = \frac{\int_{V_\alpha} y \mu_{B'}(y) dy}{\int_{V_\alpha} \mu_{B'}(y) dy}, \quad (3.17)$$

onde, V_α é dado pela Equação (3.18).

$$V_\alpha = \{y \in V \mid \mu_{B'}(y) \geq \alpha\}. \quad (3.18)$$

tal que, α é uma constante.

A *defuzzificação* centro de gravidade apresenta como ponto positivo, o fato de possuir uma plausibilidade intuitiva, além de uma continuidade comprovada, porém, exige um esforço computacional elevado para o seu cálculo, o que pode ser justificado, devido a função de pertinência $\mu_{B'}(y)$ ser irregular, pelo fato das Equações (3.16) e (3.17) apresentarem integrais de resoluções complexas e conseqüentemente com dificuldades computacionais para suas implementações.

3.2.7 - Intervalo de Confiança Fuzzy

Seja X uma variável aleatória tendo uma distribuição com parâmetros $\theta_1, \dots, \theta_n$. Dado \tilde{X} como sendo uma variável aleatória *fuzzy*. Então, \tilde{X}_h^L e \tilde{X}_h^U são variáveis aleatórias para todo $h \in [0; 1]$. Segundo Wu (2009), \tilde{X} possui a mesma distribuição de X com parâmetros $\tilde{\theta}_1, \dots, \tilde{\theta}_n$, se \tilde{X}_h^L e \tilde{X}_h^U tem a mesma distribuição como X com parâmetros $[\tilde{\theta}_1]_h^L, \dots, [\tilde{\theta}_n]_h^L$ e $[\tilde{\theta}_1]_h^U, \dots, [\tilde{\theta}_n]_h^U$, respectivamente.

Seja X_1, \dots, X_n variáveis aleatórias independentes e identicamente distribuídas. Dado $L(\mathbf{X})$ e $U(\mathbf{X})$ como sendo duas estatísticas, tal que, $L(\mathbf{X}) \geq U(\mathbf{X})$, onde $\mathbf{X} = (X_1, \dots, X_n)$. Caso o intervalo aleatório $[L(\mathbf{X}), U(\mathbf{X})]$ satisfaz,

$$P_\theta\{L(\mathbf{X}) \geq \theta \geq U(\mathbf{X})\} = 1 - \alpha,$$

então $[L(\mathbf{x}), U(\mathbf{x})]$ é um intervalo de confiança para θ , com um coeficiente de confiança igual a $1 - \alpha$, onde $\mathbf{x} = (x_1, \dots, x_n)$ e cada x_i é o valor observado de X_i para $i = 1, \dots, n$.

Dado $\tilde{X}_1, \dots, \tilde{X}_n$ como sendo variáveis aleatórias independentes e identicamente distribuídas com parâmetro *fuzzy* $\tilde{\theta}$. Seja \tilde{x}_i o valor observado de \tilde{X}_i , para $i = 1, \dots, n$ onde cada \tilde{x}_i é um número *fuzzy* canônico para $i = 1, \dots, n$. Portanto, verifica-se que $[\tilde{x}_i]_h^L$ e $[\tilde{x}_i]_h^U$ são os valores observados de $[\tilde{X}_i]_h^L$ e $[\tilde{X}_i]_h^U$, respectivamente, para todo $h \in [0; 1]$. Desta forma, é possível concluir que: $[\tilde{X}_i]_h^L \equiv \tilde{X}_{ih}^L$ e $[\tilde{X}_i]_h^U \equiv \tilde{X}_{ih}^U$. Neste contexto, pode-se afirmar segundo Wu (2009), que $\tilde{X}_{1h}^L, \dots, \tilde{X}_{nh}^L$ são variáveis aleatórias independentes e identicamente distribuídas, e $\tilde{X}_{1h}^U, \dots, \tilde{X}_{nh}^U$ são variáveis aleatórias independentes e identicamente

distribuídas. Por conseguinte, segue que $\tilde{\mathbf{x}}_h^L = (\tilde{x}_{1h}^L, \dots, \tilde{x}_{nh}^L)$ e $\tilde{\mathbf{x}}_h^U = (\tilde{x}_{1h}^U, \dots, \tilde{x}_{nh}^U)$. Então, para dois grupos de valores observados $\tilde{x}_{1h}^L, \dots, \tilde{x}_{nh}^L$ e $\tilde{x}_{1h}^U, \dots, \tilde{x}_{nh}^U$, se torna possível construir um intervalo de confiança $[L(\tilde{\mathbf{x}}_h^L), U(\tilde{\mathbf{x}}_h^L)]$ para $\tilde{\theta}_h^L$ com um coeficiente de confiança $1 - \alpha$ e um intervalo de confiança $[L(\tilde{\mathbf{x}}_h^U), U(\tilde{\mathbf{x}}_h^U)]$ para $\tilde{\theta}_h^U$ com um coeficiente de confiança $1 - \alpha$.

Mediante as definições estabelecidas acima, se chega aos intervalos de confiança aproximados, mediante o teorema central do limite (TCL), com coeficiente de confiança dado por $1 - \alpha$, onde $0 \leq \alpha \leq 1$, de acordo com as Equações (3.19) e (3.20).

$$\text{I.C}(\tilde{\mu}_h^L; 95\%) = \left[\bar{x}_h^L - Z_\alpha \times \left(\frac{s_{hL}}{\sqrt{n}} \right); \bar{x}_h^L + Z_\alpha \times \left(\frac{s_{hL}}{\sqrt{n}} \right) \right] \equiv [L(\tilde{\mathbf{x}}_h^L); U(\tilde{\mathbf{x}}_h^L)] \quad (3.19)$$

e

$$\text{I.C}(\tilde{\mu}_h^U; 95\%) = \left[\bar{x}_h^U - Z_\alpha \times \left(\frac{s_{hU}}{\sqrt{n}} \right); \bar{x}_h^U + Z_\alpha \times \left(\frac{s_{hU}}{\sqrt{n}} \right) \right] \equiv [L(\tilde{\mathbf{x}}_h^U); U(\tilde{\mathbf{x}}_h^U)], \quad (3.20)$$

onde, \bar{x}_h^L e \bar{x}_h^U representam a média entre os valores mínimo e máximo dos vértices das bases dos triângulos, respectivamente.

A partir da definição de número *fuzzy* triangular que foi utilizada neste trabalho; o valor Z_α representa o grau de confiança com o qual o intervalo será construído, em função de um nível de significância α , e ainda, este termo indica o valor tabelado da distribuição Normal(0; 1), tal que, para cada $0 \leq \alpha \leq 1$, há um valor tabelado correspondente de Z_α , como por exemplo, se $\alpha = 0,05$, então $Z_\alpha = 1,96$, assim, o grau de confiança do intervalo construído é de 95%.

Os termos s_{hL} e s_{hU} equivalem ao desvio padrão para os valores dos vértices mínimo e máximo, respectivamente, das bases dos triângulos, os quais correspondem aos conjuntos *fuzzy* construídos, sendo que, como as bases dos triângulos formados no processo de *fuzzificação* são equidistantes, isso implicará em $s_{hL} = s_{hU}$, pois, apesar de que, em cada base dos triângulos os valores mínimo e máximo são diferentes dos demais, a amplitude (diferença entre os valores dos vértices mínimo e máximo) será sempre a mesma, logo a variância e por conseguinte o desvio padrão será constante em todas as bases dos triângulos; o termo n é o número de intervalos *fuzzy*, que foram construídos na rotina de *fuzzificação* dos dados de interesse.

3.3 - Conclusão do Capítulo

Neste capítulo, foi descrita a teoria matemática e conceitual que fundamenta a metodologia dos sistemas *fuzzy*, a qual irá sustentar o desenvolvimento das ferramentas propostas neste trabalho, mediante definições, teoremas, corolários e lemas, estabelecidos na literatura. Todavia, o algoritmo de identificação híbrido que justifica, além de ser o objeto principal da proposta de Tese elaborada neste estudo, irá considerar em sua estrutura física uma parte proveniente dos sistemas *fuzzy*, fomentada plenamente pelo referencial teórico apresentado neste capítulo, combinado com a metodologia clássica de série temporal a ser discutida na sequência. Isto suscitará as proposições de modelagem e identificação *fuzzy* temporal, mediante os respectivos resultados encontrados no desenvolvimento deste estudo, sob a forma simulada ou experimental.

Capítulo 4

Série Temporal Fuzzy

4.1 - Introdução

Apesar da metodologia de *Box-Jenkins* apresentar a vantagem de produzir previsões precisas para curto prazo, ela exige um número mínimo de 50 observações na série e preferencialmente 100 observações, para obter previsões confiáveis em um modelo sazonal autoregressivo integrado de médias móveis (SARIMA). Em situações práticas, o ambiente é incerto e muda rapidamente, além disso, o analista é obrigado geralmente a prever situações usando, uma pequena quantidade de informações sobre um fenômeno que se deseja conhecer melhor o seu comportamento.

Uma forma de contornar essa limitação consiste em utilizar o modelo de regressão linear *fuzzy* proposto por Tanaka e Ishibuchi (1992). Este modelo constitui a base para uma estratégia, proposta por Tseng e Tzeng (2002), a qual trata da modelagem de uma família de modelos SARIMA *fuzzy*, denotados por $FSARIMA(p, d, q)(P, D, Q)_S$.

Outra proposta em discussão, no meio acadêmico, resulta em modelar séries temporais quando as observações históricas são imprecisas, ou representadas na forma lingüística, e que necessariamente consiste na utilização de equações relacionais *fuzzy*.

A partir dos estudos iniciais implementados por Song e Chissom (1993a, 1993b e 1994), vários pesquisadores têm concentrado seus esforços, para o desenvolvimento de diferentes métodos (algoritmos) de previsão *fuzzy*. Neste sentido, este capítulo destina-se a apresentação e discussão da metodologia dos modelos de previsão, para série temporal mediante a metodologia *fuzzy*, tal como, se apresentam nas Seções 5.2 e 5.3, onde se encontram os conceitos, definições e algoritmos utilizados por *Song e Chissom*.

Também são apresentadas propostas de outros autores com pesquisas e desenvolvimentos teóricos, considerados relevantes sobre este tema na literatura especializada, face à possibilidade de caracterizar uma série temporal *fuzzy*.

A Seção 5.4 trata de métodos propostos na literatura, com o intuito de identificar a ordem dos modelos de série temporal *fuzzy*, passando pela formalização das medidas de dependência entre conjuntos *fuzzy*. Uma proposta inédita para identificação e previsão de modelos autoregressivos *fuzzy*, se observa na Seção 5.5.

4.2 - Metodologia Sobre Série Temporal Fuzzy

Um modelo de série temporal *fuzzy* foi apresentado por Tanuwijaya e Chen (2009) e Yanhua e Lixia (2010). Além disso, Song e Chissom (1993a) descrevem o que representa um modelo auto regressivo de primeira ordem, o qual assume a caracterização dada pela Equação (4.1),

$$Z(t + 1) = R(t + 1; t) \circ Z(t), \quad (4.1)$$

onde, $Z(t)$ e $Z(t + 1)$ representam os valores observados de uma série temporal *fuzzy* $\{Z(t)\}$, no instante t e $t + 1$, respectivamente, no entanto, estes valores observados $Z(t)$ e $Z(t + 1)$ caracterizam conjuntos *fuzzy*, os quais são determinados a partir de um *universo de discurso*, $R(t + 1; t)$ define a relação *fuzzy* utilizada e “ \circ ” indica o operador “Max-Mim” ou outro operador expressivo utilizado (ver, Boukezzoula et al. 2007).

De maneira generalizada, se pode representar a Equação (4.1), sob a forma da Equação (4.2),

$$\begin{aligned} Z(t + 1) = & R(t + 1; t) \circ Z(t) \cup R(t; t - 1) \circ Z(t - 1) \\ & \cup \dots \cup R(t - k + 1; t - k) \circ Z(t - k), \end{aligned} \quad (4.2)$$

onde, \cup representa o operador união e $k > 1$ é um inteiro, tal como, visto em Yang, et. al. (2010). Segundo, Chi, et. al. (2010), desde os primeiros modelos de séries temporais propostos por Song e Chissom (1993a), Song e Chissom (1993b), Song e Chissom (1994), com base nos trabalhos de Zadeh (1965), Zadeh (1975a), Zadeh (1975b) e Zadeh (1975c), numerosos estudos têm sido realizados para melhorar a precisão das previsões e reduzir a sobrecarga computacional.

A classe de modelos para descrever uma série temporal *fuzzy* se assemelha à análise de série temporal clássica, tal que, o objetivo é identificar uma família de modelos de série temporal *fuzzy*. Como mencionado acima existe mais de uma metodologia para modelar uma série temporal *fuzzy*.

Nesta seção, apresenta-se a proposta que utiliza equações relacionais *fuzzy* na obtenção de tais modelos. Existem vários métodos distintos para determinar uma matriz de relação

fuzzy desenvolvidos principalmente por Dubois e Prade (1991), Wu (1986) e Mamdani (1977). Como o método de *Mamdani* apresenta simplicidade computacional em obter as relações *fuzzy*, se faz a opção por este método.

Sejam I e J conjuntos de índices para a variável linguística $F(t-1)$ e $F(t)$ respectivamente. Neste sentido, as definições a seguir são necessárias, para que se torne razoável uma formalização dos conceitos pertinentes, a metodologia desenvolvida em relação a uma série temporal *fuzzy*.

Definição 4.1. Seja a relação de conjuntos *fuzzy* caracterizada por P_i , onde $i = 1; \dots; r$, $\bigcup_{i=1}^r P_i = \Omega$, uma partição sobre Ω , L_i , onde $i = 1; \dots; r$ com r termos linguísticos. Então $W = \mu_1; \mu_2; \dots; \mu_r$ e $K = v_1; v_2; \dots; v_r$, são conjuntos *fuzzy*, para toda relação *fuzzy* entre W e K , que é definida sob a forma $R = W^T \circ K = [R_{ij}]_{r \times r}$, onde μ_i e v_j são os valores de pertinência em relação à W e K respectivamente, tal que, T representa a operação de transposição e $R_{ij} = \min(\mu_i; v_j)$.

Definição 4.2. Dado R uma reta real, para todo $X(t) \subset R$, e $t = \dots; 0; 1; 2; \dots, n$, o universo de discurso definido para o conjunto *fuzzy* $f_i(t)$. Seja $F(t)$ sobre $f_i(t)$, onde $i = 1; 2; \dots; n$, definido como uma série temporal de $X(t)$ para todo t . $F(t)$ pode ser uma variável linguística, onde $f_i(t)$, com $i = 1; 2; \dots; n$, representa possíveis valores linguísticos de $F(t)$ (Jasim et al. (2012), Sah and Degtiarev (2005), Chen and Chung (2006)).

Definição 4.3. Dado $F(t)$ caracterizado por $F(t-1)$, tal que, $F(t-1) \rightarrow F(t)$, para todo t . Esta relação é denotada por $F(t) = F(t-1) \circ R(t, t-1)$, onde “ \circ ” é um operador. E ainda, $R(t; t-1)$ é uma relação *fuzzy* entre $F(t)$ e $F(t-1)$, para todo t , denominada modelo de primeira ordem de $F(t)$ (Chen e Hsu (2004), Nasser et al. (2008)).

Definição 4.4. Seja $F(t-1) \rightarrow A_i$ e $F(t) \rightarrow A_j$, para todo $t, j = 0; 1; 2; \dots; n$. Então, a relação $F(t) \rightarrow F(t-1)$ é uma relação lógica $A_i \rightarrow A_j$ (Sah e Degtiarev (2005)).

Definição 4.5. Dado $R(t; t-1)$ um modelo de primeira ordem de $F(t)$. Para todo t , $R(t; t-1) = R(t-1; t-2)$, logo $F(t)$ é uma série temporal *fuzzy* invariante. Caso contrário, $F(t)$ é uma série temporal *fuzzy* variante (Chen and Chung (2006)).

Definição 4.6. Seja $F(t)$ uma série temporal *fuzzy*, para todo $t = 0; 1; \dots$. Dado T o tamanho de trajetória da série temporal, tal que $F(t-1) = A_i$ e $F(t) = A_j$. Então, a

relação *fuzzy* lógica (RFL): $F(t-1) \rightarrow F(t)$, é descrita por $A_i \rightarrow A_j, i, j = 1; 2; \dots; p$, tal que p é o número de intervalos ou subintervalos (lado esquerdo, LHS), e A_j lado direito (RHS) de RFL (Efendi et al. (2015)).

Definição 4.7. Dado $A_i \rightarrow A_j, A_i \rightarrow A_k, \dots, A_i \rightarrow A_p$, como relações *fuzzy* lógicas (RFLs) de mesmo termo das LHS. O agrupamento de lógicas *fuzzy* ordenadas (LFG), colocando todas as RFLs juntas em RHS do LFG. Então,

$$A_i \rightarrow A_j, A_i \rightarrow A_k, \dots, A_i \rightarrow A_p; \quad i, j, k, \dots, n \quad (n \leq T),$$

para todo T igual a trajetória da série temporal. (Efendi et al. (2015))

4.2.1 - Modelo Fuzzy Autoregressivo de ordem 1 descrito em “Song et al. (1997)”

Para todo $\mu_j[Z(t)] \in F(t)$, onde, $j \in J, \exists \mu_i[Z(t-1)] \in F(t-1)$, e $i \in I$, tal que, exista:

- (a) Uma relação *fuzzy* $R_{ij}(t-1; t)$;
- (b) $\mu_j[Z(t)] = \mu_i[Z(t-1)] \circ R_{ij}(t-1; t)$, onde o operador “ \circ ” representa a composição de qualquer $s-t$ norma (em particular Mamdani “Max-Mim”).

Então, se para qualquer instante t , a variável lingüística $F(t)$ é influenciada apenas pela variável lingüística $F(t-1)$, tem-se um modelo *fuzzy* autoregressivo (FAR) de ordem 1, ou seja, “FAR(1)”, denotado por $\mu_i[Z(t-1)] \rightarrow \mu_j[Z(t)]$ ou de forma equivalente $F(t-1) \rightarrow F(t)$.

Equações relacionais *fuzzy* estabelecidas por “Song et al. (1997)”. Seja $F(t)$ um modelo *fuzzy* autoregressivo de ordem 1. E ainda, dada a Equação (4.3),

$$R(t; t-1) = \bigcup_{i,j} R_{ij}(t; t-1), \quad (4.3)$$

tal que, \bigcup representa uma s -norma. Neste sentido, $R(t; t-1)$ é denominada de matriz de relação *fuzzy* entre $F(t)$ e $F(t-1)$, que resulta na equação relacional *fuzzy* expressa pela Equação (4.4).

$$F(t) = F(t-1) \circ R(t; t-1). \quad (4.4)$$

Seja $F(t)$ uma série temporal *fuzzy*. Admitindo-se a existência de duas equações relacionais *fuzzy*, tais como, se observa nas Equações (4.5) e (4.6),

$$R_1(t; t-1) = \bigcup_{i,j} R_{ij}^1(t; t-1) \quad (4.5)$$

$$e \quad R_2(t; t-1) = \bigcup_{i,j} R_{ij}^2(t; t-1), \quad (4.6)$$

entre $F(t)$ e $F(t-1)$. Se para todo, $\mu_j[Z(t)] \in F(t)$, onde, $j \in J$, existe um, $\mu_i[Z(t-1)] \in F(t-1)$, onde $i \in I$, assim como existem as relações *fuzzy* apresentadas nas Equações (4.7) e (4.8),

$$R_{ij}^1(t; t-1) \quad (4.7)$$

$$e \quad R_{ij}^2(t; t-1), \quad (4.8)$$

tal que, se chega as expressões dadas pelas Equações (4.9) e (4.10),

$$\mu_j[Z(t)] = \mu_i[Z(t-1)] \circ R_{ij}^1(t; t-1) \quad (4.9)$$

$$e \quad \mu_j[Z(t)] = \mu_i[Z(t-1)] \circ R_{ij}^2(t; t-1). \quad (4.10)$$

Por conseguinte, se diz que, as Equações (4.11) e (4.12) são consideradas iguais.

$$R_1(t; t-1) \quad (4.11)$$

$$e \quad R_2(t; t-1). \quad (4.12)$$

Segundo Huarng (2001), ao assumir que o conjunto *fuzzy* $F(t)$ é influenciado unicamente por:

$$F(t-1) \quad e \quad F(t) = F(t-1) \circ R(t; t-1),$$

para qualquer instante de tempo t , se $R(t; t-1)$ é invariante sob translação do tempo, então, dada a Equação (4.13),

$$R(t_1; t_1-1) = R(t_2; t_2-1), \quad (4.13)$$

esta é verdadeira e conseqüentemente, $F(t)$ é definido como uma série temporal *fuzzy* invariante.

Para Chen (1996), uma relação lógica *fuzzy* é estabelecida, caso $F(t-1) = A_1$ e $F(t) = A_j$. Portanto, a relação lógica *fuzzy* é caracterizada mediante a conexão entre dois termos consecutivos, no caso, $F(t)$ e $F(t-1)$, que pode ser denotado por $A_i \rightarrow A_j$, onde A_i é denominado de lado esquerdo e A_j se denomina de lado direito da relação lógica *fuzzy*, tal como, se pode observar em Lago Neto (2011), o qual garante que estas associações lógicas *fuzzy*, permitem ser agrupadas para estabelecer diferentes associações *fuzzy*.

Um modelo de série temporal *fuzzy* sazonal de primeira ordem, tal como, apresentado

por Song (1999) descreve que, para $F(t)$ sendo uma série temporal *fuzzy*, ao admitir-se existência de sazonalidade nesta série, ou seja, em $F(t), t = 1; 2; 3; \dots; N$, o modelo autoregressivo *fuzzy* sazonal de primeira ordem resulta em $F(t - m) \rightarrow F(t)$, tal que, “ m ” indica o período sazonal observado.

Um modelo de série temporal *fuzzy* sazonal de ordem “ p ”, assim como, definido por Chen (2002) indica que, para $F(t)$ sendo uma série temporal *fuzzy*, em $F(t), t = 1; 2; 3; \dots; N$, caso $F(t)$ seja resultante de $F(t - 1); F(t - 2); F(t - 3); \dots$, como também de $F(t - p)$, logo esta associação lógica *fuzzy* é determinada segundo a Equação (4.16),

$$F(t - 1); F(t - 2); F(t - 3); \dots; F(t - p) \rightarrow F(t), \quad (4.14)$$

tal que, esta associação é considerada um modelo de série temporal *fuzzy* sazonal de ordem “ p ”.

4.3 - Algoritmos de Previsão Para Série Temporal Fuzzy

A disponibilidade de algoritmos na literatura, tal como, os citados a seguir, é de suma importância quanto à previsão em série temporal *fuzzy* via equações relacionais, disponibilizando desta forma uma opção de escolha, pelo algoritmo mais conveniente ao interesse do pesquisador, em termos de funcionalidade e retorno preditivo significativo.

Deve haver na literatura algoritmos que exijam menor esforço computacional para sua implementação, sem que os conceitos e definições alcançados pela metodologia de sistemas *fuzzy*, combinadas com as técnicas de série temporal sejam comprometidos em sua eficiência.

Neste contexto, destaca-se o algoritmo proposto por Chen (2002), o qual supera os demais algoritmos apresentados a seguir, visto que, desempenha de forma eficaz uma rotina de previsão em série temporal *fuzzy*, além de possuir uma simplicidade computacional relevante, para o processo de previsão em série temporal *fuzzy*, sobretudo, quando comparado com os algoritmos sugeridos por Song e Chisson (1993a), como também de Song e Chisson (1993b), os quais atuaram de forma pioneira neste tipo de previsão *fuzzy* em série temporal.

A justificativa para este fato é devido Chen (2002), ter utilizado operações aritméticas, ao invés da composição “Max-Min”, para realizar assim previsões *fuzzy* de uma deter-

minada característica observável via série temporal, que resultou em uma rotina de programação mais simples, em termos de implementação computacional dos procedimentos necessários à esta previsão *fuzzy*.

4.3.1 - Algoritmo de Song e Chissom (1993a) e (1993b)

Contornando a dificuldade inicial de se trabalhar com uma série de dados descrita necessariamente por termos linguísticos, Song e Chissom (1993a), como também Song e Chissom (1993b) estabeleceram o conceito de lógica *fuzzy*, aliada ao problema de previsão em série temporal, suscitando propor uma nova metodologia de previsão, definida sobre série temporal *fuzzy*.

Dentre os principais algoritmos de previsão em modelos de série temporal *fuzzy* via equações relacionais, se torna possível estabelecer segundo Song e Chissom (1993a), algumas etapas para descrever a construção de um modelo de série temporal *fuzzy*, sob a forma:

$$F(t) = F(t - 1) \circ R(t, t - 1),$$

que pode ser considerado como um procedimento inicial. As etapas são descritas como:

1. Definir o universo de discurso U no qual os dados históricos estão compreendidos, tal como, descrito na Equação (4.15) e ainda, em quais conjuntos *fuzzy* estes dados serão definidos,

$$U = [D_{(\min)} - D_{(1)}; D_{(\max)} + D_{(2)}], \quad (4.15)$$

tal que, $D_{(\min)}$ é o menor valor observado na série de dados; D_1 e D_2 são números pequenos selecionados de forma subjetiva, ou seja, dependem do especialista que está realizando o procedimento de *fuzzificação*; e $D_{(\max)}$ é o maior valor identificado na série de dados.

2. Particionar o universo U em vários intervalos de mesmo comprimento;
3. Definir os conjuntos *fuzzy* sobre o universo U ;
4. *Fuzzificar* os dados históricos;
5. Determinar as relações *fuzzy*;
6. Obter os resultados das previsões;
7. Interpretar os resultados obtidos;

8. Defuzzificar os valores preditos.

A partir da metodologia desenvolvida inicialmente por Song e Chisson (1993a), como em Song e Chisson (1993b), os quais propuseram o uso de equações relacionais *fuzzy*, para possibilitar a previsão em série temporal *fuzzy*, foi possível observar uma quantidade significativa de propostas, que passaram a ser sugeridas na literatura por diversos autores, no sentido de contornar também o problema de previsão em série temporal *fuzzy*, porém, estas propostas que surgiram posteriormente à Song e Chisson (1993a), assim como em Song e Chisson (1993b), caracterizam um melhoramento ou aperfeiçoamento do algoritmo originalmente proposto para o problema de previsão em série temporal *fuzzy*. Neste sentido, se pode destacar na literatura os algoritmos de previsão para um modelo de série temporal *fuzzy* via equações relacionais listados abaixo.

4.3.2 - Algoritmo de Chen (1996)

Segundo Lago Neto (2011), Shyi-Ming Chen no ano de 1996, defendeu a utilização de um algoritmo caracterizado pela substituição das operações *max-min*, por operações aritméticas, com o intuito de reduzir a complexidade computacional agregada ao procedimento originalmente proposto por Song e Chisson (1993a), como também em Song e Chisson (1993b). Seguindo uma tendência de aprimoramento do algoritmo de Song e Chisson proposto em 1993, o qual é empregado até os dias atuais por diversos pesquisadores, Chen (1996) estabelece um método que modifica o de Song e Chisson, compreendendo três passos essenciais descritos a seguir.

Passo 1: Definir o universo de discurso e seu respectivo número de intervalos de mesma dimensão (amplitude);

O universo de discurso é obtido de forma idêntica ao proposto por Song e Chisson (1993a) e representado pela Equação (4.15). A amplitude do intervalo é dada pela expressão $at = \{\text{limite superior do intervalo } i - \text{limite inferior do intervalo } i\}$, onde $i = 1; 2; \dots; n$.

Passo 2: Definir os conjuntos *fuzzy* a partir do universo de discurso; *fuzzificação* da série de dados históricos e identificação das relações *fuzzy* existentes na série temporal *fuzzificada*.

A definição dos conjuntos *fuzzy* e a *fuzzificação* da série seguem os mesmos procedimentos estabelecidos por Song e Chisson (1993a), assim como, em Song e Chisson (1993b). No entanto, a identificação das relações *fuzzy* existentes na série temporal *fuzzificada*, sofre uma mudança em relação a Song e Chisson (1993a), como em Song e Chisson (1993b), pois estes utilizam em seu algoritmo os operadores “*Max - Mim*”, diferentemente de Chen (1996), que adota as relações lógicas *fuzzy* $A_i \rightarrow A_j$ sendo classificadas em grupos de relações *fuzzy*, tal como, se pode verificar em Lago Neto (2011), pg. 48.

Passo 3: Obtenção das Previsões *fuzzy*

As Previsões *fuzzy* diferentemente ao proposto por Song e Chisson (1993a), assim como, em Song e Chisson (1993b), que utilizaram uma matriz de relações *fuzzy* $R(t; t - 1)$, neste passo, Chen (1996) determina as previsões da série temporal *fuzzy* no instante t , por três heurísticas, isto é,

Primeira Heurística: $A_i \rightarrow A_j$, onde A_i e A_j são conjuntos *fuzzy*;

Segunda Heurística: $A_i \rightarrow A_{j_1}; A_i \rightarrow A_{j_2}; \dots; A_i \rightarrow A_{j_r}$, tal que, $A_i, A_{j_1}, A_{j_2}, \dots, A_{j_r}$ são conjuntos *fuzzy*;

Terceira Heurística: $A_i \rightarrow \emptyset$, sendo A_i um conjunto *fuzzy*,

onde, estas três heurística são observadas com detalhes em Lago Neto (2011), pg. 48 e pg. 49.

4.3.3 - Algoritmo de Chen (2002)

No ano de 2002, Shyi-Ming Chen apresentou uma nova proposta de algoritmo para previsão em série temporal *fuzzy*, via equações relacionais. O primeiro passo do seu algoritmo é semelhante ao proposto por Song e Chisson (1993a), assim como, ao de Song e Chisson (1993b). A partir do segundo passo, se pode observar mudanças pontuais nos algoritmos propostos anteriormente, passando pela definição dos conjuntos *fuzzy* no universo de discurso “ U ”, assim como, pela *fuzzificação* da série temporal e identificação dos grupos de relações *fuzzy* existentes na série temporal *fuzzificada*. O terceiro passo deste algoritmo descreve a obtenção das previsões *fuzzy*, as quais são produzidas via regras heurísticas, de forma similar ao algoritmo de Chen (1996), como garante Lago Neto (2011) que descreve o passo a passo para a determinação do algoritmo de Chen (2002), o qual conta ainda com o quarto passo (*defuzzificação* dos valores previstos).

4.3.4 - Algoritmo de Liu (2007)

No ano de 2007, Hao-Tien Liu propôs um algoritmo para previsão em série temporal mediante equações relacionais, com algumas alterações em sua composição em relação a proposta apresentada por Shyi-Ming Chen em 1996, dentre estas alterações se pode destacar:

Diferença 1: A quantidade de elementos necessários para a partição do universo de discurso “ U ” estabelecida por Huarng (2001), foi introduzida no processo de *fuzzificação* dos dados no algoritmo de Liu (2007), para desta forma superar a condição de subjetividade implementada por Song e Chisson (1993a), além de Song e Chisson (1993b), assim como, por Chen (1996), quanto à determinação da quantidade e tamanho dos intervalos necessários.

Diferença 2: Liu (2007), sugere a utilização de número *fuzzy* trapezoidal, tal como descrito na Definição 3.2.6, em substituição aos conjuntos *fuzzy* A_i , com o intuito de possibilitar a composição dos valores lingüísticos da série temporal *fuzzy*;

Diferença 3: Em contrapartida à Chen (1996), o qual calcula as previsões *fuzzy* mediante a adoção de três regras heurísticas, as quais utilizam os pontos médios dos intervalos, Liu (2007) gera previsões *fuzzy* a partir de operações aritméticas de soma e produto entre números *fuzzy*, e ainda, estas previsões são obtidas mediante a introdução do conceito de α -nível para valores intervalares, o que permite à rotina de previsão *fuzzy*, agregar o conhecimento de um especialista ao algoritmo de previsão em série temporal *fuzzy*.

A rotina para obtenção do algoritmo de previsão em série temporal *fuzzy*, via equações relacionais tal como proposto por Liu (2007), esta descrita em oito passos, que compreendem desde, a determinação do universo de discurso “ U ” no passo 1, até o calculo dos valores previstos no passo 8, como se pode observar em Lago Neto (2011), pg. 53 a 56.

O estudo de série temporal *fuzzy* desperta o interesse de vários pesquisadores na comunidade científica há algum tempo, onde se pode observar um numero grandioso de propostas sobre este tema, desde o algoritmo proposto por Song e Chisson (1993a). Dentre as propostas iniciais de algoritmo de previsão para série temporal *fuzzy*, se pode perceber uma exclusividade à modelos autoregressivos de primeira ordem, algo que restringiria a

aplicação dos procedimentos de previsão *fuzzy*, a partir dos algoritmos propostos originalmente por Song e Chisson (1993a). Porém, esta restrição foi superada com a proposta de Chen (2002), pois este propôs um modelo de previsão para série temporal *fuzzy* de ordem superior a dois, visto que, não necessariamente os dados coletados na natureza remeterão à modelos de série temporal de baixa ordem, ou seja, $p = 1$ ou 2 , onde “ p ” representa a ordem do parâmetro auto-regressivo no modelo de série temporal. Neste sentido, surge a necessidade de identificar de forma eficiente e precisa, a ordem da série de dados históricos que represente alguma característica observável em um cenário de interesse.

4.4 - Alguns Métodos Conhecidos e Disponíveis na Literatura Para Identificação dos Modelos de Série Temporal Fuzzy

Segundo Lago Neto (2011), para que a metodologia clássica de série temporal possa ser aplicada de forma eficiente, há de se atender a suposição de linearidade dos dados correspondentes à série histórica, além de um número mínimo de observações da série igual a 50, tal como, defendem também Tseng et al. (2001) e Tseng et al. (2002). No entanto, ao se utilizar a metodologia *fuzzy* combinada com os preceitos de série temporal, as suposições de linearidade e de um número mínimo de observações para a série são desprezadas, passando a serem irrelevantes para o processo de modelagem *fuzzy*, as duas suposições anteriormente citadas, necessárias então no caso da modelagem clássica exclusivamente. Surge, porém, a necessidade inerente ao processo de modelagem, no diz respeito à identificação correta da ordem do modelo de previsão em série temporal *fuzzy*, pois, na literatura não existe ainda um procedimento padrão a ser seguido, para identificar de forma exata a ordem do modelo de série temporal. Existem na literatura várias propostas de identificação da ordem de processos *fuzzy* autoregressivos, dos quais se destacam, Chen (2002), Song (2003), Yu e Huarng (2008), Egrioglu et al. (2009), Lago Neto (2011) e Carvalho Jr. (2012).

O procedimento de identificação escolhido para apresentar a sua rotina de identificação, em um modelo de série temporal *fuzzy*, dentre os disponíveis na literatura e citados anteriormente é o desenvolvido por Song (2003). Este autor utilizou como ferramentas fundamentais no processo de identificação da série temporal *fuzzy* a função de autocorrelação amostral, a função de autocorrelação parcial, além do valor médio da função de auto-

correlação amostral, esta última atuando como uma medida de dependência entre os conjuntos *fuzzy*.

Porém, a partir dos resultados obtidos pelos procedimentos empregados por Song (2003), se tornou possível perceber que a função de autocorrelação amostral e função de autocorrelação parcial, não indicam de forma exata a ordem de um processo autoregressivo, estas funções apenas sugerem uma possível ordem para o modelo de série temporal *fuzzy*, sendo que, se pode imaginar “ n ” modelos, mediante um ou mais *lag's* significativos em um intervalo de valores entre $\frac{2}{\sqrt{T}}$, com 95% de probabilidade, onde T é o tamanho da série temporal. Neste sentido, se faz necessário estabelecer uma medida de dependência que atue de forma relevante, ou seja, uma relação de continuidade cronológica, entre os conjuntos *fuzzy* originados a partir de uma série histórica para os dados observados.

4.4.1 - Algumas Medidas de Dependência Conhecidas Entre Conjuntos Fuzzy

Usualmente na metodologia de série temporal os valores históricos na abordagem clássica, se dão mediante uma variável quantitativa, logo, os valores característicos de um processo estocástico dado por $X(t)$ suscitam representações numéricas, diferentemente do que ocorre com a metodologia *fuzzy*, onde os valores da série temporal são representados por uma variável qualitativa, isto é, por termos linguísticos, para os quais se tem uma função f_i , onde $i = 1; 2; 3; \dots; m$, de maneira que, para cada instante de tempo t , se tem valores linguísticos diferentes (conjuntos).

Dada uma série temporal *fuzzy* $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$, tal que, m representa a quantidade dos possíveis valores linguísticos para um processo dinâmico $X(t)$ e T representa o tamanho da série temporal, assim como se observa em Song (2003).

Medida de Dependência 1.

Chama-se de função de autocorrelação amostral *defuzzificada*, segundo Song (2003), para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Portanto, a função de autocorrelação amostral para os valores *defuzzificados* é expressa segundo a Equação (4.16), como pode ser encontrado também em Lago Neto (2011).

$$\hat{\rho}^{DFz}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{T-k} \{DF[f_i(t)] - \bar{x}\} \{DF[f_i(t+k)] - \bar{x}\}}{\sum_{t=1}^T \{DF[f_i(t)] - \bar{x}\}^2}, \quad (4.16)$$

onde, $k = 0; 1; 2; \dots$ e “ DF ” indica o operador de *defuzzificação* mediante o centro de gravidade, e ainda, \bar{x} é a média amostral de $DF[f_i(t_1)]; \dots ; DF[f_i(t_T)]$, que é expressa pela Equação (4.17).

$$\bar{x} = \frac{\sum_{t=1}^T DF[f_i(t)]}{T}. \quad (4.17)$$

Medida de Dependência 2.

Denota-se por função de autocorrelação amostral com α -nível (CA), segundo Song (2003), caso exista um $\alpha \in [0; 1]$ para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Neste sentido, a função de autocorrelação amostral com α -nível é expressa segundo a Equação (4.18), tal como, se observa em Lago Neto (2011).

$$\hat{\rho}^{CA}(k) = \frac{\sum_{b=1}^B \rho^{CA}(k; \alpha; b)}{B}, \quad (4.18)$$

onde, $k = 0; 1; 2; \dots$ e $\hat{\rho}^{CA}(k) = \bar{\rho}_\alpha(k)$, e ainda, o termo B indica o número de amostras independentes, tal que, $\rho^{CA}(k; \alpha; b)$ possui valores que são obtidos de forma análoga à Equação (4.16), apesar de que, na Equação (4.18) as expressões $DF[f_i(t)]$ e $DF[f_i(t+k)]$, apresentam condições de serem selecionadas aleatoriamente mediante o suporte dos conjuntos *fuzzy* A_{i_t} e $A_{i_{t+k}}$, com seus respectivos valores de pertinência α .

Medida de Dependência 3.

Denomina-se por função de autocorrelação amostral média aleatória, segundo Song (2003), para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Então, a função de autocorrelação amostral média aleatória para os valores *defuzzificados* são expressos segundo a Equação (4.19), como observado também em Lago

Neto (2011).

$$\hat{\rho}^{RA}(k) = \frac{\sum_{b=1}^B \rho^{RA}(k; b)}{B}, \quad (4.19)$$

onde, $k = 0; 1; 2; \dots$ e $\hat{\rho}^{RA}(k) = \bar{\rho}_r(k)$, e ainda, o termo B indica o número de amostras independentes, tal que, $\rho^{RA}(k; b)$ apresenta valores possíveis de serem obtidos analogamente à Equação (4.16), apesar de que, na Equação (4.19) as expressões $DF[f_i(t)]$ e $DF[f_i(t + k)]$, apresentam condições de serem selecionadas aleatoriamente mediante o suporte dos conjuntos *fuzzy* A_{i_t} e $A_{i_{t+k}}$, respectivamente.

Segundo Carvalho Jr. (2012), a função de autocorrelação amostral possui a capacidade de identificar a ordem do componente de médias móveis “ q ” em uma série temporal. Porém, caso a série temporal caracterize um processo autoregressivo de ordem “ p ”, de forma total ou parcial, ou seja, componente autoregressivo e de médias móveis em uma mesma série, a função de autocorrelação amostral se torna incapaz de determinar a ordem do modelo em relação ao parâmetro “ p ”. Para contornar esta restrição no processo de identificação, associado à função de autocorrelação amostral em um processo autoregressivo, deve-se utilizar formas alternativas para identificar a ordem de um modelo de série temporal, como a função de autocorrelação parcial, a função de autocorrelação inversa, ou ainda, de forma mais precisa, a função de autocorrelação estendida, pois esta segundo a Tabela (4.1) indica a ordem exata dos parâmetros p e q .

Medida de Dependência 4.

Chama-se de função de autocorrelação parcial amostral *defuzzificada*, segundo Lago Neto (2011), para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Portanto, a função de autocorrelação parcial amostral para os valores *defuzzificados* é expressa por ϕ_{kk}^{DFz} , que se obtém substituindo $\rho_X(k)$ por $\rho^{DFz(k)}$ na função de autocorrelação parcial amostral clássica, tal como, dado nas Equações (4.5) a (4.7), pelo algoritmo desenvolvido por *Durbin-Levinson*, de onde se chega as Equações (4.20) a (4.22).

$$\phi_{1;1}^{DFz} = \rho_X^{DFz}(1), \quad (4.20)$$

$$\phi_{k;k}^{DFZ} = \frac{\rho_X^{DFZ}(k) - \sum_{j=1}^{k-1} \phi_{k-1;j}^{DFZ} \rho_X^{DFZ}(k-j)}{1 - \sum_{j=1}^{k-1} \phi_{k-1;j}^{DFZ} \rho_X^{DFZ}(j)} \quad (4.21)$$

e

$$\phi_{k;j}^{DFZ} = \phi_{k-1;j}^{DFZ} - \phi_{k;k}^{DFZ} \phi_{k-1;k-j}^{DFZ}, \quad (4.22)$$

tal que, $j = 1; 2; \dots; k - 1$.

As rotinas dadas pelas Equações (4.21) e (4.22) são refeitas para defasagens em $k = 2; 3; 4; \dots$ e desta forma, se pode obter $\phi_{k;k}^{DFZ}$, mediante a Equação (4.21), tal que, $\rho_X^{DFZ}(k)$ indica os valores da função de autocorrelação amostral, a partir da utilização dos dados *defuzzificados* obtidos a partir da Equação (4.16).

Medida de Dependência 5.

Denota-se por função de autocorrelação parcial com α -nível (CA), caso exista um $\alpha \in [0; 1]$ para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Neste sentido, a função de autocorrelação parcial com α -nível é expressa segundo a Equação (4.23), tal como, se observa em Lago Neto (2011).

$$\phi_{kk}^{CA}(\alpha; b) = \frac{\rho^{CA}(k; \alpha; b) - \sum_{j=1}^{k-1} \phi_{k-1;j}^{CA}(\alpha; b) \rho^{CA}(k-j; \alpha; b)}{1 - \sum_{j=1}^{k-1} \phi_{k-1;j}^{CA}(\alpha; b) \rho^{CA}(j; \alpha; b)}, \quad (4.23)$$

onde, caso $k \geq 2$, se tem origem a Equação (4.24).

$$\phi_{k;j}^{CA}(\alpha; b) = \phi_{k-1;j}^{CA}(\alpha; b) - \phi_{k;k}^{CA}(\alpha; b) \phi_{k-1;k-j}^{CA}(\alpha; b), \quad (4.24)$$

tal que, $j = 1; 2; \dots; k - 1$.

É importante destacar que, o valor da função de autocorrelação parcial com α -nível, o qual é obtido mediante a Equação (4.23), provém de forma recursiva do algoritmo de *Durbin-Levinson*, tal que, o ponto inicial do processo de obtenção é dado pela Equação (4.25).

$$\phi_{1;1}^{CA}(\alpha; b) = \rho^{CA}(1; \alpha; b) \quad (4.25)$$

As rotinas dadas pelas Equações (4.23) e (4.24) são refeitas para defasagens em $k =$

$2; 3; 4; \dots$ e $b = 1; 2; 3; \dots; B$, tal que, B representa o número de amostras selecionadas, e ainda, se pode obter $\rho_{k;\alpha;b}^{CA}$, mediante a Equação (4.16), apesar de que, $DF[f_i(t)]$ e $DF[f_i(t+k)]$, apresentam condições de serem selecionadas aleatoriamente mediante o suporte dos conjuntos fuzzy A_{i_t} e $A_{i_{t+k}}$, com seus respectivos valores de pertinência α .

Medida de Dependência 6.

Denota-se por função de autocorrelação parcial média com α -nível (CA), caso exista um $\alpha \in [0; 1]$ para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores defuzzificados oriundos de uma série temporal fuzzy, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Por conseguinte, a função de autocorrelação parcial média com α -nível é representada pela Equação (4.26), tal como, se pode observar em Lago Neto (2011).

$$\bar{\phi}^{CA}(k) = \frac{\sum_{b=1}^B \phi_{k;k}^{CA}(\alpha; b)}{B} \tag{4.26}$$

onde, $k = 1; 2; \dots$ e $\phi_{k;k}^{CA}(\alpha; b)$ é obtido a partir da Equação (4.23).

Medida de Dependência 7.

Chama-se de função de autocorrelação parcial média aleatória, segundo Lago Neto (2011), para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores defuzzificados oriundos de uma série temporal fuzzy, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Neste sentido, a função de autocorrelação parcial média aleatória para os valores defuzzificados são expressos segundo a Equação (4.27).

$$\bar{\phi}^{RA}(k) = \frac{\sum_{b=1}^B \phi_{k;k}^{RA}(b)}{B}, \tag{4.27}$$

onde, $k = 1; 2; \dots$. Porém, para que seja possível obter o valor de $\phi_{k;k}^{RA}(b)$, se deve utilizar o algoritmo de *Durbin-Levinson*, de maneira que, as funções de autocorrelações parciais médias aleatórias são obtidas, mediante a substituição em $\rho_X(k)$ por $\rho^{RA}(k; b)$, nas Equações (4.5) a (4.7), o que resulta na Equação (4.28).

$$\phi_{1;1}^{RA}(b) = \rho^{CA}(1; b). \tag{4.28}$$

Caso $k \geq 1$, se pode obter as equações de atualização a partir da Equação (4.29).

$$\phi_{k;k}^{RA}(b) = \frac{\rho^{RA}(k;b) - \sum_{j=1}^{k-1} \phi_{k-1;j}^{RA}(b) \rho^{CA}(k-j;b)}{1 - \sum_{j=1}^{k-1} \phi_{k-1;j}^{RA}(b) \rho^{CA}(j;b)}, \quad (4.29)$$

onde, dado $k \geq 2$, a Equação (4.30) é obtida.

$$\phi_{k;j}^{RA}(b) = \phi_{k-1;j}^{RA}(b) - \phi_{k;k}^{RA}(b) \phi_{k-1;k-j}^{RA}(b), \quad (4.30)$$

onde, $j = 1; 2; \dots; k - 1$. As rotinas dadas pelas Equações (4.29) e (4.30) são refeitas para defasagens em $k = 2; 3; 4; \dots$ e $b = 1; 2; 3; \dots; B$, tal que, B representa o número de amostras selecionadas, e ainda, se pode obter $\phi_{k;k}^{RA}(b)$, mediante a Equação (4.16), apesar de que, $DF[f_i(t)]$ e $DF[f_i(t+k)]$, apresentam condições de serem selecionados aleatoriamente mediante o suporte dos conjuntos fuzzy A_{i_t} e $A_{i_{t+k}}$, respectivamente.

Medida de Dependência 8.

Denomina-se de função de autocorrelação inversa fuzzy, segundo Lago Neto (2011), para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores defuzzificados oriundos de uma série temporal fuzzy, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Neste sentido, a função de autocorrelação inversa fuzzy para os valores defuzzificados é expressa pela Equação (4.31).

$$\rho_{inv}^{DFZ}(h) = \frac{-\phi_{r;h}^{DFZ} + \sum_{j=1}^{r-h} \phi_{r;j}^{DFZ} \phi_{r;j+h}^{DFZ}}{1 + \sum_{j=1}^r (\phi_{r;j}^{DFZ})^2}, \quad (4.31)$$

onde, $\phi_{r;h}^{DFZ}$ possui valores a serem obtidos recursivamente pelas Equações (4.5) a (4.7).

Medida de Dependência 9.

Denota-se por função de autocorrelação inversa com α -nível (CA), segundo Lago Neto (2011), caso exista um $\alpha \in [0; 1]$ para $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores defuzzificados oriundos de uma série temporal fuzzy, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Neste sentido, a função de autocorrelação inversa com α -nível é representada pela Equação (4.32).

$$\rho_{inv}^{CA}(h; \alpha; b) = \frac{-\phi_{r;h}^{CA}(\alpha; b) + \sum_{j=1}^{r-h} \phi_{r;j}^{CA}(\alpha; b) \phi_{r;j+h}^{CA}(\alpha; b)}{1 + \sum_{j=1}^r [\phi_{r;j}^{CA}(\alpha; b)]^2}, \quad (4.32)$$

onde, $h = 1; 2; \dots; r$ e $\phi_{r;k}^{CA}(\alpha; b)$ são obtidos de forma recursiva, de acordo com a rotina descrita nas Equações (4.23) e (4.24). Realiza-se a repetição de $\phi_{r;k}^{CA}(\alpha; b)$ para $k = 2; 3; \dots; r - 1$ e $b = 1; 2; 3; \dots; B$.

Medida de Dependência 10.

Define-se por função de autocorrelação inversa média com α -nível (CA), caso exista um $\alpha \in [0; 1]$ para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Portanto, a função de autocorrelação inversa média com α -nível é representada pela Equação (4.33), como se observa em Lago Neto (2011).

$$\bar{\rho}_{inv}^{CA}(h) = \frac{\sum_{b=1}^B \rho_{inv}^{CA}(h; \alpha; b)}{B}, \quad (4.33)$$

onde, $h = 1; 2; \dots; r$ e $\rho_{inv}^{CA}(h; \alpha; b)$ é obtido a partir da Equação (4.32).

Quanto ao calculo da função de autocorrelação inversa com α -nível, segundo Lago Neto (2011), os valores *defuzzificados* originam-se do mesmo contorno α . Então, se pode garantir que a média das $\rho_{inv}^{CA}(h; \alpha; b)$ passa a ser denominada de função de autocorrelação inversa média com α -nível.

Medida de Dependência 11.

Denota-se por função de autocorrelação inversa média aleatória, segundo Lago Neto (2011), para $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$. Neste sentido, a função de autocorrelação inversa média aleatória para os valores *defuzzificados* é definida a partir da Equação (4.34).

$$\bar{\rho}_{inv}^{RA}(h) = \frac{\sum_{b=1}^B \rho_{inv}^{RA}(h; b)}{B}, \quad (4.34)$$

onde, $h = 1; 2; 3; \dots; r$ e os valores referentes à $\rho_{inv}^{RA}(h; b)$, se tornam possíveis de serem obtidos mediante a utilização da Equação (4.35).

$$\rho_{inv}^{RA}(h; b) = \frac{-\phi_{r;h}^{RA}(b) + \sum_{j=1}^{r-h} \phi_{r;j}^{RA}(b) \phi_{r;j+h}^{RA}(b)}{1 + \sum_{j=1}^r [\phi_{r;j}^{RA}(b)]^2}, \quad (4.35)$$

onde, $h = 1; 2; \dots; r$ e ainda, $\phi_{r;k}^{RA}(b)$ podem ser obtidos de forma recursiva a partir das Equações (4.29) e (4.30), tal que, se repete o calculo de $\phi_{r;k}^{RA}(b)$ para $k = 2; 3; \dots; r - 1$ e $b = 1; 2; 3; \dots; B$. Portanto, se pode denominar de função de autocorrelação inversa média aleatória para $\bar{\rho}_{inv}^{RA}(h)$.

Lago Neto (2011) defende a idéia de selecionar a ordem de um modelo autoregressivo, mediante o suporte de cada conjunto *fuzzy* com uma quantidade distinta de valores, para que posteriormente seja possível calcular a função de autocorrelação parcial e autocorrelação inversa, em relação à cada conjunto de dados. A partir do comportamento destas funções com valores *defuzzificados* e médios, se tem estabelecida a identificação da ordem do modelo de série temporal *fuzzy*. Ainda, segundo o autor supracitado, Song (2003) garante que as Equações (4.36) e (4.37) são verdadeiras.

$$|\rho^{DF_z}(k)| \geq |\bar{\rho}^{CA}(k)| \quad (4.36)$$

e

$$|\rho^{DF_z}(k)| \geq |\bar{\rho}^{RA}(k)|. \quad (4.37)$$

Em decorrência dos resultados acima, se torna razoável esperar que a estimativa $|\hat{\phi}_{kk}^{DF_z}|$ exceda a estimativa $|\bar{\phi}_{kk}^{CA}|$. De forma imediata se pode sugerir que $\hat{\phi}_{kk}^{DF_z} \neq 0$ e $\bar{\phi}_{kk}^{CA} = 0$, ou seja, é possível qua a função de autocorrelação parcial com α -nível, indique imprecisamente um modelo autoregressivo *fuzzy* com uma ordem inferior a um *lag* “ k ”, e portanto, o processo de identificação ficaria com sua eficiência comprometida nestas condições. Neste sentido, se faz necessário estabelecer procedimentos capazes de identificar a ordem de processos autoregressivos *fuzzy*, com a maior precisão possível, sob pena, de julgar satisfatório um modelo incapaz de representar o comportamento real de uma série temporal *fuzzy*, caso o processo de identificação em alguma etapa seja equivocado, como por exemplo, pela escolha da função de autocorrelação adequada.

4.5 - Desenvolvimento e Proposição de Novas Ferramentas Para Identificação dos Modelos de Série Temporal Fuzzy

Nesta seção propõem-se novas ferramentas para auxílio na rotina de identificação, em processos autoregressivos sazonais com uma abordagem *fuzzy*. Portanto, se definem três funções de autocorrelação estendida combinadas com a metodologia de conjuntos *fuzzy*, para com isto tentar estabelecer uma forma mais precisa e eficaz, quanto ao processo de identificação da ordem de uma série temporal *fuzzy*. Porém, de uma forma diferente do que se encontra disponível na literatura atual, quando se trata de uma rotina para identificação e previsão em série temporal com abordagem *fuzzy*. Neste sentido, três ferramentas de identificação são propostas a seguir, com base nos conceitos e definições da função de autocorrelação estendida (Ver, Cryer e Chan, 2008), além das Seções 4.2 e 4.3, que tratam metodologia sobre série temporal *fuzzy* e algoritmos de previsão para série temporal *fuzzy*, respectivamente, como se pode observar posteriormente.

Ferramenta de Identificação Proposta 1.

Denomina-se função de autocorrelação estendida *fuzzy*, para todo $DF[f_i(t)]$ com $1 \leq i \leq m$ e $1 \leq t \leq T$, como sendo o conjunto de valores *defuzzificados* oriundos de uma série temporal *fuzzy*, dada por $f_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$, onde m indica o número de conjuntos *fuzzy* utilizados, onde T é o tamanho de trajetória da série temporal. Desta forma, a função de autocorrelação estendida *fuzzy* para os valores *defuzzificados* é expressa pela Equação (4.38).

$$W_{i,k,j}^{DF_Z} = Z_i^{DF_Z} - \tilde{\phi}_1^{DF_Z} Z_{i-1}^{DF_Z} - \dots - \tilde{\phi}_k^{DF_Z} Z_{i-k}^{DF_Z}, \quad (4.38)$$

onde, $k = 0; 1; 2; \dots$, $j = 0; 1; 2; \dots$ e “ DF_Z ” indica o operador de *defuzzificação* mediante o método centroide (centro de gravidade), e ainda, $\tilde{\phi}_k^{DF_Z}$, representa um estimador consistente *fuzzy* de $\phi_k^{DF_Z}$. Deve-se ressaltar que, os parâmetros $\phi_k^{DF_Z}$ são obtidos de forma iterativa, como pode ser visto em Cryer e Chan (2008), isto é, somente após a série temporal *defuzzificada* ser obtida, os estimadores *fuzzy* são determinados.

Ferramenta de Identificação Proposta 2.

Denomina-se função de autocorrelação estendida *fuzzy* com α -nível (EACF *fuzzy*-CA), caso exista para todo $\alpha \in [0; 1]$ uma função $W_{i,k,j}^{CA}(h; \alpha; b)$, onde o valor de α é definido

a priori por um especialista. Isso implicará na ativação do suporte do conjunto *fuzzy* correspondente, e contido no universo de discurso U , de onde são selecionados os valores em cada intervalo *fuzzy*, que irão compor a série temporal *fuzzy* com α -nível. Seja $Z_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$, onde m indica o número de conjuntos *fuzzy* utilizados e T é o tamanho de trajetória da série temporal. Então, a função EACF *fuzzy*-CA é representada pela Equação (4.39).

$$W_{i,k,j}^{CA}(h; \alpha; b) = Z_i^{CA}(\alpha; b) - \tilde{\phi}_1^{CA}(\alpha; b) Z_{i-1}^{CA}(\alpha; b) - \dots - \tilde{\phi}_k^{CA}(\alpha; b) Z_{i-k}^{CA}(\alpha; b), \quad (4.39)$$

onde, $h = 1; 2; \dots; r$, $b = 1; 2; \dots; r$, e $\tilde{\phi}_k^{CA}(\alpha; b)$ são obtidos de forma iterativa (Cryer e Chan (2008)), no caso, após a série temporal *crisp* ser submetida ao α -nível, com valor igual, por exemplo, a $\alpha = 0,25$; $\alpha = 0,50$ ou $\alpha = 0,75$. Em seguida ocorre a seleção aleatória, a partir do suporte dos conjuntos *fuzzy* ativados pela função de pertinência desejada, e finalmente os estimadores *fuzzy* são obtidos com a aplicação da função EACF *fuzzy*-CA.

Denomina-se função de autocorrelação estendida *fuzzy* média com α -nível (EACF *fuzzy* média-CA), para todo $\alpha \in [0; 1]$ sobre a função $\overline{W}_{i,k,j}^{CA}(h)$, como sendo o valor médio resultante de uma série temporal *fuzzy* com α -nível. Então, a função EACF *fuzzy* média-CA é representada de acordo com a Equação (4.40).

$$\overline{W}_{i,k,j}^{CA}(h) = \frac{\sum_{b=1}^B W_{i,k,j}^{CA}(h; \alpha; b)}{B}, \quad (4.40)$$

onde, $W_{i,k,j}^{CA}$ segue a Equação (4.39), $h = 1; 2; \dots; r$, $b = 1; 2; \dots; B$ e B o número de amostras.

Com o cálculo da função de autocorrelação estendida *fuzzy* com α -nível, se torna razoável aceitar que os valores resultantes originam-se do mesmo contorno α . Assim, se pode garantir que a média das funções $W_{i,k,j}^{CA}(h; \alpha; b)$, passa a ser denominada de função de autocorrelação estendida *fuzzy* média com α -nível, e conseqüentemente representada pela expressão $\overline{W}_{i,k,j}^{CA}(h)$.

Ferramenta de Identificação Proposta 3.

Denomina-se função de autocorrelação estendida *fuzzy* média aleatória (EACF *fuzzy* média aleatória-RA), para a média da EACF dentro de cada intervalo *fuzzy*, após a ativação

do suporte dos conjuntos *fuzzy* contidos no universo de discurso U . Assim, são gerados aleatoriamente T valores dentro de cada intervalo *fuzzy*, para posteriormente obter as EACF's *fuzzy* média alatória, as quais irão compor a série temporal *fuzzy*, para todo $Z_i(t)$, com $1 \leq i \leq m$ e $1 \leq t \leq T$, onde m indica o número de conjuntos *fuzzy* utilizados. Então, a função EACF *fuzzy* média alatória (EACF-RA) é definida pela Equação (4.41).

$$W_{i,k,j}^{RA}(h; b) = Z_i^{RA}(b) - \tilde{\phi}_1^{RA}(b) Z_{i-1}^{RA}(b) - \dots - \tilde{\phi}_k^{RA}(b) Z_{i-k}^{RA}(b), \quad (4.41)$$

onde, $h = 1; 2; \dots; r$, $k(\text{AR}) \geq 0$, $j(\text{MA}) \geq 0$, e ainda, $b = 1; 2; \dots; B$, B representa o número de amostras independentes, e os valores referentes a $\tilde{\phi}_k^{RA}(b)$ são obtidos de forma iterativa, de acordo com Cryer e Chan (2008). Portanto, deve ocorrer a geração aleatória de valores no suporte dos conjuntos *fuzzy* ativados, mediante uma função de pertinência desejada, e os estimadores *fuzzy* são obtidos com a aplicação da função de autocorrelação estendida *fuzzy*.

Observação 4.1. Segundo Cryer e Chan (2008), como $\tilde{\phi}$ representa um estimador consistente para ϕ , então é possível concluir que $\tilde{\phi}$ representa um estimador “*assintoticamente não-viciado*” para ϕ , ou seja, visto que, $\lim_{n \rightarrow \infty} E(\tilde{\phi}_n) = \phi$ e $\lim_{n \rightarrow \infty} Var(\tilde{\phi}_n) = 0$. Neste contexto, para um “ n ” suficientemente grande, a relação $E(\tilde{\phi}_n) = \phi$ é razoável, equivalendo assim afirmar que em média, o estimador $\tilde{\phi}$ converge para o próprio parâmetro ϕ , desde que, o comprimento de trajetória da série temporal relacionada a uma característica de interesse tenda ao infinito.

4.6 - Um Novo Algoritmo Proposto Para Identificação de Modelos Fuzzy em Série Temporal

Nesta seção serão descritos os passos para se identificar de forma eficiente, um modelo de série temporal *fuzzy*, tal como, se realizará previsões mediante este modelo de série temporal *fuzzy* indicado pelo método desenvolvido. Será medido o erro de previsão alcançado com o modelo proposto e posteriormente, serão comparados os resultados destas previsões com os encontrados na literatura, para séries conhecidas na comunidade acadêmica, como por exemplo, o número de matrículas na universidade do Alabama nos Estados Unidos da América; o preço de energia elétrica em New South Wales na Austrália; o índice TAIFEX; etc. Será comparado o resultado das previsões obtidas, com os resultados já conhecidos na literatura, a partir de outras ferramentas de previsão propostas por outros autores, medi-

ante a metodologia de série temporal clássica ou *fuzzy*. A rotina proposta para previsões de séries temporais *fuzzy* é descrita abaixo.

Passo 1: Coletar os dados de interesse que formarão a série temporal;

Passo 2: Avaliar a possível estacionariedade da série temporal;

Passo 3: Definir o universo de discurso a partir da Equação (4.42).

$$U = [Z_{(Min)} - \sigma; Z_{(Max)} + \sigma] \quad (4.42)$$

onde, $Z_{(Min)}$ e $Z_{(Max)}$ representam o menor e o maior valor da série histórica observada, respectivamente, σ indica uma medida de variabilidade individual da série histórica analisada, ou seja, o desvio padrão da série, tal que, σ é expresso pela Equação (4.43).

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (Z_i - \bar{Z})^2}{n}} \quad (4.43)$$

Passo 4: Particionar o universo de discurso em “ m ” conjuntos *fuzzy*, tal que, a amplitude ou comprimento (l) de cada intervalo que representa um conjunto *fuzzy*, é obtido mediante a Equação (4.44), tal que, o valor de “ m ” é arbitrário e definido pelo especialista geralmente dado por $m \in [5; 9]$, como defende Miller (1994).

$$l = \frac{[Z_{(Max)} + \sigma] - [Z_{(Min)} - \sigma]}{m} \quad (4.44)$$

onde, $Z_{(Min)}$ e $Z_{(Max)}$ representam o menor e o maior valor da série histórica observada, respectivamente, σ é expresso pela Equação (4.43), e ainda, “ m ” representa o número de conjuntos *fuzzy* ou termos linguísticos a serem construídos à critério do especialista;

Observação 4.2. O valor dos “ m ” conjuntos *fuzzy* a serem construídos fica a critério do pesquisador, porém, Miller (1956) defende que o número de conjuntos *fuzzy* (termos linguísticos) deve ser igual à sete, com variação de mais ou menos dois, isto é, entre 5 e 9 conjuntos, pois esta é a capacidade que um ser humano possui, segundo este autor, para a classificação de conjuntos sem que haja prejuízo quanto a memorização dos mesmos.

Passo 5: Construir os conjuntos *fuzzy* a partir da definição de número *fuzzy* triangular, tal como, descrito nas Equações (3.5) e (3.6), para a série de dados históricos à se *fuzzificar*;

Passo 6: Escolher uma das funções de autocorrelação estendida *fuzzy*, que estão sendo propostas neste trabalho na Seção 4.5, para identificar a ordem do modelo de série temporal *fuzzy*;

- (a) Caso a função escolhida seja $W_{i,k,j}^{DFZ}$, defuzzificar a série de dados históricos mediante o método centróide e em seguida aplicar a função de autocorrelação estendida *fuzzy*.
- (b) Caso a função escolhida seja $W_{i,k,j}^{CA}(h; \alpha; b)$:
 - (b1) Atribuir o alfa-nível aos conjuntos *fuzzy* construídos no passo anterior, a partir do valor de alfa definido a priori pelo especialista, por exemplo, $\alpha = 0, 25; 0, 5$ ou $0, 75$;
 - (b2) Ativar os conjuntos *fuzzy* mediante o alfa-nível utilizado no passo anterior, para posteriormente definir o intervalo *fuzzy* correspondente a cada conjunto *fuzzy* A_j que foi ativado;
 - (b3) Gerar aleatoriamente 1000 funções de autocorrelação amostral, a partir do suporte do conjunto *fuzzy* ativado, calculando a média das 1000 funções em cada intervalo *fuzzy*, aplicando em seguida a função de autocorrelação estendida *fuzzy*, sobre a série das funções de autocorrelação amostrais médias resultantes, relacionadas a cada intervalo *fuzzy*, para obter a identificação da ordem na série temporal *fuzzy* de interesse;
- (c) Caso a função escolhida seja $\overline{W}_{i,k,j}^{RA}(h)$, se faz necessário gerar aleatoriamente 1000 valores aleatórios (dentro de cada suporte ativado) e obter as funções de autocorrelação amostrais, a partir do suporte do conjunto *fuzzy* ativado no passo anterior, calculando em seguida a média das 1000 funções em cada intervalo *fuzzy*, para na sequencia aplicar a função de autocorrelação estendida *fuzzy*.

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6 (a), 6 (b) ou 6 (c);

Passo 8: Medir a capacidade preditiva do modelo identificado, mediante a escala de erro médio absoluto (MASE), tal como, está definido em Hyndman e Koehler (2006).

No passo 6 citado anteriormente, está implícita a condição de estacionariedade da série temporal, tal como discutido na Seção 4.5, a partir dos resultados, definições e implicações defendidos por Hylleberg (1995), mesmo que a série temporal seja caracterizada em um ambiente *fuzzy*. Portanto, a série temporal deve ser livre de tendência, pois esta condição

é relevante para o processo de identificação e previsão, visto que, segundo Morettin e Tolo (2004), a função de autocorrelação só deve ser utilizada caso a série seja estacionária, algo que a própria configuração da função de autocorrelação estendida satisfaz, devido a sua estrutura padrão de identificação.

Alguns autores, como por exemplo, Song e Chisson (1993a), Song (2003), dentre outros, ignoram a suposição de linearidade e estacionariedade de uma série temporal, ao combinar este processo de modelagem estocástica com a metodologia dos conjuntos *fuzzy*.

Porém, Kai e Wen-Gang (2009) propõe um modelo heurístico de ordem “ n ” em uma série temporal *fuzzy*, para qual adotaram “ n ” diferenças nas observações e as tomam como informação heurística, para melhorar ainda mais a precisão das previsões do modelo de série temporal *fuzzy*.

Este fato ratifica fortemente a relevância quanto à estabilização de uma série histórica, para consequente estimação da série temporal *fuzzy* aos dados de interesse, pois, caso os dados representem um padrão não estacionário, as previsões geradas pelo modelo identificado apresentarão baixíssima acurácia, devido aos dados da série temporal sofrerem forte influência com o passar do tempo, como defendem também, por exemplo, Carvalho Jr. et al. (2012), Askari e Montazerin (2015), Abdollahzade et al. (2015), Lee e Hong (2015), Carvalho Jr. e Costa Jr. (2017).

4.7 - Conclusão do Capítulo

Neste capítulo abordaram-se os conceitos fundamentais para o desenvolvimento do estudo sobre série temporal *fuzzy*, mediante várias definições que caracterizam esta metodologia, em termos de conjuntos, equações relacionais, passando por um modelo de série temporal autoregressivo *fuzzy* com ordem 1, até um modelo de série temporal *fuzzy* sazonal de ordem “ p ”.

Destacou-se também, alguns dos principais algoritmos de previsão *fuzzy* via série temporal, disponíveis na literatura, sobretudo, os algoritmos que surgiram em decorrência direta da proposta apresentada pelos precursores desta linha de pesquisa, que são Song e Chisson (1993a e 1993b).

A partir destes algoritmos foi possível perceber a consolidação de um roteiro, para que as previsões *fuzzy* mediante série temporal sejam obtidas, incluindo neste roteiro desde a definição do universo de discurso até a *defuzzificação* dos valores preditos, de maneira que, a proposta original desenvolvida por Song e Chisson (1993a e 1993b), suscitou o interesse de diversos pesquisadores da comunidade científica, quanto à realização de previsões *fuzzy* via um modelo de série temporal.

Também neste capítulo foram definidas várias medidas de dependência entre os conjuntos *fuzzy*, sendo estas resultantes da função de autocorrelação amostral, parcial, inversa e estendida, respectivamente, tal que, a função de autocorrelação estendida *fuzzy* foi a principal contribuição deste trabalho.

A proposição de ferramentas que auxiliem na identificação de processos autoregressivos *fuzzy* de ordem “ p ”, de maneira mais eficiente quanto aos erros de previsão, em relação às ferramentas de identificação, que se encontram disponíveis na literatura específica, também foram consideradas por este estudo.

Incorporou-se ainda na função de autocorrelação estendida o conceito de α -nível e média aleatória, além dos valores *defuzzificados* (centróide), a partir da série histórica. Neste sentido, um novo método de identificação de processos autoregressivos *fuzzy* sazonais, foi proposto neste capítulo, com o intuito de corroborar com os algoritmos de previsão existentes na literatura, quanto à identificação dos modelos de série temporal no ambiente *fuzzy*.

No próximo capítulo, a metodologia desenvolvida neste trabalho será avaliada a partir de simulações, para assim mensurar a capacidade de identificação das ferramentas propostas. Desta forma, espera-se que a eficiência na identificação de processos autoregressivos sazonais *fuzzy*, possa ser plenamente comprovada, devido às ferramentas propostas apresentarem uma determinação da ordem dos parâmetros do modelo de previsão, com resultados plenamente satisfatórios, sobretudo, em relação aos métodos de identificação estabelecidos na literatura.

Simulações Mediante as Ferramentas Desenvolvidas Para a Identificação dos Modelos de Série Temporal Fuzzy

5.1 - Simulação de séries temporais fuzzy com várias quantidades de conjuntos *fuzzy* e combinações de parâmetros ϕ_1 e ϕ_2 .

Para que seja possível estabelecer a Função de Autocorrelação Estendida (EACF), como uma ferramenta eficiente de identificação dos processos Autoregressivos (AR), realizou-se simulações das funções $W_{i,k,j}^{DFZ}$; $W_{i,k,j}^{CA}(h; \alpha = 0, 5; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$, descritas pelas Equações (4.38), (4.39) e (4.41), respectivamente. As simulações realizadas para a EACF consistiram em uma rotina de programação, para gerar 1.000 processos autoregressivos de ordem 2, com quatro pares de combinações dos parâmetros ϕ_1 e ϕ_2 , escolhidos arbitrariamente, os quais foram iguais a: 1,6 e -0,8; 0,5 e -0,5; 0,5 e 0,3; 0,5 e 0,0.

E ainda, diversos tamanhos de termos linguísticos (conjuntos fuzzy) foram empregados ao processo (5; 7; 9; 23 e 47), onde o tamanho da série ou comprimento (T) das trajetórias simuladas (série temporal) foram iguais a 24, 36 e 96, para com isto, comparar o desempenho das funções de autocorrelação estendida quanto a identificação dos processos autoregressivos de segunda ordem, a partir do calculo das probabilidades de $P \left[W_{i,k,j}^{DFZ} \right] > 2T^{-\frac{1}{2}}$, $P \left[W_{i,k,j}^{CA}(h; \alpha; b) \right] > 2T^{-\frac{1}{2}}$ e $P \left[\overline{W}_{i,k,j}^{RA}(h) \right] > 2T^{-\frac{1}{2}}$. Neste sentido, as funções de autocorrelação representadas pelas Equações (4.38), (4.39) e (4.41), são utilizadas como medidas de dependência entre os conjuntos fuzzy, e, portanto, funcionam como uma ferramenta de identificação da ordem dos processos autoregressivos de ordem 2.

5.1.1 - Definição do plano experimental

Dentro deste estudo de simulação o objetivo principal será medir o valor das probabilidades, sobre a capacidade de determinação da ordem dos modelos autoregressivos de segunda ordem, mediante cada função de autocorrelação estendida utilizada neste estudo,

isto é, $P [W_{i,k,j}^{DF_Z}] > 2T^{-\frac{1}{2}}$, $P [W_{i,k,j}^{CA}(h; \alpha; b)] > 2T^{-\frac{1}{2}}$ e $P [\overline{W}_{i,k,j}^{RA}(h)] > 2T^{-\frac{1}{2}}$. Neste contexto, a probabilidade observada em cada combinação de parâmetros ϕ_1 e ϕ_2 , para m conjuntos fuzzy, com tamanho de trajetória T , indicará o quanto a ferramenta proposta neste trabalho determinou corretamente a ordem do modelo AR(2), e, quando ϕ_1 ou ϕ_2 for igual à zero, o objetivo é avaliar se a ferramenta proposta está identificando um modelo AR(1) como um AR(2), isto é, determinando de forma errada a ordem do modelo.

Para que fosse possível avaliar o desempenho das funções definidas neste trabalho, o software Matlab®R2010B, foi utilizado na produção de amostras independentes, as quais seguem uma distribuição de probabilidade normal, com média 0 (zero) e variância 1 (um), procedimento semelhante ao utilizado por Lago Neto (2011). Neste contexto, as amostras que foram produzidas geraram observações para diversos modelos AR(2), ou seja, modelos autoregressivos de médias móveis de segunda ordem. Com o intuito de suprimir o possível efeito de inicialização dos valores durante as simulações, os 100 dados inicialmente gerados foram excluídos em cada uma das séries. Para estas simulações foram adotados os tamanhos de trajetória das séries temporais $T = 24$, $T = 36$ e $T = 96$. Arbitrariamente, com o intuito de preencher a superfície dos parâmetros autoregressivos (ϕ_1 e ϕ_2) serão adotados, respectivamente os valores (1,6 e -0,8); (0,5 e -0,5); (0,5 e 0,3); (0,5 e 0,0).

Um processo de *fuzzificação* com formato triangular será aplicado sobre a sequência de T dados simulados, os quais representarão um modelo autoregressivo de ordem 2. A justificativa para *fuzzificar* os dados autoregressivos das simulações transformando-os em observações fuzzy, é devido o propósito de gerar séries temporais fuzzy simuladas, o que necessariamente tornará possível obter resultados de modelos autoregressivos, que indicarão na prática conjuntos fuzzy.

5.1.2 - Série temporal fuzzy com função $W_{i,k,j}^{DF_Z}$, $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$, para $T = 24$, $m = 5; 7; 9; 23; 47$, $\phi_1 = 1,6; 0,5; 0,5; 0,5$ e $\phi_2 = -0,8; -0,5; 0,3; 0,0$, respectivamente.

Após diversas simulações geradas pela rotina de identificação proposta neste trabalho, se tornou possível observar o desempenho da função de autocorrelação estendida inicialmente em relação aos valores defuzzificados da série $W_{t,k,j}^{DF_Z}$, onde se percebeu que, para trajetórias de comprimento $T = 24$, o melhor desempenho observado alcançou 25%, quando o número de conjuntos fuzzy (m) foi igual a 47, com a combinação de parâmetros: $\phi_1 = 1,6$ e

$\phi_2 = -0,8$, e ainda, o segundo melhor desempenho desta função, se obteve quando $m = 23$, com a mesma combinação de parâmetros já citada, onde o desempenho da função passou a ser de 19,44%, como se observa na Tabela (5.1). Assim, um número elevado do “ m ” (23 ou 47) resulta em um bom desempenho desta função.

Tabela 5.1 - Frequência relativa (%) dos eventos $P | W_{i,k,j}^{DFZ} | > 2T^{-\frac{1}{2}}$, $P | W_{i,k,j}^{CA}(h; \alpha; b) | > 2T^{-\frac{1}{2}}$ e $P | \overline{W}_{i,k,j}^{RA}(h) | > 2T^{-\frac{1}{2}}$ em 1.000 simulações de processos autoregressivo de ordem 2, com o comprimento da série $T = 24$ ($2T^{-\frac{1}{2}} = 0,4082$) para valores de $m = 5; 7; 9; 23; 47$ (termos linguísticos ou conjuntos fuzzy) e quatro combinações dos parâmetros ϕ_1 e ϕ_2 .

EACF	m	$\phi_1 = 0,5$ e $\phi_2 = -0,5$	$\phi_1 = 1,6$ e $\phi_2 = -0,8$	$\phi_1 = 0,5$ e $\phi_2 = 0,0$	$\phi_1 = 0,5$ e $\phi_2 = 0,3$
$W_{i,k,j}^{DFZ}$	5	5,56	5,56	11, 11 ^{1;2}	5,56
	7	11, 11 ²	5,56	2,78	11, 11 ²
	9	13, 89 ¹	2,78	8,33	16, 67 ^{1;2}
	23	11, 11	19, 44 ²	5,56	8,33
	47	11, 11	25, 00 ^{1;2;3}	5,56	16, 67 ¹
$W_{i,k,j}^{CA}(h; \alpha^*; b)$	5	8,33	19, 44 ²	19, 44 ^{1;2}	5,56
	7	11, 11	25, 00 ^{1;2;3}	16,67	11, 11
	9	19, 44 ^{1;2}	5,56	5,56	13,89
	23	11, 11 ²	8,33	8,33	11, 11 ²
	*Nota: $\alpha = 0,5$	47	5,56	16,67	11, 11
$\overline{W}_{i,k,j}^{RA}(h)$	5	11, 11	8,33	8,33	13, 89 ²
	7	11, 11	11, 11 ¹	11, 11 ¹	13, 89 ²
	9	8,33	8,33	11, 11 ^{1;2}	11, 11 ²
	23	19, 44 ^{1;2;3}	11, 11 ¹	5,56	16, 67 ¹
	47	5,56	5,56	11, 11 ¹	13, 89 ²

Nota: ¹ Melhor desempenho quanto à identificação da ordem de um processo autoregressivo de ordem 2, em cada uma das três funções de autocorrelação estendida ajustadas ($W_{i,k,j}^{DFZ}$; $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$), respectivamente, em relação à combinação de parâmetros (ϕ_1 e ϕ_2) na coluna “ j ” e quantidade específica de conjuntos fuzzy (termos linguísticos) na linha “ k ”. ² Melhor desempenho quanto à identificação da ordem de um processo autoregressivo de ordem 2, em cada quantidade específica de conjuntos fuzzy (linha “ k ”), dentro de cada uma das três função de autocorrelação estendida ajustadas ($W_{i,k,j}^{DFZ}$; $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$), respectivamente, em relação as quatro cominações de parâmetros (ϕ_1 e ϕ_2), que foram simuladas. ³ Melhor desempenho geral quanto à identificação da ordem de um processo autoregressivo de ordem 2, de cada uma das três função de autocorrelação estendida ajustadas ($W_{i,k,j}^{DFZ}$; $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\overline{W}_{i,k,j}^{RA}(h)$), respectivamente.

Quando a função de identificação utilizada é $W_{i,k,j}^{CA}(h; \alpha; b)$, com $T = 24$, o melhor desempenho observado foi para combinação de parâmetros $\phi_1 = 1,6$ e $\phi_2 = -0,8$, com $m = 7$, que resultou em 25% de eficiência (Ver, Tabela (5.1)). O segundo melhor desempenho desta função foi obtido quando, se reduziu o número de termos linguísticos para 5, onde a eficiência quanto a identificação de processos AR(2) passou a ser de 19,44%, se mantendo a mesma combinação de parâmetros $\phi_1 = 1,6$ e $\phi_2 = -0,8$, que foi utilizada

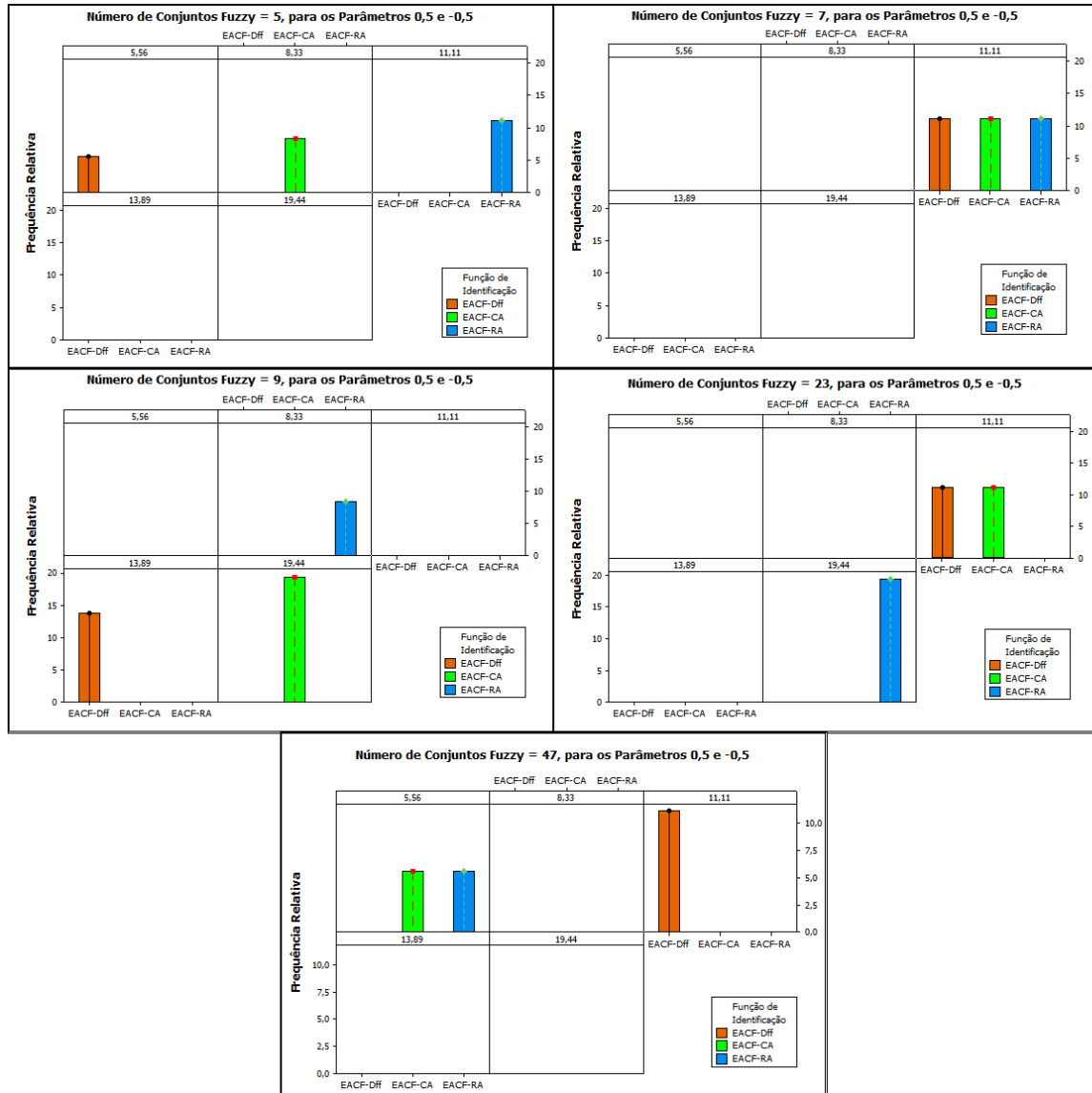
quando $m = 7$. Caso $m = 5$ e a combinação de parâmetros seja $\phi_1 = 0,5$ e $\phi_2 = 0,0$, a eficiência de 19,44% não se altera. Continuando a utilizar esta função de identificação, se torna possível alcançar a mesma eficiência de 19,44%, porém, com uma quantidade superior de conjuntos fuzzy, em relação ao que se apresentou anteriormente, no caso, $m = 9$ com $\phi_1 = 0,5$ e $\phi_2 = -0,5$, ou $m = 47$ com $\phi_1 = 0,5$ e $\phi_2 = 0,3$, se tem os mesmos 19,44% de eficiência, o que pelo princípio da parcimônia não é indicado utilizar neste cenário, pois a complexidade matemática e o esforço computacional dos processos de estimação paramétrica para $m = 9$ ou $m = 47$ serão superiores a um processo de identificação com $m = 5$, visto que, este último apresenta uma eficiência idêntica, com muito mais simplicidade no processo de identificação. Portanto, se a função para identificação tem a incorporação em sua estrutura do alfa-nível, desde que, a série temporal fuzzy seja de baixa duração, ou seja, uma série temporal de trajetória curta ($T \leq 24$), ao utilizar sete termos linguísticos, o melhor desempenho desta função quanto à identificação de processos autoregressivos é alcançado plenamente.

Segundo a Tabela (5.1), caso a função de identificação escolhida para um processo autoregressivo seja $\overline{W}_{i,k,j}^{RA}(h)$, desde que, $T \leq 24$, o melhor desempenho foi alcançado quando $m = 23$, com os parâmetros $\phi_1 = 0,5$ e $\phi_2 = -0,5$, o que gerou uma eficiência no processo de identificação igual a 19,44%. Caso se mantenha o número de termos linguísticos $m = 23$, no entanto, com a combinação de parâmetros $\phi_1 = 0,5$ e $\phi_2 = 0,3$, o segundo melhor desempenho quanto à identificação de processos AR(2) será obtido com 16,67%. Então, caso a função de autocorrelação estendida média aleatória fuzzy, seja a escolhida para identificar um modelo de processo AR(2), segundo a Tabela (5.1), o indicado é utilizar 23 conjuntos fuzzy durante o processo de *fuzzificação* da série temporal de trajetória $T = 24$. Neste sentido, se pode concluir que para uma série temporal fuzzy curta ($T \leq 24$) a eficiência da função $W_{i,k,j}^{DFZ}$ é alcançada, desde que o número de conjuntos fuzzy seja grande ($m \geq 23$). Em termos gerais, se a série temporal fuzzy for curta ($T \leq 24$), o ideal como mostra a Tabela (5.1) é utilizar para a identificação de processos AR(2), a função $W_{i,k,j}^{DFZ}$ ou $W_{i,k,j}^{CA}(h; \alpha; b)$ ou $\overline{W}_{i,k,j}^{RA}(h)$, com $m = 47$; $m = 7$; $m = 23$, respectivamente.

Mediante as informações contidas na Tabela (5.1), se torna possível representar graficamente os seus resultados, tal como, se observa nas Figuras (5.1-5.4), de maneira que, para cada combinação de parâmetros que foram utilizadas nas simulações, se tem uma

frequência relativa associada à eficiência na identificação do modelo, em um processo autoregressivo. Assim, a partir das cinco quantidade de termos linguísticos ($m = 5; 7; 9; 23$ e 47) simuladas neste estudo, com um comprimento de trajetória $T = 24$ (tamanho da série temporal simulada) observam-se os resultados a seguir.

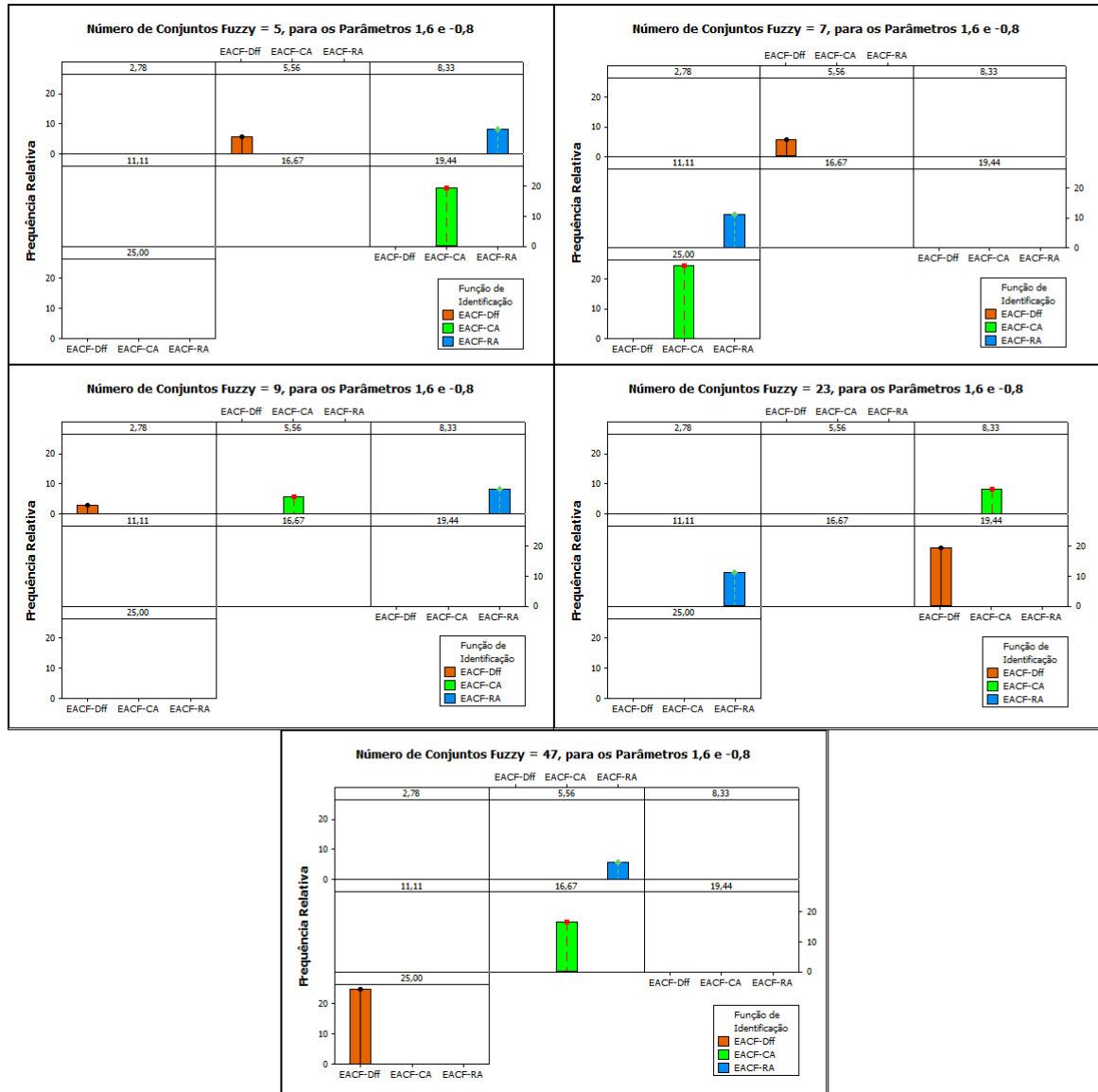
Figura 5.1 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = -0,5$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$.



Analisando a Figura (5.1) é possível perceber que quando $m = 5$ (cinco conjuntos fuzzy), a ferramenta com melhor desempenho na determinação do modelo AR (2), foi a função média aleatória com frequência de 11,11%. Com $m = 7$, as três ferramentas possuem desempenho idêntico de 11,11%. Quando o número de conjuntos fuzzy é igual

a 9, a melhor ferramenta foi a com α -nível, que determinou a ordem de um processo AR(2) com frequência de 19,44%. Porém, com $m = 23$, a função média aleatória foi a mais eficiente na identificação do processo AR(2) com 19,44%. Finalmente, quando $m = 47$ a ferramenta com dados defuzzificados retornou uma frequência de 11,11%, a qual representou o melhor resultado obtido dentro das possibilidades.

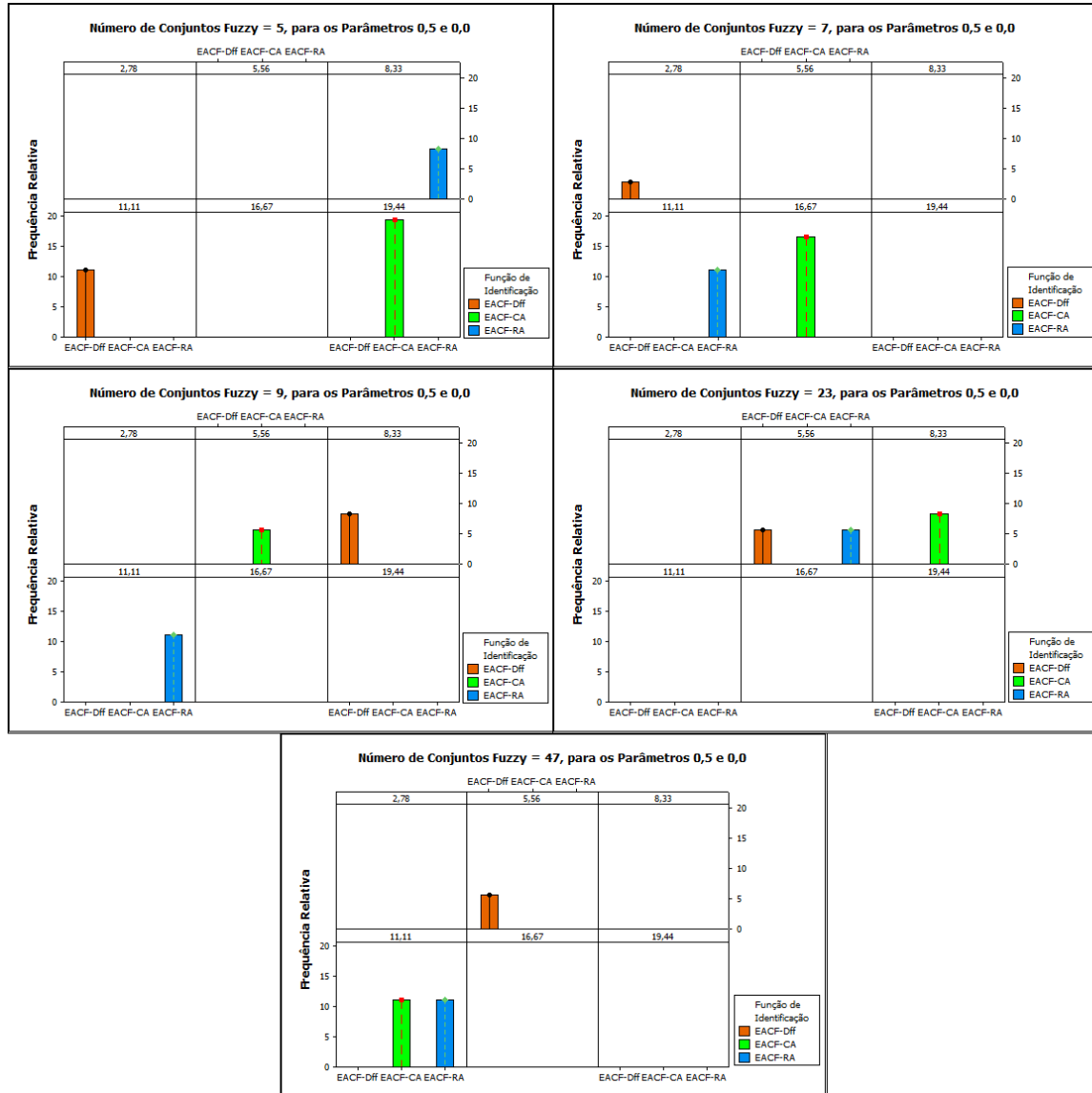
Figura 5.2 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 1,6$ e $\phi_2 = -0,8$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$.



Mediante a Figura (5.2), a ferramenta de determinação da ordem dos modelos AR(2) com α -nível, apresentou o melhor desempenho, quando $m = 5$ ou $m = 7$, com frequências de 19,44% e 25%, respectivamente. A ferramenta proposta com a função média aleatória

foi superior às demais, apenas quando o número de conjuntos fuzzy foi igual a 9, nesta combinação de parâmetros específica. Assim, quando $m = 23$ ou $m = 47$, a ferramenta de dados defuzzificados retornou os melhores resultados, com 19,44% e 25%, respectivamente.

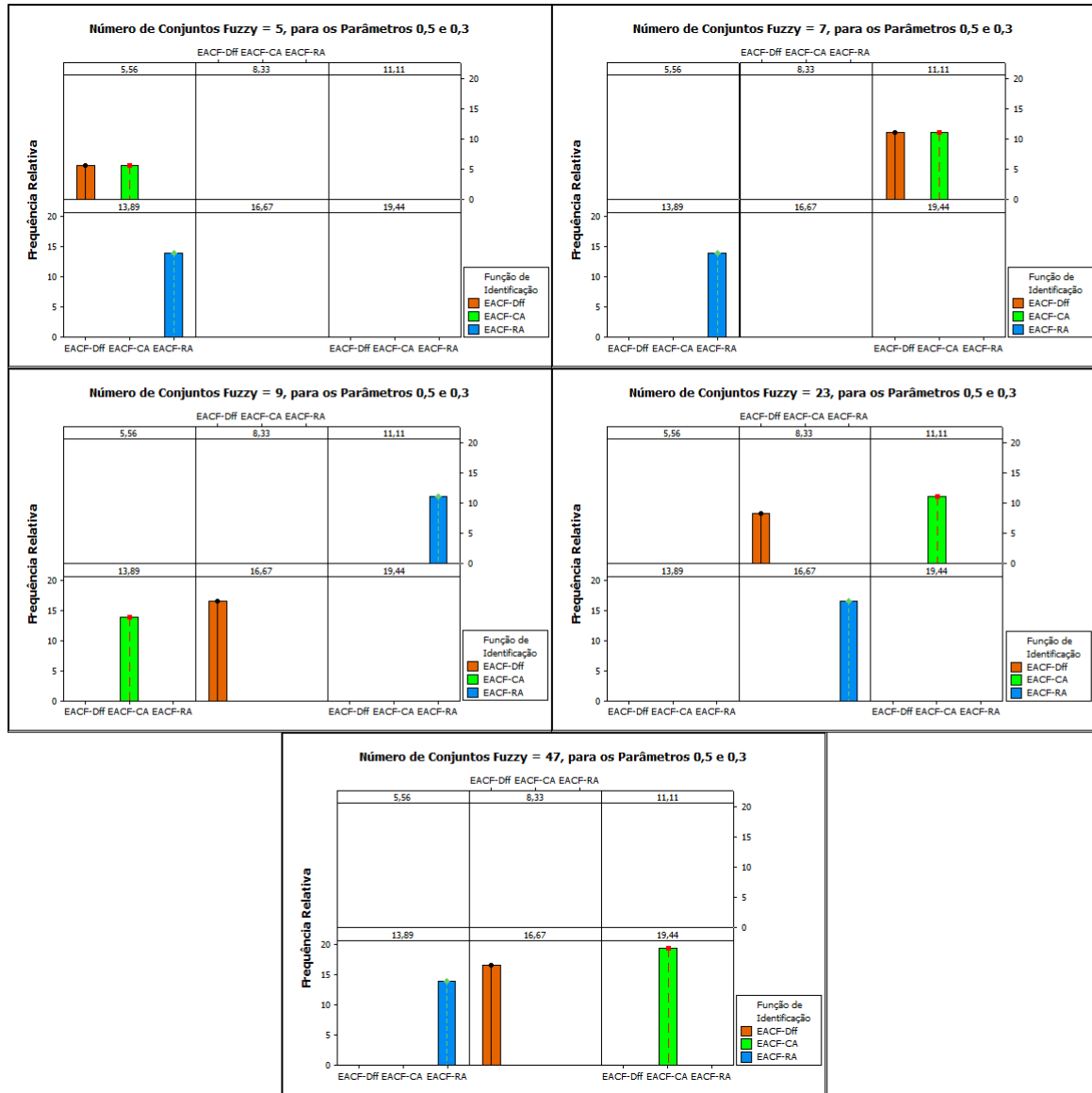
Figura 5.3 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = 0,0$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$.



A Figura (5.3) indica se as ferramentas de determinação da ordem propostas neste trabalho estão identificando processos AR(1), isto é, $\phi_1 = 0,5$ e $\phi_2 = 0,0$, como sendo processos AR(2). Assim, é possível perceber que para conjuntos fuzzy igual a: $m = 5$, $m = 7$ ou $m = 23$, a função com α -nível identificou equivocadamente o processo AR(1), como sendo um AR(2), com frequência de 19,44%, 16,67% e 8,33%, respectivamente. Por

sua vez, a maior frequência de identificação errada de AR(1) como AR(2), quando $m = 9$, foi observada com a função média aleatória. Finalmente, quando o $m = 47$, as funções com α -nível e média aleatória se equivaleram, quanto à determinação equivocada de um processo AR(2) com frequência de 11,11%.

Figura 5.4 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = 0,3$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 24$.



Analisando a Figura (5.4), é possível perceber que a função média aleatória determinou de forma mais satisfatória a ordem dos processos AR(2), em relação as demais funções, pois apresentou uma frequência na identificação dos processos igual a 13,89%, 13,89% e

16,67%, para $m = 5$, $m = 7$ e $m = 23$, respectivamente. Quando o número de conjuntos fuzzy foi de $m = 9$, o melhor resultado observado foi à frequência de 16,67%, para a função com dados *defuzzificados*. Por fim, quando $m = 47$, a função para determinação da ordem com α -nível, foi a que apresentou o melhor resultado para identificação do processo AR(2) com uma frequência de 19,44%.

5.1.3 - Série temporal fuzzy com função $W_{i,k,j}^{DF_Z}$, $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\bar{W}_{i,k,j}^{RA}(h)$, para $T = 36$, $m = 5; 7; 9; 23; 47$, $\phi_1 = 1,6; 0,5; 0,5; 0,5$ e $\phi_2 = -0,8; -0,5; 0,3; 0,0$, respectivamente.

Com um comprimento de trajetória $T = 36$, o melhor desempenho da função $W_{i,k,j}^{DF_Z}$ para os valores *defuzzificados* da série, ocorreu quando a combinação de parâmetros foi igual a $\phi_1 = 1,6$ e $\phi_2 = -0,8$, com $m = 47$, o que resultou em 30,56% de eficiência, seguido da mesma combinação de parâmetros anterior, porém, com $m = 7$, que gerou 27,78% de eficiência na identificação, como pode ser verificado na Tabela (5.2). Assim, caso a opção dentre as ferramentas de identificação de processos autoregressivos seja pela função $W_{i,k,j}^{DF_Z}$, o indicado é utilizar $m = 47$ ou $m = 7$, para assim obter uma eficiência de aproximadamente 30,56% ou 27,78%, respectivamente, caso a série possua trajetória $T = 36$. Porém, como a diferença entre o desempenho desta função com $m = 47$ e $m = 7$, não é muito elevada, ou seja, 2,78%, se torna mais prático e simples utilizar 7 conjuntos fuzzy, quando $T = 36$ e a função de identificação seja $W_{i,k,j}^{DF_Z}$ para um processo AR(2).

Quando a função de identificação escolhida foi $W_{i,k,j}^{CA}(h; \alpha; b)$, o melhor desempenho alcançado se deu quando $m = 5$ com os parâmetros $\phi_1 = 0,5$ e $\phi_2 = -0,5$, em uma trajetória $T = 36$, pois esta estrutura gerou 27,78% de eficiência, como mostra a Tabela (5.2). Caso o número de conjuntos fuzzy seja aumentado para 9, com esta mesma combinação de parâmetros o desempenho desta função reduz pouco, passa a ser de 25%, porém com uma partição maior do universo de discurso em relação à $m = 5$, ou seja, um número maior de termos linguísticos a serem formados no processo de *fuzzificação* da série temporal. Neste contexto, isto representará um ganho em termos de alocação de cada valor do universo de discurso, em um conjunto fuzzy específico. Logo, cada conjunto terá valores mais próximos entre si, o que irá resultar em uma maior homogeneidade dos valores observados em cada conjunto. O mesmo desempenho de 25% quanto à identificação de um processo AR(2) pela função $W_{i,k,j}^{CA}(h; \alpha; b)$, se observa na Tabela (5.2) quando o $m = 7$, em combinação aos parâmetros $\phi_1 = 0,5$ e $\phi_2 = 0,0$, logo, caso a opção seja por uma quantidade mediana

de conjuntos *fuzzy* entre $5 \leq m \leq 9$, com razoável desempenho quanto à identificação, se pode optar em utilizar na partição do universo de discurso $m = 7$.

Tabela 5.2 - Frequência relativa (%) dos eventos $P | W_{i,k,j}^{DFZ} | > 2T^{-\frac{1}{2}}$, $P | W_{i,k,j}^{CA}(h; \alpha; b) | > 2T^{-\frac{1}{2}}$ e $P | \bar{W}_{i,k,j}^{RA}(h) | > 2T^{-\frac{1}{2}}$ em 1.000 simulações de processos autoregressivo de ordem 2, com o comprimento da série $T = 36$ ($2T^{-\frac{1}{2}} = 0,3333$) para valores de $m = 5; 7; 9; 23; 47$ (termos linguísticos ou conjuntos fuzzy) e quatro combinações dos parâmetros ϕ_1 e ϕ_2 .

EACF	m	$\phi_1 = 0,5$ e $\phi_2 = -0,5$	$\phi_1 = 1,6$ e $\phi_2 = -0,8$	$\phi_1 = 0,5$ e $\phi_2 = 0,0$	$\phi_1 = 0,5$ e $\phi_2 = 0,3$
$W_{i,k,j}^{DFZ}$	5	19,44	22,22 ²	11,11	22,22 ²
	7	22,22	27,78 ²	13,89	16,67
	9	16,67	11,11	19,44 ²	13,89
	23	22,22 ²	16,67	13,89	13,89
	47	19,44	30,56 ^{2;3}	11,11	13,89
$W_{i,k,j}^{CA}(h; \alpha^*; b)$	5	27,78 ^{2;3}	22,22	19,44	13,89
	7	11,11	16,67	25,00 ^{1;2}	16,67
	9	25,00 ²	16,67	8,33	8,33
	23	16,67 ²	13,89	16,67 ²	13,89
	47	22,22 ²	13,89	16,67	11,11
$\bar{W}_{i,k,j}^{RA}(h)$	5	19,44	22,22	25,00 ^{1;2}	13,89
	7	16,67	33,33 ^{1;2;3}	22,22	27,78 ¹
	9	19,44 ²	19,44 ²	19,44 ²	19,44 ²
	23	30,56 ^{1;2}	13,89	13,89	19,44
	47	11,11	11,11	25,00 ^{1;2}	25,00 ²

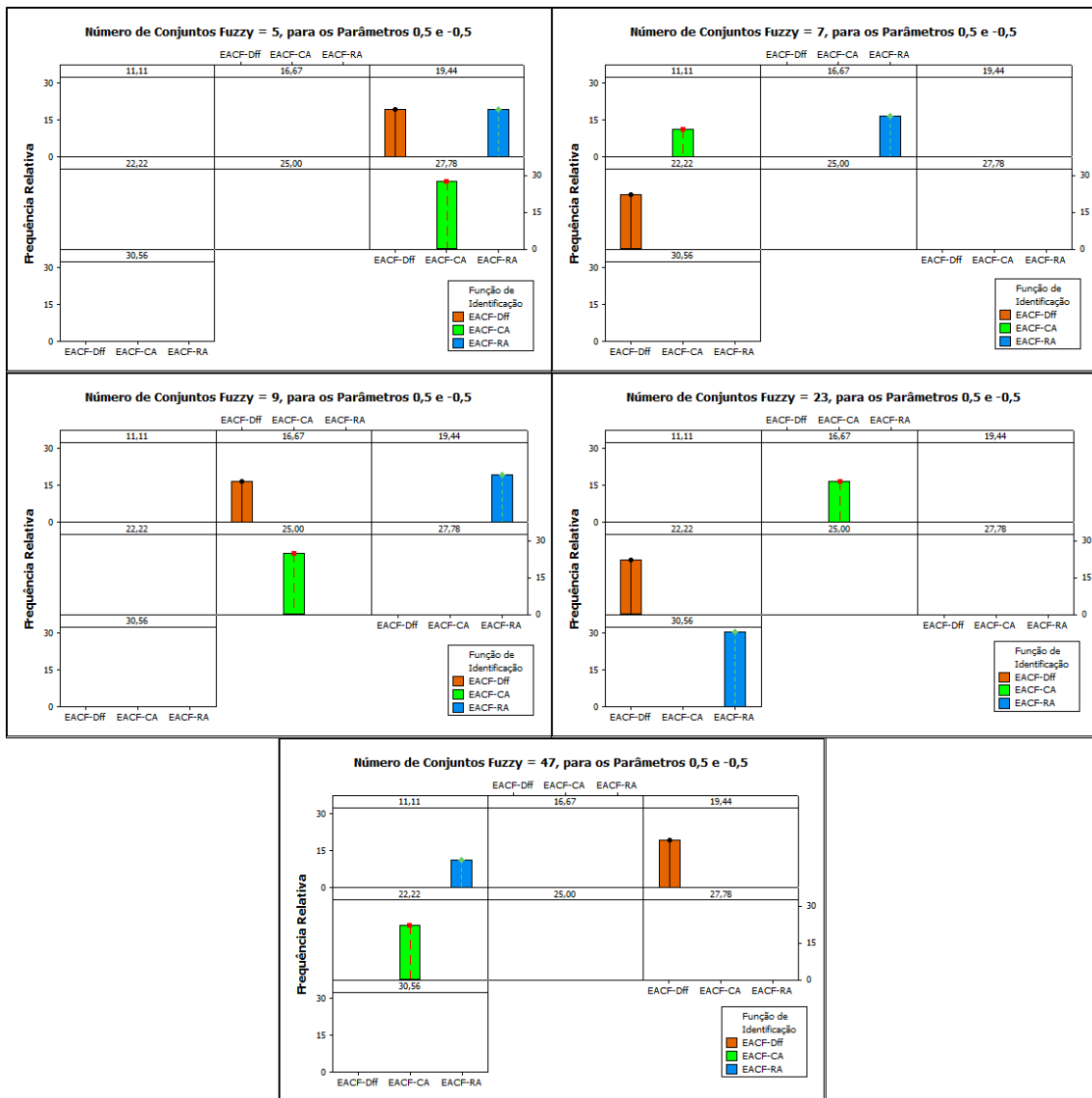
Nota: 1, 2 e 3 são tais como definidos na Tabela 5.1.

Para a função de identificação $\bar{W}_{i,k,j}^{RA}(h)$, o melhor desempenho segundo a Tabela (5.2) foi alcançado com a combinação de parâmetros $\phi_1 = 1,6$, $\phi_2 = -0,8$ e $m = 7$, que resultou em uma eficiência de 33,33%. Em segundo lugar quanto ao desempenho desta função, com $m = 23$, $\phi_1 = 0,5$ e $\phi_2 = -0,5$, se obteve 30,56%, o que em termos práticos representa uma rotina menos eficiente, não apenas pelo resultado do desempenho em si, mas também pela quantidade superior de conjuntos *fuzzy*, em relação à estrutura de identificação anterior, que necessitou de apenas 7 conjuntos para obter um desempenho superior de 33,33%. Caso se deseje trabalhar com a função de identificação $\bar{W}_{i,k,j}^{RA}(h)$ e $m = 7$, por se tratar de uma quantidade mediana entre 5 e 9 conjuntos, segundo as simulações a eficiência desta ferramenta será em torno de 27,78%.

Portanto, caso a série seja considerada de trajetória mediana $24 < T \leq 36$, a função de identificação que apresentou o melhor desempenho dentre as três utilizadas nas simulações, foi a função $\bar{W}_{i,k,j}^{RA}(h)$ com $m = 7$, pois mediante esta função se obteve o maior desempenho quanto à identificação de processos AR(2), que ficou em torno de 33,33%. Ou seja, a maior eficiência quanto à identificação a partir de uma quantidade pequena de termos linguísticos

no processo de *fuzzificação*. Em segundo lugar, se pode utilizar esta mesma função para obter um desempenho de 30,56%, porém com 23 termos linguísticos. Ou ainda, se pode chegar ao mesmo desempenho de 30,56%, mediante a função $W_{i,k,j}^{DFz}$ com 47 conjuntos fuzzy. Neste sentido, caso $T = 36$ se deve utilizar a função $\overline{W}_{i,k,j}^{RA}(h)$ com $m = 7$, para se obter o maior desempenho possível na identificação de processos autoregressivos dentre as três ferramentas utilizadas nas simulações.

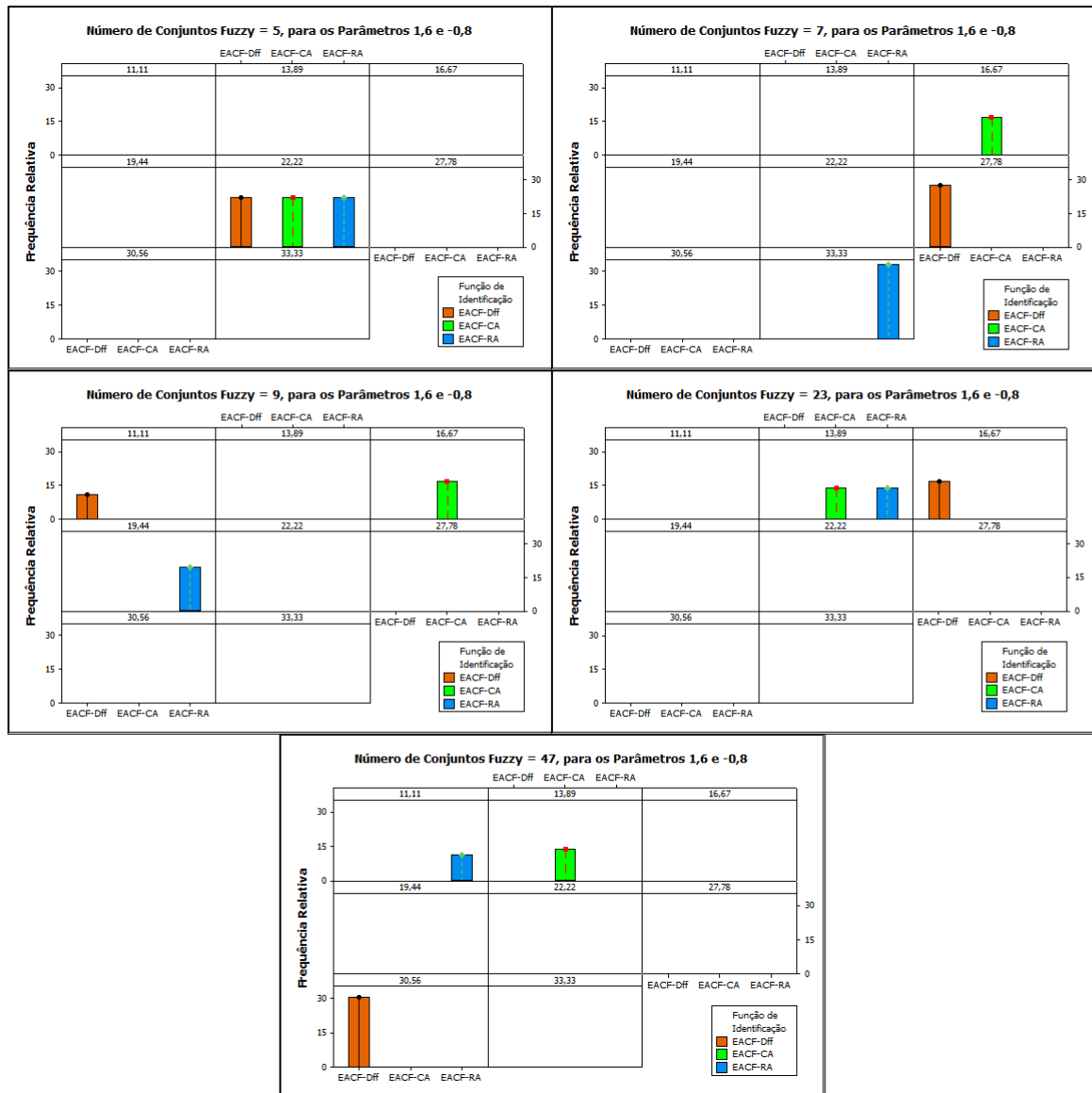
Figura 5.5 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = -0,5$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$.



Com base nas informações contidas na Tabela (5.2), se torna possível representar graficamente os seus resultados, tal como, se observa nas Figuras (5.5-5.8), de maneira que,

para cada combinação de parâmetros que foram utilizados nas simulações, se tem uma frequência relativa associada à eficiência na identificação da ordem, em um processo autoregressivo com ordem 2, a partir das cinco quantidade de termos linguísticos ($m = 5; 7; 9; 23$ e 47) simulados neste estudo, em um comprimento de trajetória $T = 36$ (tamanho da série temporal simulada).

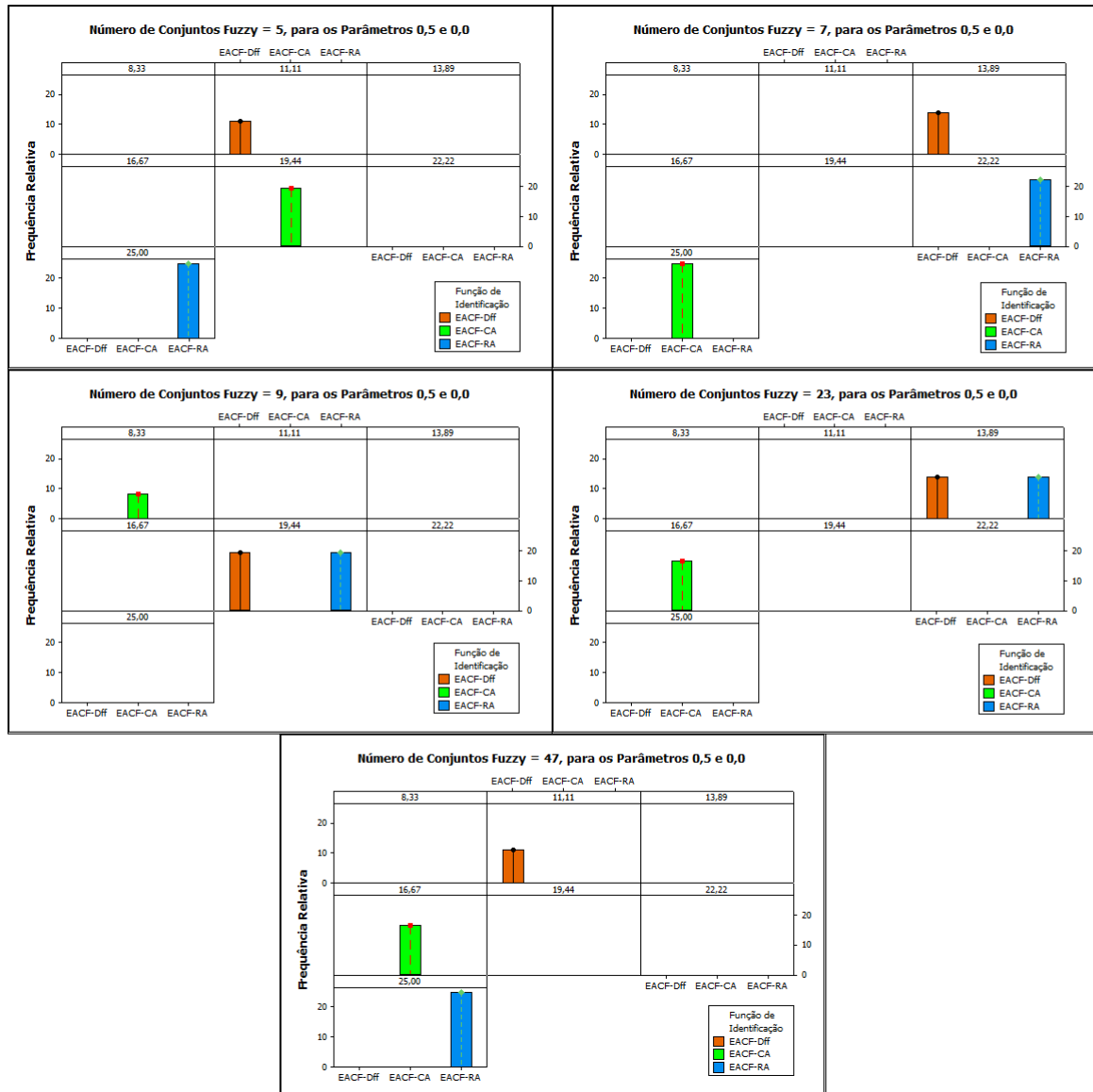
Figura 5.6 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 1,6$ e $\phi_2 = -0,8$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$.



Pela Figura (5.5), é possível perceber que a função com α -nível possui desempenho superior às demais, quando $m = 5$ (27,78%); $m = 9$ (25%) e $m = 47$ (22,22%). Com sete conjuntos fuzzy, a melhor função foi a com dados defuzzificados (16,67%), e quando

$m = 23$, a ferramenta mais eficiente na determinação do processo AR(2) foi a função média aleatória (16,67%).

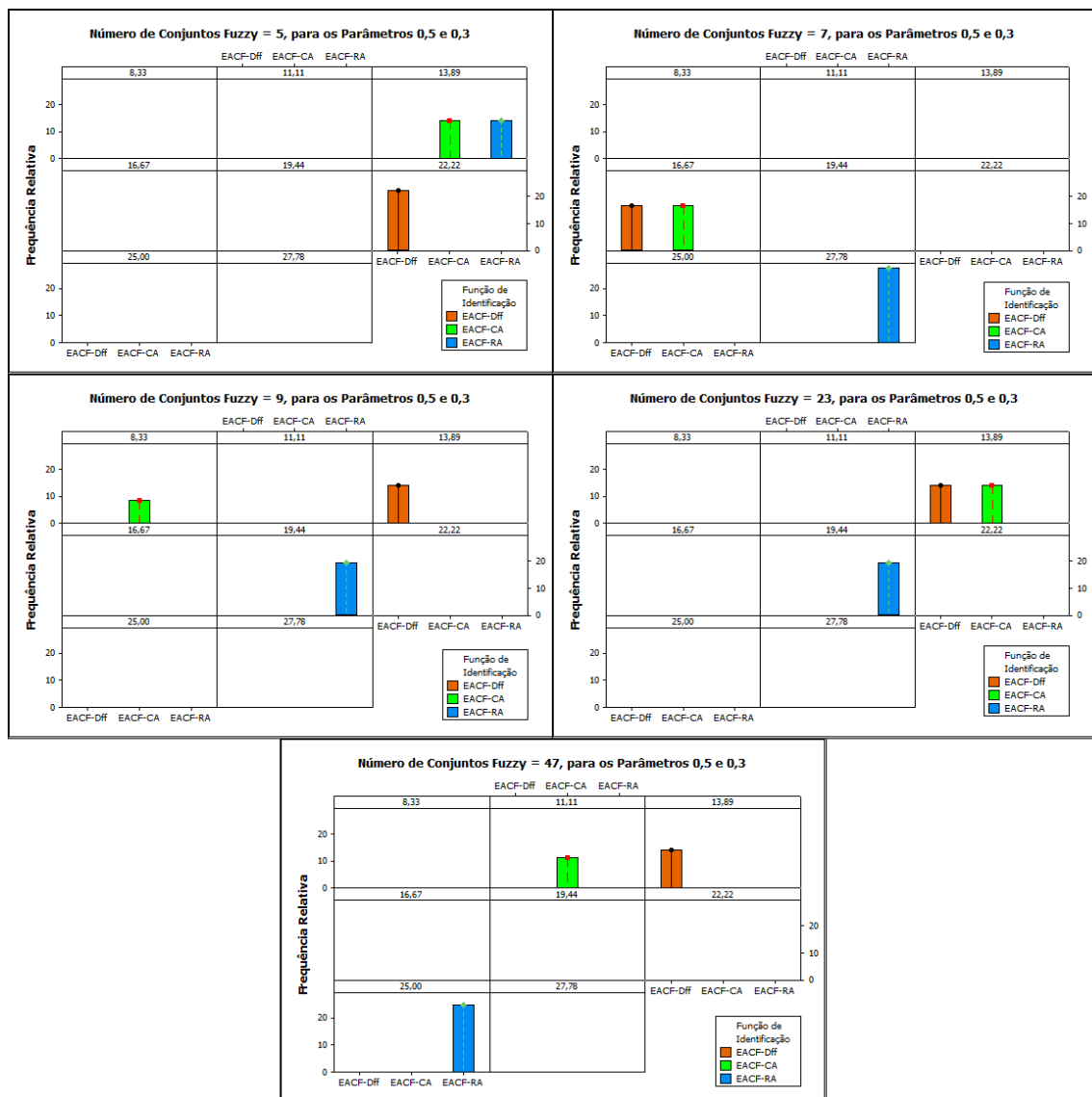
Figura 5.7 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = 0,0$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$.



Na Figura (5.6) é possível verificar que o desempenho das três funções propostas foi semelhante, quando o número de conjuntos fuzzy foi igual a 5 (22,22%). No entanto, aumentando o número de conjuntos fuzzy para 7 ou 9, a função média aleatória é a mais eficiente na determinação do processo AR(2), com frequências de 33,33% e 19,44%, respectivamente. E ainda, com $m = 23$ ou $m = 47$, a função com melhor desempenho foi a com dados defuzzificados, pois a frequência observada foi de 16,67% e 30,56%, respectivamente.

Na Figura (5.7) testa-se o erro na determinação da ordem dos processos AR(1), como se fossem processos AR(2), onde a função média aleatória apresentou os piores resultados (25%), quando o número de conjuntos fuzzy foi $m = 5$ ou $m = 47$. Porém, quando $m = 7$ ou $m = 23$, a função com pior desempenho, isto é, que identificou com mais erro um processo AR(2), foi a função com α -nível retornando uma frequência de 25% e 16,67%, respectivamente. Para um valor de $m = 9$, tanto a função com dados defuzzificados, quanto a função média aleatória apresentaram os piores desempenhos com frequência de 19,44%.

Figura 5.8 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = 0,3$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 36$.



Os resultados apresentados na Figura (5.8) indicam que a função média aleatória apresentou o melhor desempenho, quanto a identificação satisfatória dos processos AR(2). Quando o número de conjuntos fuzzy foi $m = 7$, $m = 9$, $m = 23$ ou $m = 47$, a função média aleatória retornou os melhores resultados iguais a: 27,78%, 19,44%, 19,44% e 25%, respectivamente. Para $m = 5$, a função com α -nível e a função média aleatória são equivalente quantitativamente, com desempenho de 13,89% na eficiência da determinação correta para ordem dos processos AR(2).

5.1.4 - Série temporal fuzzy com função $W_{i,k,j}^{DF_Z}$, $W_{i,k,j}^{CA}(h; \alpha; b)$ e $\bar{W}_{i,k,j}^{RA}(h)$, para $T = 96$, $m = 5; 7; 9; 23; 47$, $\phi_1 = 1,6; 0,5; 0,5; 0,5$ e $\phi_2 = -0,8; -0,5; 0,3; 0,0$, respectivamente.

Quanto a uma série temporal de longa duração onde $T = 96$, os melhores desempenhos observados após as simulações foram alcançados segundo a Tabela (5.3), para a função $W_{i,k,j}^{DF_Z}$, quando $\phi_1 = 1,6$ e $\phi_2 = -0,8$, com $m = 9$, tal que o desempenho desta função de identificação ficou em torno de 55,56%. E ainda, o segundo melhor desempenho obtido quanto à identificação de um processo autoregressivo chegou a 52,78%, com os parâmetros $\phi_1 = 0,5$ e $\phi_2 = -0,5$, para o mesmo número de termos linguísticos observados na composição anterior ($m = 9$). Portanto, caso a série temporal fuzzy seja de longa duração ($T > 36$), a eficiência quanto a identificação de um processo AR(2), a partir da função $W_{i,k,j}^{DF_Z}$, é obtida para um número de conjuntos fuzzy dentro do intervalo $5 \leq m \leq 9$, o que satisfaz o paradigma estabelecido por Miller (1994). Este autor por sua vez sugere um número 7 mais ou menos 2, que um ser humano é capaz de classificar (definir) em um processo de forma a memorizá-los sem risco de perda, o que pode ser associado plenamente na rotina de *fuzzificação* quanto ao número de conjuntos fuzzy.

Se a opção pela identificação de um processo autoregressivo for mediante a função $W_{i,k,j}^{CA}(h; \alpha; b)$, em uma série temporal onde $T = 96$, o melhor desempenho observado após as simulações realizadas, segundo a Tabela (5.3), ocorreu quando $\phi_1 = 1,6$ e $\phi_2 = -0,8$, com $m = 23$, visto que, o desempenho ficou em 50,00%. Em segundo lugar se observou um desempenho de 47,22%, com os mesmos 23 termos linguísticos para o processo de *fuzzificação* da série temporal, porém, com os parâmetros $\phi_1 = 0,5$ e $\phi_2 = 0,3$. Estes resultados sugerem a necessidade de estabelecer uma quantidade razoável de termos linguísticos, ou seja, uma quantidade eficiente de conjuntos fuzzy a ser definida para que esta função con-

significa alcançar um desempenho significativo, em termos de identificação para um processo autoregressivo de uma série temporal fuzzy.

Para a função de autocorrelação estendida fuzzy $\overline{W}_{i,k,j}^{RA}(h)$, o melhor desempenho observado desta função foi atingido quando os parâmetros $\phi_1 = 0,5$ e $\phi_2 = 0,0$; com $m = 47$, o que resultou em 55,56% de eficiência, segundo a Tabela (5.3). Porém, o segundo melhor desempenho desta função de identificação foi obtido em três situações específicas: quando os parâmetros $\phi_1 = 0,5$ e $\phi_2 = -0,5$; com $m = 47$, ou quando os parâmetros $\phi_1 = 1,6$ e $\phi_2 = -0,8$; com $m = 5$ e também para $m = 9$, onde o desempenho desta função alcançou 47,22%, de eficiência quanto à identificação dos processos simulados (Tabela, 5.3). Portanto, caso a eficiência no desempenho da função na identificação do processo igual a 47,22% seja considerada satisfatória, visto que, a diferença de 8,34% entre o primeiro e o segundo melhor desempenho desta função não é tão elevada se for levado em consideração que, para se obter 55,56% de desempenho da função se faz necessário uma quantidade de termos linguístico igual a quarenta e sete, em contrapartida ao desempenho de 47,22% que pode ser obtido com apenas cinco ou nove termos linguístico, a opção com as quantidades menores de conjuntos fuzzy deve ser escolhida durante a fuzzificação da série histórica, desde que a função $\overline{W}_{i,k,j}^{RA}(h)$ seja utilizada na rotina de identificação do modelo.

Tabela 5.3 - Frequência relativa (%) dos eventos $P | W_{i,k,j}^{DF_Z} | > 2T^{-\frac{1}{2}}$, $P | W_{i,k,j}^{CA}(h; \alpha; b) | > 2T^{-\frac{1}{2}}$ e $P | \overline{W}_{i,k,j}^{RA}(h) | > 2T^{-\frac{1}{2}}$ em 1.000 Simulações de processos autoregressivo de ordem 2, com o comprimento da série $T = 96$ ($2T^{-\frac{1}{2}} = 0,2041$) para valores de $m = 5; 7; 9; 23; 47$ (termos linguísticos ou conjuntos fuzzy) e quatro combinações dos parâmetros ϕ_1 e ϕ_2 .

EACF	m	$\phi_1 = 0,5$ e $\phi_2 = -0,5$	$\phi_1 = 1,6$ e $\phi_2 = -0,8$	$\phi_1 = 0,5$ e $\phi_2 = 0,0$	$\phi_1 = 0,5$ e $\phi_2 = 0,3$
$W_{i,k,j}^{DF_Z}$	5	38,89	30,56	27,78	41,67 ²
	7	36,11	33,33	38,89	41,67 ²
	9	52,78 ¹	55,56 ^{1;2;3}	30,56	19,44
	23	33,33	36,11 ²	30,56	27,78
	47	38,89 ²	38,89 ²	33,33	33,33
$W_{i,k,j}^{CA}(h; \alpha^*; b)$	5	30,56 ²	25,00	27,78	30,56 ²
	7	41,67 ²	41,67 ²	30,56	30,56
	9	27,78	30,56	33,33 ²	33,33 ²
	23	36,11	50,00 ^{2;3}	38,89	47,22 ¹
	*Nota: $\alpha = 0,5$	47	38,89 ²	36,11	30,56
$\overline{W}_{i,k,j}^{RA}(h)$	5	41,67	47,22 ²	44,44	27,78
	7	33,33	36,11	38,89 ²	38,89 ²
	9	38,89	47,22 ²	38,89	36,11
	23	36,11	36,11	36,11	38,89 ²
	47	47,22	44,44	55,56 ^{1;2;3}	30,56

Nota: 1, 2 e 3 são tais como definidos na Tabela 5.1.

Em virtude das informações contidas na Tabela (5.3), se torna possível representar graficamente os seus resultados, tal como, se observam nas Figuras (5.9-5.12), de maneira que, para cada combinação de parâmetros que foram utilizados nas simulações, se tem uma frequência relativa associada à eficiência na identificação do modelo, em um processo autoregressivo com ordem 2, a partir das cinco quantidade de termos linguísticos ($m = 5; 7; 9; 23$ e 47) simuladas neste estudo, com um comprimento de trajetória $T = 96$ (tamanho da série temporal simulada).

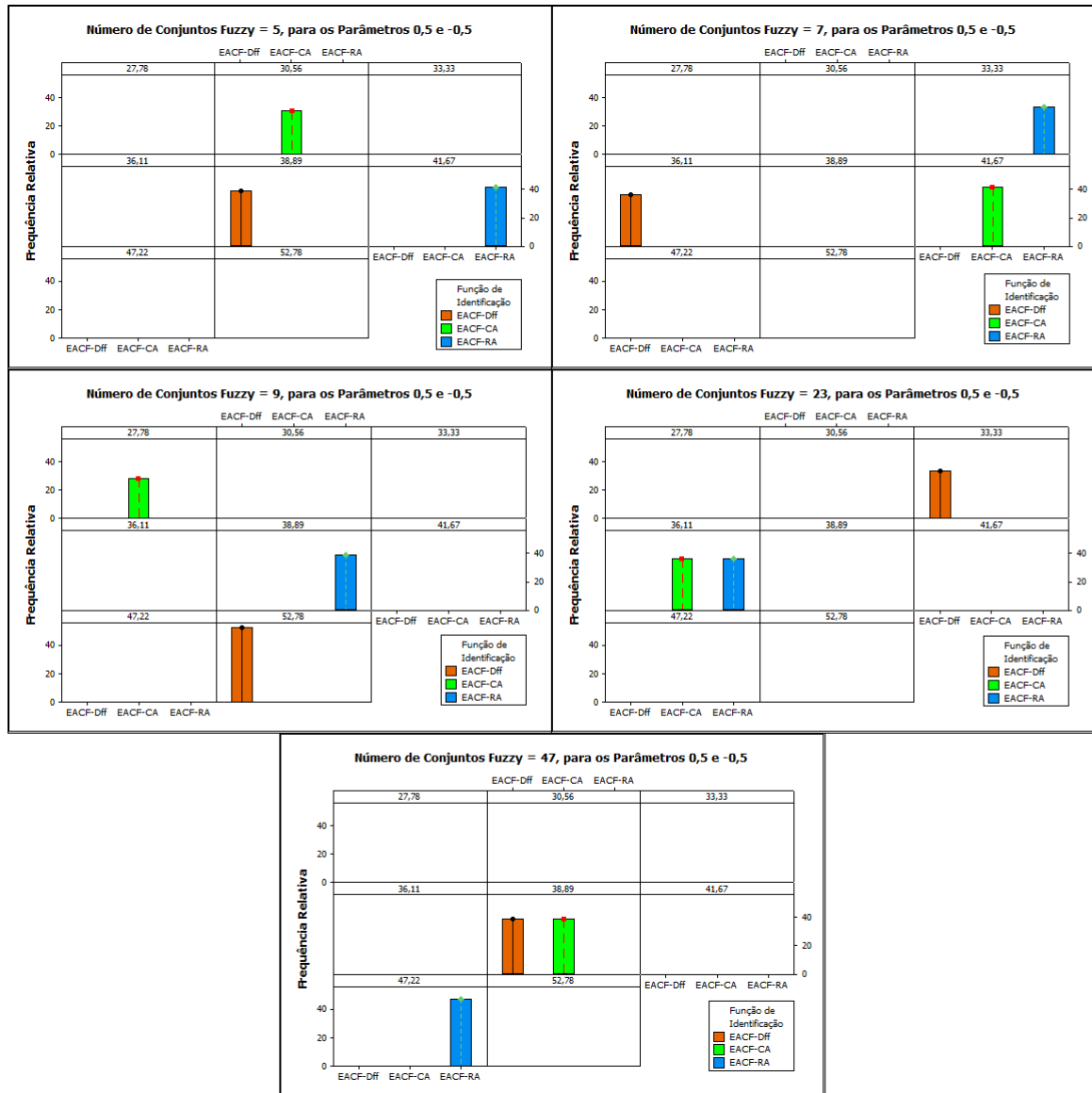
Neste sentido, se pode concluir que, caso a série temporal fuzzy seja curta, portanto, uma trajetória com $T \leq 24$, a identificação de um processo AR(2) deve ser realizada pela função $W_{i,k,j}^{DFZ}$ ou $W_{i,k,j}^{CA}(h; \alpha; b)$, com 47 e 7 termos linguísticos, respectivamente, para que a eficiência deste procedimento seja máxima. Para uma série temporal de trajetória mediana, ou seja, uma série com $24 < T \leq 36$, o indicado segundo as simulações realizadas é utilizar a função $\bar{W}_{i,k,j}^{RA}(h)$, com 7 conjuntos fuzzy para que o desempenho quanto à identificação de um processo AR(2) seja o melhor possível.

Quanto a séries longas, no caso, uma série temporal onde $T > 36$, o mais indicado é utilizar $W_{i,k,j}^{DFZ}$ ou $\bar{W}_{i,k,j}^{RA}(h)$, como função de identificação com 9 e 47 conjuntos fuzzy, respectivamente, pois desta forma o desempenho da rotina de identificação de processos AR(2) será maximizado, e o comportamento destas funções com as configurações citadas acima resultará em uma melhor eficiência, quanto à identificação do modelo temporal.

A partir dos resultados das simulações apresentados nas Tabelas 5.1, 5.2 e 5.3, se tornou possível traçar um perfil para a rotina de identificação de processos ARMA($p; q$), mediante a função de autocorrelação estendida fuzzy, que está sendo proposta neste trabalho, tal como, se apresentam nas Figuras (5.13-5.15). Neste contexto, apresentam-se três fluxogramas com o passo a passo que deve ser seguido, de maneira que, se alcance uma identificação ótima (com maior precisão possível), o que como foi possível perceber a partir das simulações está diretamente relacionado com o comprimento de trajetória da série temporal. Ou seja, de acordo com o número de observações que a série temporal apresenta é incorporado pela função de identificação mediante os valores (*defuzzificados*, alfa-nível e média aleatória) da série, uma quantidade específica de conjuntos fuzzy a serem construídos. A quantidade de conjuntos fuzzy construídos no processo de *fuzzificação* da série, como se constatou nas simulações também é relevante para o bom desempenho da função

de identificação dos modelos de série temporal fuzzy, sobretudo, para contribuir com a acurácia das previsões realizadas pelo modelo sugerido na rotina de identificação. Então, há dois fatores principais associados ao desempenho do procedimento de identificação: o tamanho de trajetória da série temporal e o número de conjuntos fuzzy construídos.

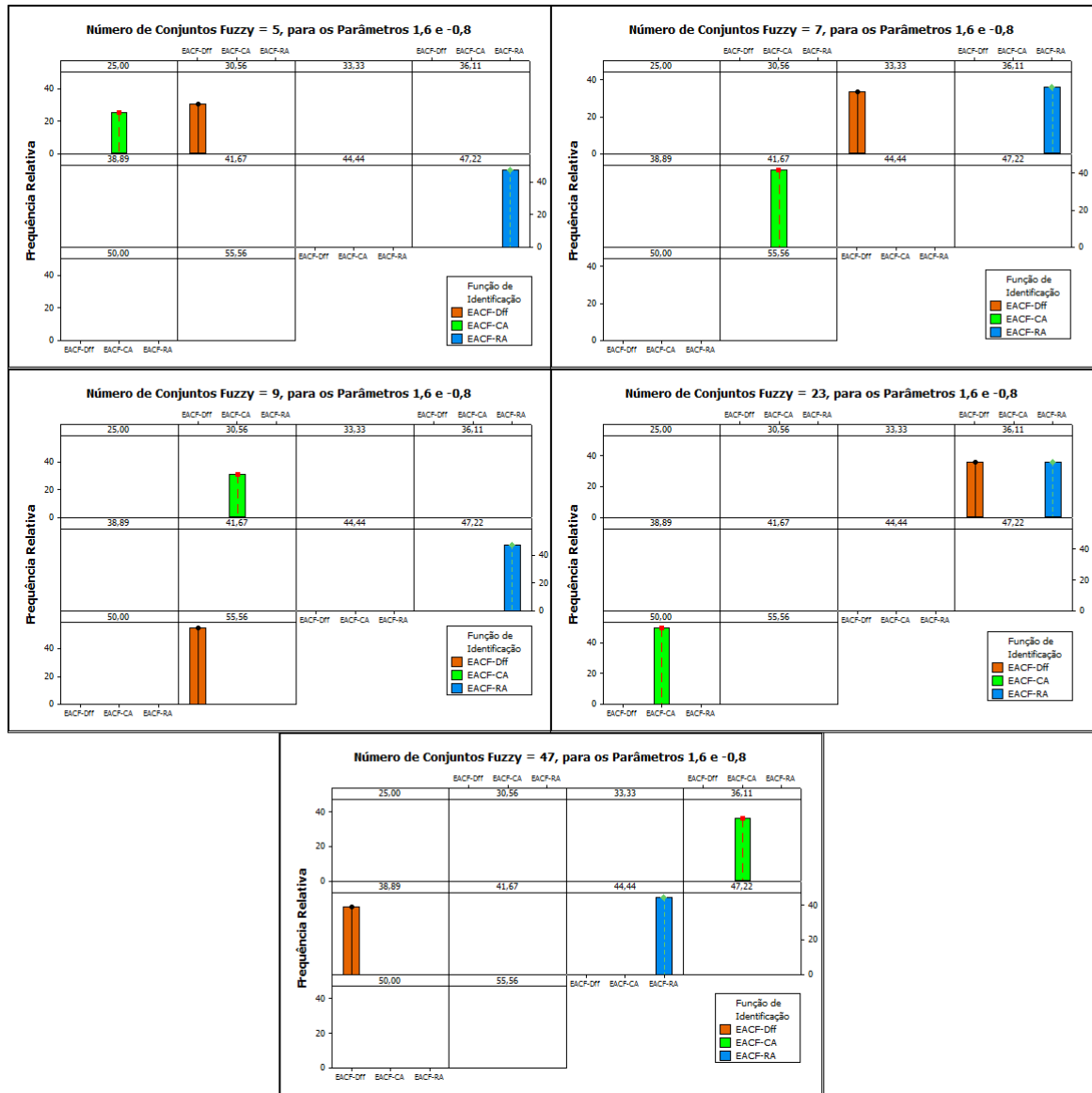
Figura 5.9 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = -0,5$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$.



Mediante a Figura (5.9) é possível verificar que a função média aleatória, retorna o melhor desempenho quanto a determinação da ordem de processos AR(2), quando o número de conjuntos fuzzy é $m = 5$ ou $m = 23$ ou $m = 47$, com frequências de 41,67%, 36,11% (igualmente à função com α -nível) e 47,22%, respectivamente. Com sete conjuntos fuzzy,

a função de melhor desempenho na determinação do processo AR(2), foi o método com α -nível retornando uma eficiência de 41,67%. Finalmente, para $m = 9$, a função com dados defuzzificados apresentou um desempenho de 52,78%, que foi superior as demais funções.

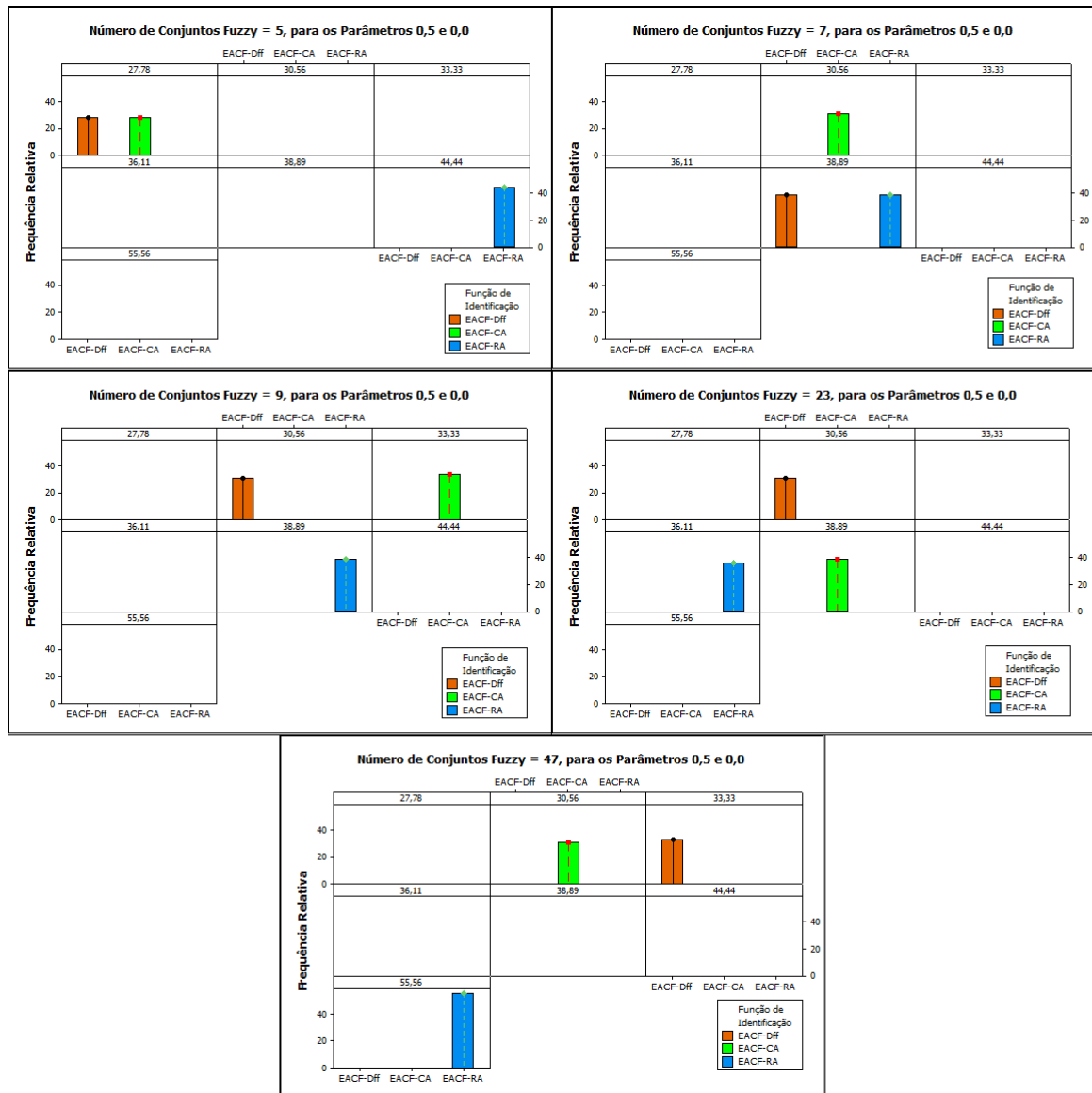
Figura 5.10 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 1,6$ e $\phi_2 = -0,8$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$.



De acordo com a Figura (5.10), em duas situações ($m = 5$ ou $m = 47$), a função média aleatória apresentou resultados superiores as demais funções com 47,22% e 44,44%, respectivamente, na frequência de acerto para determinação do processo AR(2). Com $m = 7$ ou $m = 23$, a melhor ferramenta de identificação foi a função com α -nível, pois, apresentou uma frequência de identificação igual a 41,67% e 50%, respectivamente. Por

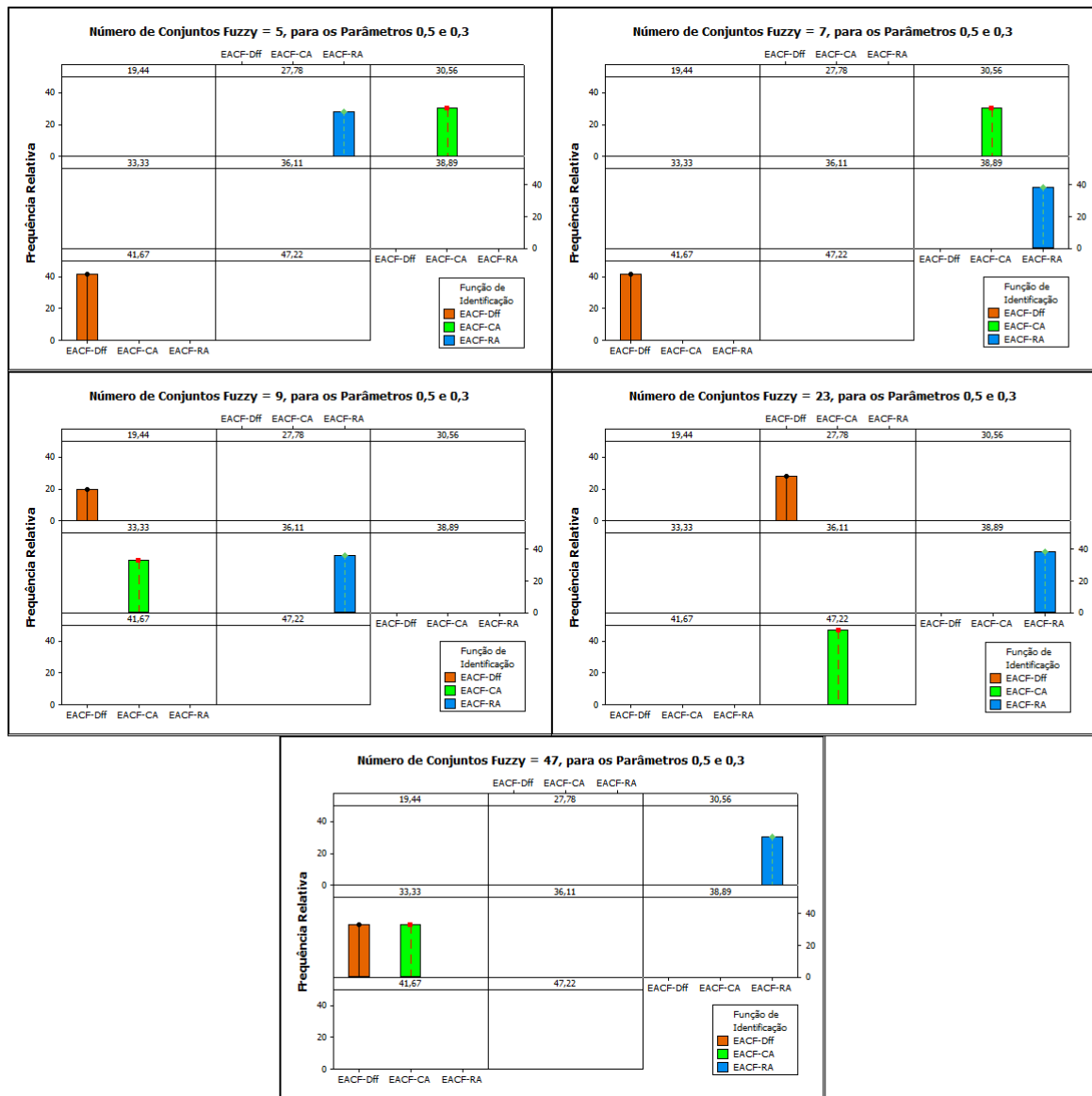
fim, quando $m = 9$, a melhor função foi a que considera os dados *defuzzificados* (55,56%).

Figura 5.11 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = 0,0$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$.



Pela Figura (5.11), observa-se que a função média aleatória foi a que mais falhou ao identificar um processo AR(1), como se fosse um AR(2), com uma frequência de 44,44% ($m = 5$), 38,89% ($m = 7$), 38,89% ($m = 9$) e 50,56% ($m = 47$). Com 23 conjuntos fuzzy, a ferramenta que mais falhou no processo de identificação do AR(2), foi a função com α -nível. Vale destacar, que a função com os dados *defuzzificados* apresentou os menores erros de identificação para os diferentes tamanhos de conjuntos fuzzy simulados.

Figura 5.12 - Representações gráficas dos resultados das simulações dos processos AR(2) com $\phi_1 = 0,5$ e $\phi_2 = 0,3$, para valores de $m = 5; 7; 9; 23; 47$ e comprimento de trajetória $T = 96$.



Analisando a Figura (5.12), é possível identificar o melhor desempenho entre as funções propostas neste trabalho, quando a ferramenta utilizada é a com dados *defuzzificados* para um número de conjuntos *fuzzy* representado por $m = 5$, $m = 7$ ou $m = 47$ (juntamente à função com α -nível), com frequências de determinação corretas de processos AR(2) iguais a: 41,67%, 41,67% e 33,33%, respectivamente. Com nove conjuntos *fuzzy* o melhor desempenho é da função média aleatória (36,11%) e para $m = 23$, a função com α -nível se destaca 47,67% de eficiência superando o desempenho das demais funções utilizadas.

5.2 - Conclusão do Capítulo

Neste capítulo se buscou testar a eficiência das ferramentas de identificação em modelos de série temporal *fuzzy*, mediante a realização de simulações de diversos processos autoregressivos, para tal, se utilizou os métodos de auxílio da identificação que foram propostos no Capítulo 5. Neste contexto, quatro combinações de parâmetros para os modelos autoregressivos de segunda ordem, assim como, três tamanhos de trajetória para a série temporal e cinco quantidades de conjuntos *fuzzy*, foram simulados e medidas a capacidade de identificação, em relação a cada uma das configurações que foram processadas. A partir da frequência relativa retornada em cada rotina estabelecida e submetida as três ferramentas propostas neste estudo, a acurácia destas ferramentas foram avaliadas, sejam elas oriundas de dados *defuzzificados*, dados com *alfa-nível* em três estágios de ativação, e ainda, para função de autocorrelação amostral média aleatória.

A função $W_{i,k,j}^{DF_Z}$ apresentou uma capacidade razoável de identificação em uma série temporal *fuzzy*, com trajetórias de tamanho 24, quando o número de conjuntos *fuzzy* foi igual a 47, assim como, para trajetórias de comprimento 36. Porém, quando o tamanho da trajetória foi 96, o melhor desempenho observado se deu com 9 conjuntos *fuzzy*. Neste contexto, se pode concluir que para séries curtas ($T = 24$ e $T = 36$) o número de conjuntos *fuzzy* a ser utilizado no processo deve ser elevado, algo que não é necessário caso a série seja de longa duração, isto é, para uma série temporal *fuzzy* com $T > 36$, desde que, a função de identificação utilizada seja $W_{i,k,j}^{DF_Z}$. Quanto a função $W_{i,k,j}^{CA}(h; \alpha; b)$ em séries curtas, portanto para $T = 24$ e $T = 36$, o número de conjuntos *fuzzy* necessários para se obter uma melhor eficiência na identificação dos processos, foi alcançado com 7 e 5 conjuntos, respectivamente. Porém, para uma trajetória $T = 96$, o melhor desempenho se obteve com 23 conjuntos *fuzzy*. Então, se pode concluir que para séries curtas ou longas, o número de conjuntos *fuzzy* (m) varia em torno de $5 \leq m \leq 23$, caso, a função de autocorrelação estendida *fuzzy* escolhida para auxiliar na identificação seja $W_{i,k,j}^{CA}(h; \alpha; b)$. Em relação à função $\bar{W}_{i,k,j}^{RA}(h)$ o melhor desempenho, isto é, a maior capacidade de identificação dos processos foi observada quando $T = 24$ e $T = 96$, para $m = 23$ e $m = 47$, respectivamente. Logo, uma quantidade considerável de conjuntos *fuzzy* deve ser construída, para que esta função obtenha um bom desempenho quanto à identificação dos processos. Porém, caso $T = 36$, faz-se a construção de 7 conjuntos na rotina de *fuzzificação*, ou seja, em trajetórias

muito curtas ($T = 24$) ou medianas ($T = 96$), o número de conjuntos *fuzzy* é elevado, porém, se o comprimento da trajetória é mediano ($T = 36$), com 7 conjuntos *fuzzy* é suficiente para se obter êxito na rotina de identificação, mediante a função $\overline{W}_{t,k,j}^{RA}(h)$.

A partir dos resultados obtidos com as simulações, uma rotina de identificação sobre modelos autoregressivos *fuzzy*, se tornou possível de estabelecer considerando inicialmente, o tamanho de trajetória da série analisada, devido sua influencia direta no número de conjuntos *fuzzy* a serem construídos. Assim sendo, de acordo com o comprimento da série temporal, uma das três ferramentas propostas neste estudo será a mais indicada, como a melhor opção para obter-se a eficiência máxima no processo de identificação, dos modelos de previsão via série temporal *fuzzy*. Neste momento, a presença de um especialista contribuirá bastante para a tomada de decisão, isto é, a escolha da função mais conveniente para o cenário em questão, como por exemplo, caso a trajetória $T = 36$, se faz necessário a construção de apenas 5 conjuntos *fuzzy*, ao utilizar a função $W_{i,k,j}^{CA}(h; \alpha; b)$. No entanto, para este mesmo $T = 36$, se a opção for pela função $\overline{W}_{i,k,j}^{RA}(h)$ ou $W_{i,k,j}^{DFZ}$, se faz necessário a construção de 7 ou 47 conjuntos *fuzzy*, respectivamente. Consequentemente, a execução da rotina de identificação deve considerar a sensibilidade de um especialista, tal que, a eficiência do processo de identificação *versus* o tempo de duração para implementação da *fuzzificação* dos dados, necessariamente devam ser avaliados e otimizados, para que o esforço computacional exigido seja minimizado, durante a rotina de identificação que foi proposta. É importante ressaltar que a escolha correta da ferramenta de identificação, implica em utilizar um modelo de previsão com o menor erro possível, isto é, um modelo que capture plenamente toda a dependência necessária nos dados observados. Logo, a identificação equivocada do modelo de série temporal, seja em um universo de discurso clássico ou no ambiente *fuzzy*, inevitavelmente irá influenciar de forma direta na capacidade preditiva do modelo de série temporal estimado.

No próximo capítulo, todas as ferramentas de identificação propostas neste trabalho, serão aplicadas em séries reais disponíveis na literatura. Assim, acredita-se ser possível validar os resultados que foram obtidos nas simulações deste estudo. Logo, a partir de previsões e projeções realizadas pelos modelos de série temporal *fuzzy*, resultantes das rotinas de identificação, a eficiência e funcionalidade das ferramentas desenvolvidas neste trabalho serão novamente avaliadas e possivelmente validadas.

Figura 5.13 - Fluxograma para a rotina de identificação em modelos $ARMA(p; q)$, proposta neste trabalho via simulações realizadas para os valores *defuzzificados* da série histórica.

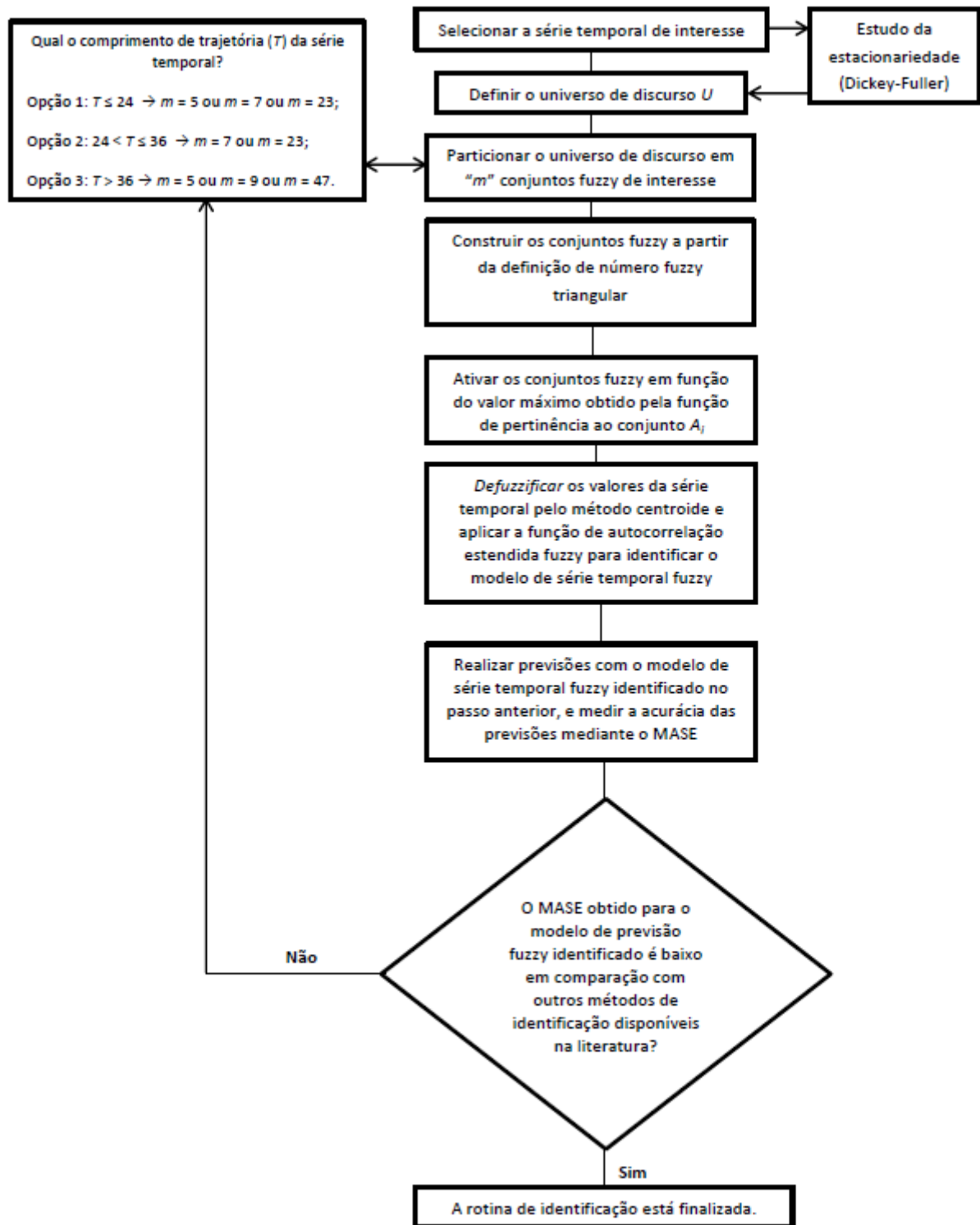


Figura 5.14 - Fluxograma para a rotina de identificação em modelos $ARMA(p; q)$, proposta neste trabalho via simulações realizadas para os valores da série histórica com *alfa-nível*.

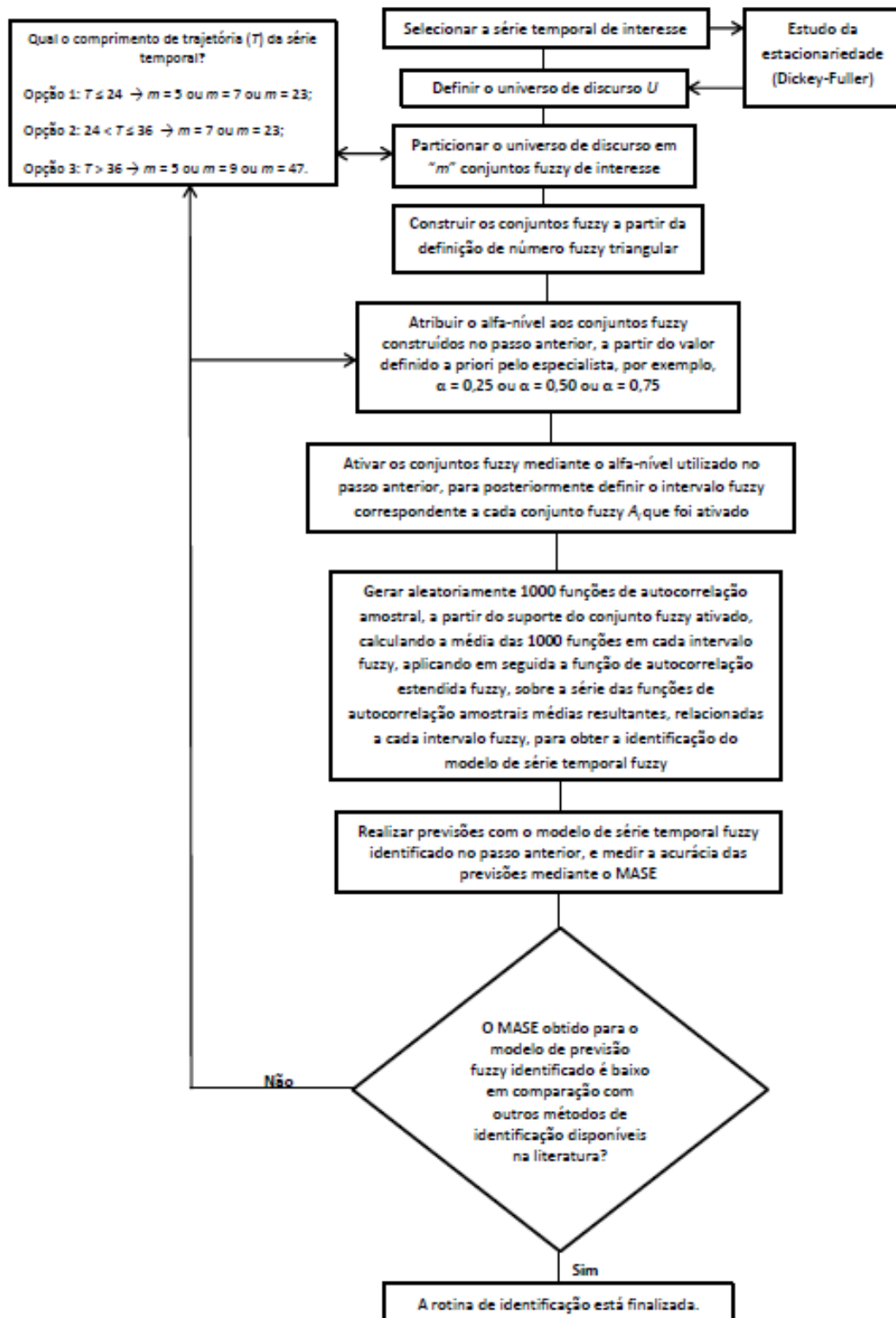
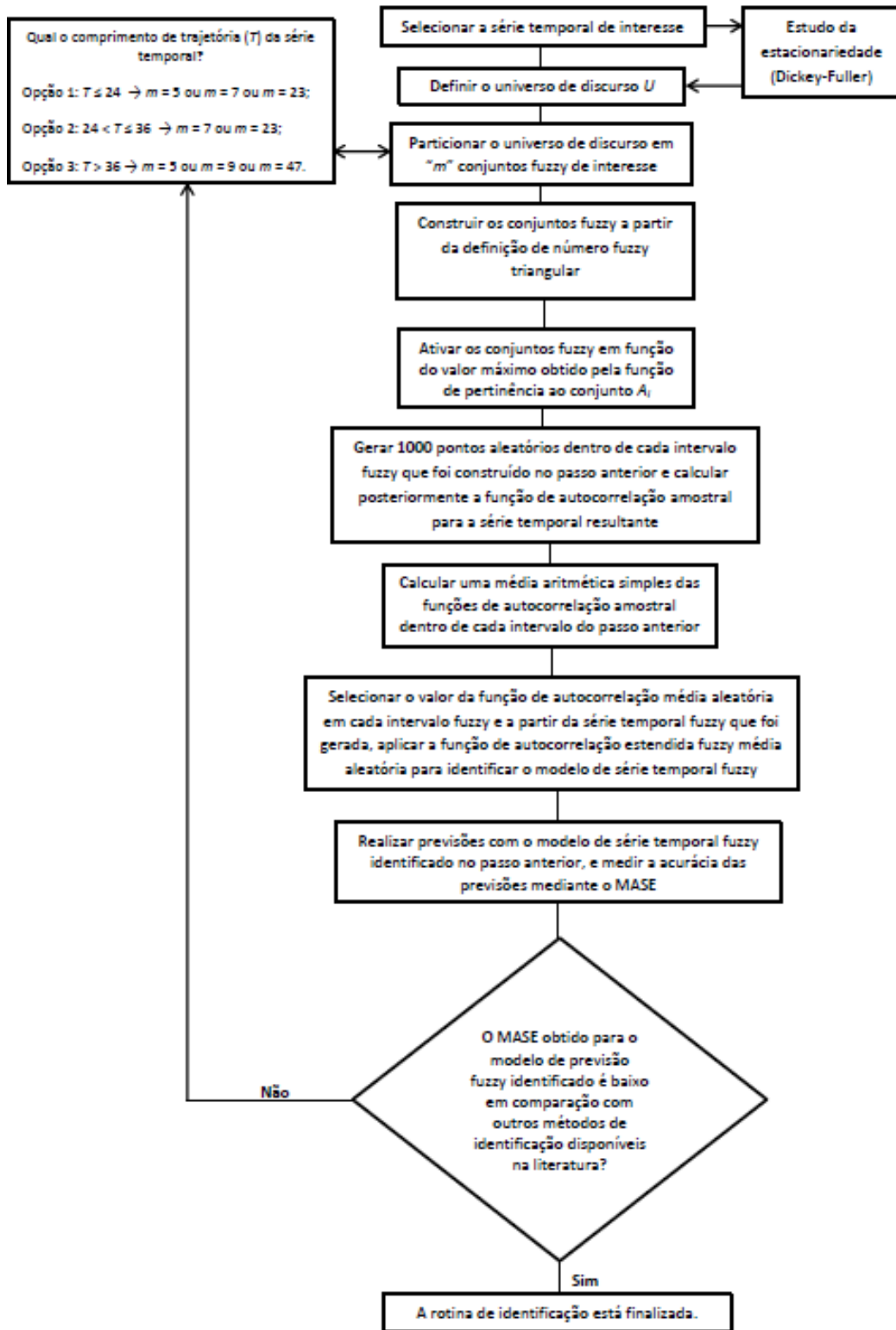


Figura 5.15 - Fluxograma para a rotina de identificação em modelos $ARMA(p; q)$, proposta neste trabalho via simulações realizadas para os valores da série fac média aleatória.



Capítulo 6

Utilização Sobre Dados Reais dos Métodos Propostos Para a Identificação de Modelos de Série Temporal Fuzzy

6.1 Estudo de Caso e Comparação dos Resultados Obtidos

Após todos os resultados obtidos com as simulações no Capítulo 5, neste capítulo serão aplicadas as ferramentas propostas no Capítulo 4 deste trabalho, em séries de dados reais disponíveis na literatura. Desta forma, se busca validar os algoritmos de identificação dos modelos de série temporal *fuzzy*, que foram desenvolvidos neste estudo. Além disso, a acurácia dos modelos de previsão *fuzzy* será avaliada, a partir da geração de previsões, mediante um modelo sugerido e ajustado com o auxílio das metodologias de identificação, que são apresentadas neste trabalho. A precisão dos modelos de previsão *fuzzy* será avaliada, com o auxílio da escala de erro médio absoluto (MASE), a qual corresponde a uma medida de variabilidade do modelo de previsão ajustado. Também serão realizadas comparações dos resultados das previsões, pelos métodos propostos neste trabalho, em relação às previsões obtidas a partir de outros métodos de identificação já consolidados, que se encontram disponíveis na literatura de previsão por série temporal. Neste contexto, será estabelecida neste trabalho uma avaliação, quanto ao desempenho de todos os procedimentos propostos, em relação aos métodos de identificação mediante modelos de série temporal clássico ou *fuzzy* conhecidos, e utilizados atualmente por parte da comunidade científica.

6.1.1 - Estudo de caso 1: Previsão fuzzy para o preço da energia elétrica em New South Wales, na Austrália.

Com o intuito de comprovar a eficiência e aplicabilidade das ferramentas de identificação propostas neste trabalho, uma base de dados disponível em Dong et al. (2011) foi utilizada neste estudo de caso. Trata-se de uma série temporal, contendo os preços de energia

elétrica coletados a cada meia hora, em um sistema elétrico na Austrália. Portanto, durante um dia inteiro de uma segunda-feira escolhida aleatoriamente, a coleta de dados resultou em uma série temporal com 48 instantes de registro, do preço em dólar da energia elétrica consumida na região de New South Wales. As ferramentas de identificação que foram desenvolvidas no Capítulo 5, deste trabalho serão utilizadas, sobre esta base de dados diários do preço de energia elétrica. Ao final deste estudo experimental, os procedimentos realizados tornarão possível estabelecer um modelo de série temporal *fuzzy*, com capacidade preditiva que se acredita ser suficiente, para prever plenamente a característica numérica utilizada, durante a validação do método proposto.

Para avaliar o desempenho alcançado pelo modelo de previsão, tal que, este seja identificado pela metodologia proposta neste estudo, uma comparação será realizada entre o erro de previsão, que é gerado pelo modelo sugerido na rotina de identificação, em relação a três métodos de previsão por série temporal. A base de dados utilizada na validação do método proposto, considera o preço da energia elétrica em New South Wales na Austrália apresentada por Dong et al. (2011).

Passo 1: Coletar os dados de interesse que formarão a série temporal.

Como a série temporal dos preços diários de energia elétrica, em New South Wales na Austrália está disponível na literatura, se inicia a rotina descrita na Seção 4.6, a partir do Passo 2.

Passo 2: Construir o gráfico da série temporal dos preços diários de energia elétrica, em New South Wales na Austrália durante o período de tempo observado, como mostra a Figura 6.1.

Passo 3: Definir o universo de discurso.

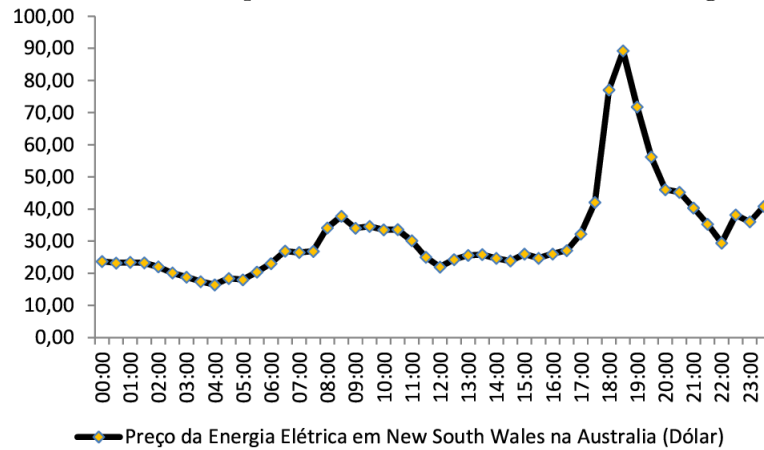
Para a série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, como $Z_{(min)} = 16,43$, $Z_{(max)} = 89,24$, $D_1 = 12,135$ e $D_2 = 12,16$. Então, $U = [4, 295; 101, 400]$.

Passo 4: Particionar o universo de discurso em “ m ” conjuntos *fuzzy*.

Seguindo a rotina de identificação de processos ARMA($p; q$), que foi estabelecida na Seção 5.1, mais exatamente pelo fluxograma (ver, Figura 5.13). Segundo esta rotina para

um tamanho de trajetória $T > 36$ ($T = 48$, neste estudo de caso), se utiliza m igual a 5, 7 ou 9. Neste contexto, o valor de $m = 7$ será o adotado, como recomenda Miller (1994). Além disto, $m = 7$ representa uma quantidade média entre os valores mínimo e máximo possíveis, além de ser um valor amplamente utilizado na literatura, o que facilita a comparação com outras propostas de identificação. Ainda, neste exemplo a função de pertinência será calculada pela definição de número *fuzzy* triangular, por se acreditar que esta função aumenta a acurácia das previsões. A amplitude será $l = 24, 27$.

Figura 6.1 - Série temporal diária referente ao preço em dólar da energia elétrica em New South Wales na Austrália, durante 48 instantes de tempo distintos no período das 00:00 AM a 23:30 PM, coletados em intervalos de tempo de meia em meia hora de uma segunda feira.



Passo 5: Construir os conjuntos *fuzzy* correspondentes a quantidade estabelecida no Passo 4 e *fuzzificar* a série histórica a partir da definição de número *fuzzy* triangular.

Seja A_j , onde $j = 1; 2; \dots; 6; 7$, o j -ésimo conjunto *fuzzy* sobre o universo de discurso U . Contudo, caso A_j seja um número *fuzzy* triangular (ver, Definição 3.5), se chega que, $u_j = [a_j; b_j; c_j]$. Assim, a pertinência de Z_i , para $i = 0; 1; \dots; n$, ($n \leq T$) em relação a A_j , pela Equação (6.1) é.

$$A_j = \begin{cases} 0, & \text{se } Z_i \leq a_j; \\ \frac{Z_i - a_j}{b_j - a_j}, & \text{se } a_j < Z_i \leq b_j; \\ \frac{c_j - Z_i}{c_j - b_j}, & \text{se } b_j < Z_i \leq c_j. \\ 0, & \text{se } Z_i > c_j. \end{cases} \quad (6.1)$$

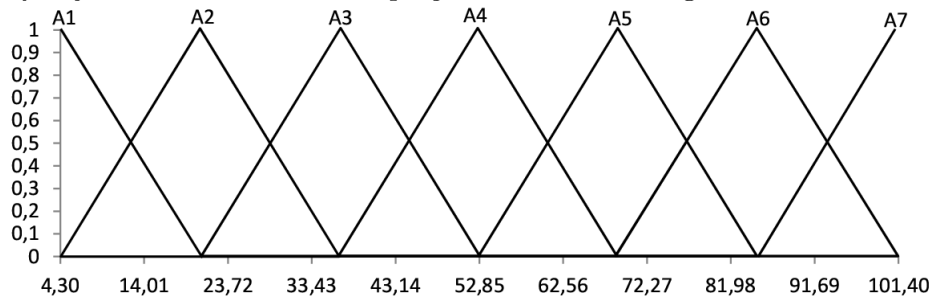
Para os sete intervalos resultantes de $m = 7$, onde devido a função de pertinência a ser

calculada, para cada preço diário de energia elétrica em relação aos conjuntos A_j , ser proveniente da definição de número *fuzzy* triangular, cada intervalo representa a base de um dos sete triângulos que serão construídos (ver, Figura 6.2). Neste sentido, os intervalos são dados por:

$$u_1 = [-11, 89; 4, 30; 20, 47], \quad u_2 = [4, 30; 20, 47; 36, 66], \quad u_3 = [20, 47; 36, 66; 52, 84], \\ u_4 = [36, 66; 52, 84; 69, 03], \quad u_5 = [52, 84; 69, 03; 85, 21], \quad u_6 = [69, 03; 85, 21; 101, 40] \text{ e} \\ u_7 = [85, 21; 101, 40; 117, 60].$$

Observação 6.1. Foi adotada a pertinência máxima para ativar os sete conjuntos *fuzzy* da Figura (6.2). O método adotado de *defuzzificação* da série temporal *fuzzy* é o centróide.

Figura 6.2 - Conjuntos *fuzzy* pela definição de número *fuzzy* triangular, para o preço em dólar da energia elétrica em New South Wales na Austrália, onde o eixo Y representa a pertinência ao conjunto *fuzzy* A_i e o eixo X indica o preço em dólar da energia elétrica.



Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para identificar o modelo de série temporal *fuzzy*, sobre os dados históricos de interesse.

(a) Para a função $W_{i,k,j}^{DFZ}$, se deve *defuzzificar* a série histórica mediante o método centróide, e em seguida aplicar a função de autocorrelação estendida *fuzzy* à série *defuzzificada*.

Seguindo o passo 6(a), da rotina descrita na Seção 5.6, se obteve os valores de $W_{i,k,j}^{DFZ}$, com suas probabilidades de significância (*p-valor*) correspondentes, que são apresentados na Tabela (6.1) e Tabela (6.2), respectivamente.

Tabela 6.1 - Resultados da função $W_{i,k,j}^{DFZ}$ para o preço da energia elétrica em New South Wales.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,055	0,101	0,044	-0,069	-0,036	-0,130
AR(1)	-0,428	0,068	0,055	-0,075	0,012	-0,014
AR(2)	-0,307	0,349	0,046	-0,065	0,000	-0,011
AR(3)	0,346	0,495	0,186	-0,063	-0,010	-0,030
AR(4)	-0,375	-0,366	-0,110	0,284	0,027	-0,030
AR(5)	-0,284	-0,148	-0,108	0,205	0,319	-0,035

Tabela 6.2 - Probabilidades de significância para os resultados da Tabela (6.1).

<i>Lag</i>	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,706	0,492	0,769	0,642	0,807	0,382
AR(1)	0,004*	0,683	0,767	0,663	0,940	0,934
AR(2)	0,040*	0,083*	0,798	0,674	1,000	0,956
AR(3)	0,022*	0,001*	0,289	0,688	0,952	0,903
AR(4)	0,014*	0,016*	0,482	0,191	0,904	0,899
AR(5)	0,066*	0,348	0,547	0,368	0,049*	0,871

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Mediante a rotina de *fuzzificação* realizada sobre os dados do preço da energia elétrica (Tabela 6.3), necessariamente terem resultado em relações *fuzzy* lógicas para pares de valores da série, também foram estabelecidos os grupos destas relações *fuzzy* lógicas, como mostra a Tabela (6.4). Logo, se torna possível identificar os grupos de relações *fuzzy* lógicas, com a transição entre os conjuntos A_j durante todo o processo de *fuzzificação*. Devido aos resultados obtidos na Tabela (6.2), a um nível de significância $\alpha = 10\%$, se pode afirmar que a ordem do modelo ARMA identificada pela função $W_{i,k,j}^{DFZ}$ é: ARMA(2; 1), ARMA(3; 1) e ARMA(3; 2). Neste sentido, a série do preço de energia elétrica será ajustada por um modelo ARIMA (3; 1; 2), com o auxílio de um *software* estatístico para testar a significância dos parâmetros p e q , identificados pela função de autocorrelação estendida *fuzzy*. Desta forma, se chegará ao modelo de série temporal *fuzzy* com maior capacidade preditiva, em relação ao preço (dólar) da energia elétrica em New South Wales na Austrália, pois desta forma se acredita que, além do modelo ser significativo (estatisticamente diferente de zero) a um nível $\alpha = 10\%$, os erros de previsão gerados por este modelo implicarão em um MASE muito pequeno, o que aumenta a acurácia do modelo de previsão *fuzzy* quanto ao preço da energia elétrica estudado.

A partir deste momento, a estacionariedade da série em estudo deve ser comprovada, para atender aos pressupostos estabelecidos a uma série temporal. Pelos resultados da Tabela (6.5), o teste da raiz unitária de *Dickey-Fuller* aumentado (ver, Hylleberg, 1995), quando a série temporal dos preços de energia elétrica constam sem diferenciação ($d = 0$), indicam não-estacionariedade da série, ou seja, o valor da estatística de teste *tau* do modelo utilizado foi -1,66346 para um *p-valor* de 0,44299. Desta forma, não se rejeitou a hipótese nula de raiz unitária, o que resultou em uma série temporal não estacionária. No

Tabela 6.3 - Dados históricos referentes ao preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de *fuzzificação* realizado a partir de um número *fuzzy* triangular, relação *fuzzy* lógica e os preços em dólar da energia elétrica *defuzzificados*.

Horário observado	Preço em dólar	Pertinência máxima	Intervalo correspondente	Série Temporal <i>fuzzy</i>	Relação <i>fuzzy</i> lógica	Preços <i>defuzzificados</i>
00:00	23,72	0,799	[4,30; 20,47; 36,66]	A_2	-	20,417
00:30	23,24	0,829	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,486
01:00	23,42	0,818	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	19,870
01:30	23,28	0,826	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,200
02:00	22,03	0,904	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,333
02:30	20,11	0,978	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	21,066
03:00	18,86	0,900	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,162
03:30	17,42	0,811	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,607
04:00	16,43	0,750	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,484
04:30	18,41	0,873	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,526
05:00	18,00	0,847	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,082
05:30	20,36	0,993	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	21,079
06:00	23,05	0,841	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,478
06:30	26,93	0,601	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,616
07:00	26,48	0,629	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,346
07:30	26,88	0,604	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,633
08:00	34,16	0,845	[20,49; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	36,333
08:30	37,76	0,932	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	37,015
09:00	34,07	0,840	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,664
09:30	34,61	0,873	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	37,040
10:00	33,58	0,810	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,806
10:30	33,60	0,811	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	37,256
11:00	30,13	0,596	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,517
11:30	25,07	0,716	[4,30; 20,47; 36,66]	A_2	$A_3 \rightarrow A_2$	20,650
12:00	21,98	0,907	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,129
12:30	24,26	0,766	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,626
13:00	25,58	0,684	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,121
13:30	25,83	0,669	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,508
14:00	24,66	0,741	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,330
14:30	23,83	0,792	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,868
15:00	25,99	0,659	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,335
15:30	24,71	0,738	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,317
16:00	26,05	0,655	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,142
16:30	27,10	0,590	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	20,602
17:00	32,23	0,726	[20,49; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	36,443
17:30	42,06	0,666	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,353
18:00	77,08	0,502	[52,85; 69,03; 85,21]	A_5	$A_3 \rightarrow A_5$	69,091
18:30	89,24	0,751	[69,03; 85,21; 101,40]	A_6	$A_5 \rightarrow A_6$	85,282
19:00	71,77	0,831	[52,85; 69,03; 85,21]	A_5	$A_6 \rightarrow A_5$	69,250
19:30	56,21	0,792	[36,66; 52,84; 69,03]	A_4	$A_5 \rightarrow A_4$	52,957
20:00	46,05	0,580	[36,66; 52,84; 69,03]	A_4	$A_4 \rightarrow A_4$	53,281
20:30	45,25	0,531	[36,66; 52,84; 69,03]	A_4	$A_4 \rightarrow A_4$	53,052
21:00	40,35	0,772	[20,49; 36,66; 52,84]	A_3	$A_4 \rightarrow A_3$	36,174
21:30	35,33	0,918	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,632
22:00	29,42	0,552	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	37,292
22:30	38,17	0,907	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,610
23:00	36,02	0,960	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,586
23:30	40,85	0,741	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	36,798

entanto, quando se realiza uma diferença na série temporal ($d = 1$), o comportamento da série de energia elétrica passa a apresentar estacionariedade, face ao valor da estatística

Tabela 6.4 - Grupos de relações *fuzzy* lógicas ativadas para os dados da Tabela (6.3).

Grupo ativado	Relações <i>fuzzy</i> lógicas
Grupo 1	$A_2 \rightarrow A_2, A_2 \rightarrow A_3$
Grupo 2	$A_3 \rightarrow A_2, A_3 \rightarrow A_3, A_3 \rightarrow A_5$
Grupo 3	$A_4 \rightarrow A_3, A_4 \rightarrow A_4$
Grupo 4	$A_5 \rightarrow A_4, A_5 \rightarrow A_6$
Grupo 5	$A_6 \rightarrow A_5$

de teste *tau* ser igual a -3,42882, para uma probabilidade de significância de 0,014712 (ver, Tabela 6.6). Logo, mediante a rejeição da hipótese nula de raiz unitária, a um nível de significância de 10%, se pode garantir que a metodologia de série temporal pode ser plenamente utilizada, sem qualquer prejuízo ao modelo que será identificado.

Tabela 6.5 - Teste da raiz unitária de *Dickey-Fuller* aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	<i>Tau</i>	<i>p-valor</i>
1	OLS	0,000016	0,082431	46	-1,66346	0,44299
2	COV	0,000016	0,002438	46	-1,66346	0,44299
3	COV	0,000016	-0,002529	46	-1,66346	0,44299
4	COV	0,000016	0,001310	46	-1,66346	0,44299

Tabela 6.6 - Teste da raiz unitária de *Dickey-Fuller* aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	<i>Tau</i>	<i>p-valor</i>
1	OLS	0,000017	0,0002709	45	-3,42882	0,014712*
2	COV	0,000017	0,0000004	45	-3,42882	0,014712*
3	COV	0,000017	-0,0000109	45	-3,42882	0,014712*
4	COV	0,000017	0,0000057	45	-3,42882	0,014712*

Nota: * Valor significativo para $\alpha = 10\%$.

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6 (a);

Pelos resultados da Tabela (6.7), se pode concluir que o modelo ajustado é significativo, visto que, os testes para os coeficientes do modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₃, a partir da estatística de teste *t-Student* são significativos. Uma exceção é o coeficiente de média móvel de ordem 1, o qual possui probabilidade de significância de 0,905, ou seja, como $p - valor = 0,905 > \alpha = 0,10$, então, este coeficiente não é significativo, logo, MA(1) é estatisticamente igual a zero a um nível de significância de 10%. Neste sentido, o termo MA(1) não deve ser considerado no modelo, o que exclui a possibilidade de ajuste pelos modelos identificados com a função $W_{t,k,j}^{DFZ}$, no caso, um ARMA(2; 1) e ARMA(3; 1), como sugeridos nas Tabelas (6.1 e 6.2). Porém, a outra opção de identificação para o

modelo de previsão *fuzzy*, no caso com o termo MA(2), não pode ser desconsiderada, pois este componente é estatisticamente diferente de zero para $\alpha = 10\%$, tal como, a constante (termo independente) do modelo ajustado. Isto implica em manter este e os demais termos no modelo de previsão ajustado à série temporal *fuzzy* do preço da eletricidade.

Em relação ao teste Q *Ljung-Box* na Tabela (6.8), adotando um nível de significância igual a 10%, se deve rejeitar a hipótese nula quando o componente $s = 12$ (*lag* igual a 12), pois neste caso, $p - \text{valor} = 0,061 < \alpha = 0,10$. Desta forma, há uma relação de dependência entre os erros da série temporal *fuzzy* observada, quando o período sazonal da série é igual a 12, que neste caso, corresponde ao 12^o (décimo segundo) instante de tempo coletado a cada 30 minutos, desde o início da série onde os preços de energia elétrica passaram a ser observados continuamente. Este fato possibilita a utilização do modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₃, como ferramenta de previsão do preço em dólar da energia elétrica em New South Wales na Austrália, dentro do período estudado, a um nível de significância de 10%, a partir do teste estatístico Q realizado.

Tabela 6.7 - Teste de significância dos parâmetros do modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₃.

Modelo	Coefficiente	Erro padrão	Estatística t	p -valor
AR(1)	0,270	0,144	1,880	0,067
AR(2)	0,740	0,103	7,160	0,000
AR(3)	-0,527	0,138	- 3,830	0,000
SAR(13)	-0,990	0,049	-20,280	0,000
MA(1)	-0,015	0,122	-0,120	0,905**
MA(2)	0,958	0,148	6,470	0,000
SMA(13)	-0,709	0,192	-3,690	0,001
Constante	0,348	0,133	2,620	0,012

Nota: ** Valor não significativo (estatisticamente igual a zero) para $\alpha = 10\%$.

Tabela 6.8 - Teste Q (*Ljung-Box*) para o modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₂.

Fonte de variação	<i>Lag</i>	Qui-quadrado	Grau de liberdade	p -valor
Teste Q " <i>Ljung-Box</i> "	12,000	9,000	4,000	0,061*
⋮	24,000	20,700	16,000	0,192
⋮	36,000	30,800	28,000	0,327
⋮	48,000	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Como consequência do teste Q *Ljung-Box*, com os resultados apresentados pela Tabela (6.8), para o caso do componente sazonal ser aumentado de $s = 12$, para $s = 13$ a 24 ou $s = 25$ a 36. A condição necessária de independência entre os resíduos dos preços

de energia elétrica, então passa a ser atendida, pois $p - valor = 0,192 > \alpha = 0,10$ e $p - valor = 0,327 > \alpha = 0,10$, respectivamente. Logo, isto implica em não rejeitar a hipótese H_0 (os resíduos são independentes entre si), o que evidencia uma ausência de associação entre os erros, algo que também é indicado pelo teste Q *Ljung-Box* quando $s = 37$ a 48 . Portanto, o teste não conseguiu calcular o valor absoluto da estatística, no período sazonal com instante de tempo para observação do preço de energia elétrica igual a $T = 48$, então ratificando o ajuste e utilização do modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₃.

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

A partir do momento que o modelo de previsão *fuzzy* têm seus parâmetros estimados, se deve realizar previsões mediante este modelo, para que desta forma a sua capacidade preditiva possa ser avaliada e este modelo seja validado. Realiza-se a comparação com os três modelos de previsão propostos por Dong et al. (2011), na Tabela (6.9).

Analisando os resultados das previsões obtidas pelo método proposto neste estudo (Tabela 6.9), é possível perceber uma razoável acurácia do modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₃, em relação ao método SARIMA (Dong et al., 2011). Pois, o modelo proposto gerou um erro de previsão (MASE) igual a 1,48, contra o erro do método SARIMA (Dong et al., 2011) que retornou valor de 1,76. No entanto, os outros dois métodos propostos por Dong et al. (2011), vistos na Tabela (6.9), se mostraram com melhor desempenho para prever o preço da energia elétrica em New South Wales na Austrália. Isto é, a capacidade preditiva dos métodos SA-ARIMA e EMD-SA-ARIMA são superiores ao método FSARIMA(3; 1; 2)(1; 0; 1)₁₃, o qual foi ajustado a partir dos valores *defuzzificados*, visto que, estes dois métodos obtiveram MASE igual a 0,99 e 1,02, respectivamente. Isso indica maior capacidade preditiva dos modelos SA-ARIMA e EMD- SA-ARIMA, em relação ao método FSARIMA(3; 1; 2)(1; 0; 1)₁₃, para o preço de energia elétrica.

Objetivando reafirmar o bom ajuste dos dados *defuzzificados* ao modelo de previsão *fuzzy* que foi ajustado, será utilizado o conceito de intervalo de confiança *fuzzy* (Wu, 2009), para 6 projeções referentes a instantes de tempo subsequentes ao período, que foi considerado para ajustar o modelo de previsão *fuzzy*, mediante as Equações (6.2) e (6.3). Assim, mediante os preços *defuzzificados* da energia elétrica e adotando um grau de confiança igual a 99% para a construção dos intervalos, logo, $Z_\alpha = 2,58$, como são 7 conjuntos

fuzzy, logo $m = 7$, $\bar{x}_h^L = 40,70$ e $\bar{x}_h^U = 64,97$, e ainda, $s_{hL} = s_{hU} = 24,27$. Então, os intervalos de confiança *fuzzy* serão dados pelas Equações (6.2) e (6.3).

Tabela 6.9 - Comparação dos resultados de previsão pelo método proposto (preços *defuzzificados*) em relação aos três métodos propostos em Dong et al. (2011).

Instante observado em horas	Preço da energia em dólar	Método SA-ARIMA (Dong et al. 2011)	Método EMD-SA-ARIMA (Dong et al. 2011)	Método SARIMA (Dong et al. 2011)	Método proposto: Treinamento	Método proposto: Validação
00:00	23,72	24,30	24,35	22,99	-	-
00:30	23,24	23,68	23,95	21,69	24,76	-
01:00	23,42	22,55	22,85	22,39	22,65	-
01:30	23,28	23,04	23,32	20,62	32,96	-
02:00	22,03	21,55	21,75	16,17	20,57	-
02:30	20,11	17,70	17,81	13,85	12,19	-
03:00	18,86	15,71	15,75	12,30	16,30	-
03:30	17,42	14,36	14,35	11,78	20,11	-
04:00	16,43	13,91	13,84	11,48	18,63	-
04:30	18,41	13,90	13,78	12,90	13,86	-
05:00	18,00	14,99	14,80	16,04	14,73	-
05:30	20,36	17,63	17,34	17,69	18,81	-
06:00	23,05	19,33	18,94	26,70	36,23	-
06:30	26,93	27,11	26,47	30,95	32,24	-
07:00	26,48	30,23	29,41	30,29	21,62	-
07:30	26,88	29,74	28,82	38,95	30,60	-
08:00	34,16	37,73	36,43	42,32	30,86	-
08:30	37,76	41,11	39,54	40,29	33,19	-
09:00	34,07	39,24	37,61	40,25	41,61	-
09:30	34,61	39,36	37,58	39,26	29,25	-
10:00	33,58	38,67	36,78	39,08	22,22	-
10:30	33,60	38,79	36,76	37,10	39,31	-
11:00	30,13	36,91	34,84	34,04	27,33	-
11:30	25,07	33,97	31,96	34,08	36,13	-
12:00	21,98	33,61	31,50	33,17	8,64	-
12:30	24,26	33,11	30,92	32,27	23,01	-
13:00	25,58	32,53	30,27	31,12	33,92	-
13:30	25,83	31,70	29,39	29,82	33,23	-
14:00	24,66	30,64	28,30	28,76	21,71	-
14:30	23,83	29,50	27,15	28,46	26,93	-
15:00	25,99	29,55	27,10	28,11	31,79	-
15:30	24,71	29,10	26,58	29,17	30,52	-
16:00	26,05	30,35	27,63	32,13	25,94	-
16:30	27,10	33,27	30,18	36,93	28,80	-
17:00	32,23	38,24	34,56	60,08	24,50	-
17:30	42,06	58,05	52,28	84,17	31,46	-
18:00	77,08	84,13	75,51	76,65	57,59	-
18:30	89,24	78,83	70,50	65,37	80,95	-
19:00	71,77	67,17	59,85	51,50	82,18	-
19:30	56,21	53,27	47,30	46,89	62,63	-
20:00	46,05	48,00	42,47	43,82	36,44	-
20:30	45,25	45,57	40,18	37,35	40,74	-
21:00	40,35	39,71	34,88	34,61	-	34,01
21:30	35,33	36,66	32,10	30,89	-	46,13
22:00	29,42	32,50	28,36	35,12	-	38,74
22:30	38,17	37,76	32,83	30,39	-	30,63
23:00	36,02	33,08	28,66	32,97	-	46,26
23:30	40,85	36,51	31,52	25,75	-	43,45
MASE		1,02	0,99	1,76	-	1,48

$$\begin{aligned} \text{I.C}(\tilde{\mu}_h^L; 99\%) &= \left[40,70 - 2,58 \times \left(\frac{24,27}{\sqrt{7}} \right); 40,70 + 2,58 \times \left(\frac{24,27}{\sqrt{7}} \right) \right] \\ &\equiv [17,03; 64,37] \end{aligned} \quad (6.2)$$

e

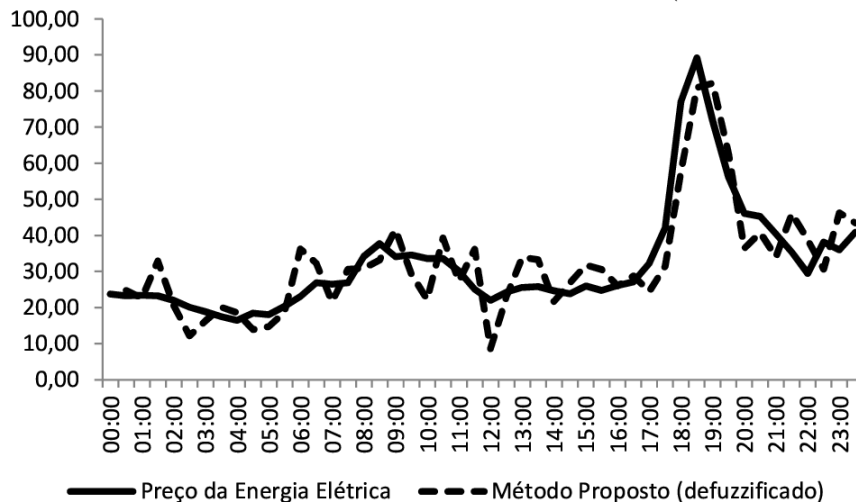
$$\begin{aligned} \text{I.C}(\tilde{\mu}_h^U; 99\%) &= \left[64,97 - 2,58 \times \left(\frac{24,27}{\sqrt{7}} \right); 64,97 + 2,58 \times \left(\frac{24,27}{\sqrt{7}} \right) \right] \\ &\equiv [41,30; 88,64]. \end{aligned} \quad (6.3)$$

Tabela 6.10 - Resultados das projeções para o preço em dólar da energia elétrica em New South Wales na Austrália, nos seis instantes de tempo posteriores a série histórica observada.

Horário da projeção	Projeção para o preço de energia	Conjunto <i>fuzzy</i>	Suporte u_i	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
00:00	15,20	A_1	[4,295; 28,56]	[17,03 ; 64,37]	[41,30; 88,64]
00:30	18,62	A_1	[4,295; 28,56]	[17,03 ; 64,37]	[41,30; 88,64]
01:00	22,83	A_1	[4,295; 28,56]	[17,03 ; 64,37]	[41,30; 88,64]
01:30	35,27	A_2	[16,43; 40,70]	[17,03 ; 64,37]	[41,30; 88,64]
02:00	37,60	A_2	[16,43; 40,70]	[17,03 ; 64,37]	[41,30; 88,64]
02:30	37,66	A_2	[16,43; 40,70]	[17,03 ; 64,37]	[41,30; 88,64]

Com o auxílio da construção dos intervalos de confiança *fuzzy*, o bom ajuste do modelo de previsão *fuzzy* que foi proposto, pôde ser constatado como se observa na Tabela (6.10), pois todas as seis projeções realizadas pelo modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₃, que foi identificado pela rotina de previsão desenvolvida neste estudo, estão entre os 6 intervalos que foram construídos com 99% de confiança, ou seja, com uma margem de erro tipo 1 (considerar que o verdadeiro valor não está contido no intervalo) igual a 1%, o que indica uma razoável e satisfatória condição para prever os valores dos preços em dólar da energia elétrica em New South Wales na Austrália, mediante o método de previsão que está sendo proposto, com o auxílio dos valores do preço em dólar de energia elétrica *defuzzificados*, durante uma das etapas do processo de identificação do modelo de previsão. Corroborando com os resultados obtidos com a metodologia de intervalo de confiança *fuzzy* que foi adotada anteriormente, na Figura (6.3) que trata dos valores reais versus os valores de previsão pelo método 1 proposto, para os preços em dólar da energia elétrica em New South Wales na Austrália, onde se percebe que o modelo identificado conseguiu capturar (adquirir) plenamente, grande parte da dependência inerente aos dados históricos de energia elétrica, pois em vários instantes de observação do preço de energia elétrica real, os valores de previsão gerados pelo modelo FSARIMA(3; 1; 2)(1; 0; 1)₁₃ estão bem próximos, também sugerindo um bom ajuste do modelo *fuzzy* sazonal que foi identificado.

Figura 6.3 - Série temporal do preço em dólar da energia elétrica em New South Wales na Austrália *versus* os valores das previsões pelo o método proposto (preços *defuzzificados*).



Passo 6: (b1) Obter a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy*, com alfa-nível pré-estabelecido, em seguida aplicar a função de autocorrelação estendida *fuzzy*.

Seguindo o passo 6(b1), da rotina descrita na Seção 5.6 e analisando os resultados obtidos nas Tabelas (6.11) e (6.12), tal que, os parâmetros sejam $p > 0$ e $q > 0$, se torna possível afirmar que o modelo identificado pela função de autocorrelação estendida *fuzzy* $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$ foi um ARMA(1; 4), pois esta combinação de parâmetros do modelo é estatisticamente diferente de zero, a um nível de significância de 10%, visto que, o p -valor = 0,087 < 10%, logo, se deve rejeitar a hipótese de que estes parâmetros ($p = 1$ e $q = 4$), sejam estatisticamente iguais a zero, ratificando assim a plena utilização do modelo identificado pela rotina proposta neste trabalho.

Figura 6.4 - Conjuntos *fuzzy* construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_j e o eixo X o preço em dólar da energia elétrica, com alfa-nível igual a 0,25 para ativar o suporte dos respectivos conjuntos *fuzzy*.

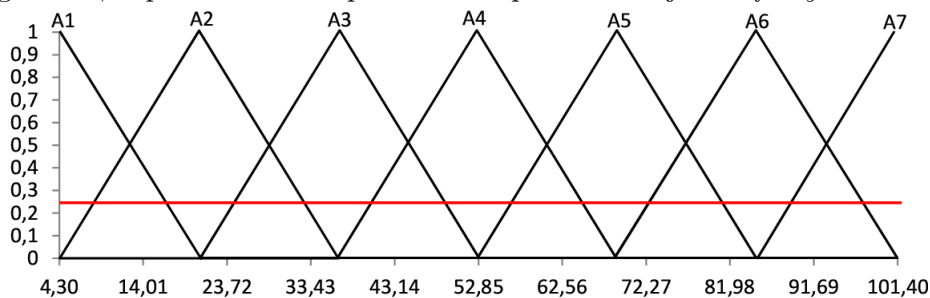


Tabela 6.11 - Resultados da função $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$ para o preço da energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,502	0,009	0,087	-0,057	-0,192	0,090
AR(1)	-0,480	0,112	0,118	0,033	-0,289	0,077
AR(2)	-0,235	-0,217	-0,282	0,063	-0,093	0,021
AR(3)	-0,472	0,039	-0,242	0,131	-0,073	0,013
AR(4)	-0,160	-0,172	-0,296	0,124	-0,029	-0,027
AR(5)	-0,529	0,333	-0,344	-0,103	-0,105	-0,258

Tabela 6.12 - Probabilidade de significância para os resultados da Tabela (6.11).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,001*	0,960	0,628	0,752	0,285	0,626
AR(1)	0,001*	0,559	0,470	0,869	0,087*	0,704
AR(2)	0,115	0,146	0,193	0,791	0,701	0,919
AR(3)	0,002*	0,814	0,310	0,624	0,784	0,949
AR(4)	0,293	0,258	0,224	0,642	0,912	0,891
AR(5)	0,001*	0,109	0,122	0,631	0,615	0,198

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

O processo de *fuzzificação* da série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, quando o valor do alfa-nível é igual a 0,50 ativou exatamente os suportes dos mesmos conjuntos *fuzzy* obtidos com alfa-nível igual a 0,25, conseqüentemente gerando resultados idênticos aos apresentados na Tabela (6.11) e (6.12), respectivamente.

Passo 6: (b2) Se obtém a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy*, com alfa-nível pré-estabelecido (Tabela 6.13) e em seguida aplicando a função de autocorrelação estendida *fuzzy* sobre esta série. Seguindo a rotina descrita na Seção 5.6, se obteve os valores de $W_{i,k,j}^{CA}(h; \alpha = 0,50; b)$, com *p-valor* correspondente, como se pode observar junto as Tabelas (6.15) e (6.16).

Figura 6.5 - Conjuntos *fuzzy* construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_j e o eixo X o preço em dólar da energia elétrica, com alfa-nível igual a 0,50 para ativar o suporte dos respectivos conjuntos *fuzzy*.

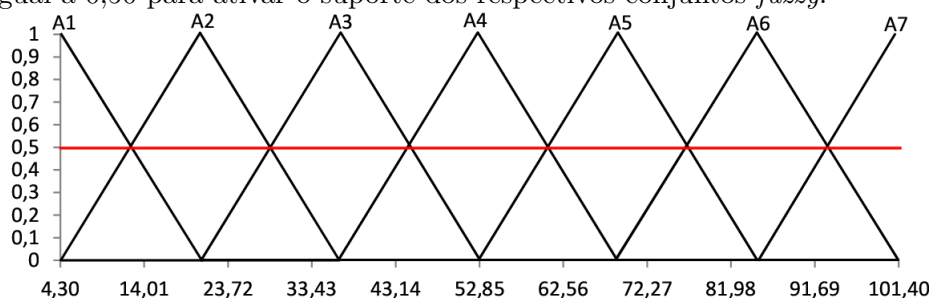


Tabela 6.13 - Processo de *fuzzificação* realizado a partir de um número *fuzzy* triangular, relação *fuzzy* lógica e a série dos preços em dólar da energia elétrica com alfa-nível de 0,25.

Horário observado	Preço em dólar	Intervalo correspondente	Série temporal <i>fuzzy</i>	Relação <i>fuzzy</i> lógica	Série $\alpha = 0,25$
00:00	23,72	[4,30; 20,47; 36,66]	A_2	-	0,005955
00:30	23,24	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005562
01:00	23,42	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005287
01:30	23,28	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000703
02:00	22,03	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002299
02:30	20,11	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001936
03:00	18,86	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,004998
03:30	17,42	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003898
04:00	16,43	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003122
04:30	18,41	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005025
05:00	18,00	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001250
05:30	20,36	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,004273
06:00	23,05	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003120
06:30	26,93	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000933
07:00	26,48	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002343
07:30	26,88	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000209
08:00	34,16	[20,47; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,000071
08:30	37,76	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,002297
09:00	34,07	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000741
09:30	34,61	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,004999
10:00	33,58	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000500
10:30	33,60	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,002157
11:00	30,13	[4,30; 20,47; 36,66]	A_2	$A_3 \rightarrow A_2$	0,000928
11:30	25,07	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000417
12:00	21,98	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001278
12:30	24,26	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001118
13:00	25,58	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002899
13:30	25,83	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003864
14:00	24,66	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002507
14:30	23,83	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000165
15:00	25,99	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003720
15:30	24,71	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000802
16:00	26,05	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005638
16:30	27,10	[20,47; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,003409
17:00	32,23	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,002447
17:30	42,06	[52,83; 69,03; 85,21]	A_5	$A_3 \rightarrow A_5$	0,003933
18:00	77,08	[69,03; 85,21; 101,40]	A_6	$A_5 \rightarrow A_6$	0,002012
18:30	89,24	[52,83; 69,03; 85,21]	A_5	$A_6 \rightarrow A_5$	0,001263
19:00	71,77	[36,66; 52,83; 69,03]	A_4	$A_5 \rightarrow A_4$	0,000644
19:30	56,21	[36,66; 52,83; 69,03]	A_4	$A_4 \rightarrow A_4$	0,002103
20:00	46,05	[36,66; 52,83; 69,03]	A_4	$A_4 \rightarrow A_4$	0,005501
20:30	45,25	[20,47; 36,66; 52,83]	A_3	$A_4 \rightarrow A_3$	0,004632
21:00	40,35	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,001077
21:30	35,33	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,002548
22:00	29,42	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000943
22:30	38,17	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,003372
23:00	36,02	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,002485
23:30	40,85	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,002773

Assim, face a coincidência dos conjuntos *fuzzy* ativados, quando $\alpha = 0,25$ e $\alpha = 0,50$, a identificação da função $W_{i,k,j}^{CA}(h; \alpha = 0,50; b)$, também retornou um modelo ARMA(1; 4).

Passo 6: (b3) A série temporal *fuzzy* é obtida pela ativação dos conjuntos *fuzzy*, com alfa-

Tabela 6.14 - Série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de *fuzzificação* realizado a partir de um número *fuzzy* triangular, relação *fuzzy* lógica e a série dos preços em dólar da energia elétrica com alfa-nível de 0,50.

Horário observado	Preço em dólar	Intervalo correspondente	Série temporal <i>fuzzy</i>	Relação <i>fuzzy</i> lógica	Série $\alpha = 0,50$
00:00	23,72	[4,30; 20,47; 36,66]	A_2	-	0,002021
00:30	23,24	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002986
01:00	23,42	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001676
01:30	23,28	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002491
02:00	22,03	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001029
02:30	20,11	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000346
03:00	18,86	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003578
03:30	17,42	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005044
04:00	16,43	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,004317
04:30	18,41	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003697
05:00	18,00	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002175
05:30	20,36	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003863
06:00	23,05	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001543
06:30	26,93	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000547
07:00	26,48	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003766
07:30	26,88	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001376
08:00	34,16	[20,47; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,004942
08:30	37,76	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,003813
09:00	34,07	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,004975
09:30	34,61	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,003314
10:00	33,58	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000572
10:30	33,60	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000104
11:00	30,13	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,001724
11:30	25,07	[4,30; 20,47; 36,66]	A_2	$A_3 \rightarrow A_2$	0,002762
12:00	21,98	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000106
12:30	24,26	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002444
13:00	25,58	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001870
13:30	25,83	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003499
14:00	24,66	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005537
14:30	23,83	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000022
15:00	25,99	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001265
15:30	24,71	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003338
16:00	26,05	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000115
16:30	27,10	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,004729
17:00	32,23	[20,47; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,002707
17:30	42,06	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,004600
18:00	77,08	[69,03; 85,21; 101,40]	A_6	$A_3 \rightarrow A_6$	0,001420
18:30	89,24	[69,03; 85,21; 101,40]	A_6	$A_6 \rightarrow A_6$	0,002985
19:00	71,77	[52,83; 69,03; 85,21]	A_5	$A_6 \rightarrow A_5$	0,001045
19:30	56,21	[36,66; 52,83; 69,03]	A_4	$A_5 \rightarrow A_4$	0,001266
20:00	46,05	[36,66; 52,83; 69,03]	A_4	$A_4 \rightarrow A_4$	0,003353
20:30	45,25	[36,66; 52,83; 69,03]	A_4	$A_4 \rightarrow A_4$	0,002033
21:00	40,35	[20,47; 36,66; 52,84]	A_3	$A_4 \rightarrow A_3$	0,003985
21:30	35,33	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,005624
22:00	29,42	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,001739
22:30	38,17	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000751
23:00	36,02	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000291
23:30	40,85	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000034

nível pré-estabelecido (Tabela 6.19) e posteriormente aplicando a função de autocorrelação

estendida *fuzzy* sobre esta série. Seguindo a rotina descrita na Seção 5.6, se obteve os valores de $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$, com *p-valor* correspondente (Tabelas (6.17) e (6.18)).

Tabela 6.15 - Resultados da função $W_{i,k,j}^{CA}(h; \alpha = 0,50; b)$ para o preço da energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,502	0,009	0,087	-0,057	-0,192	0,090
AR(1)	-0,480	0,112	0,118	0,033	-0,289	0,077
AR(2)	-0,235	-0,217	-0,282	0,063	-0,093	0,021
AR(3)	-0,472	0,039	-0,242	0,131	-0,073	0,013
AR(4)	-0,160	-0,172	-0,296	0,124	-0,029	-0,027
AR(5)	-0,529	0,333	-0,344	-0,103	-0,105	-0,258

Tabela 6.16 - Probabilidade de significância para os resultados da Tabela (6.15).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,001*	0,960	0,628	0,752	0,285	0,626
AR(1)	0,001*	0,559	0,470	0,869	0,087*	0,704
AR(2)	0,115	0,146	0,193	0,791	0,701	0,919
AR(3)	0,002*	0,814	0,310	0,624	0,784	0,949
AR(4)	0,293	0,258	0,224	0,642	0,912	0,891
AR(5)	0,001*	0,109	0,122	0,631	0,615	0,198

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Seguindo o passo 6(b3) da rotina descrita na Seção 5.6, os valores da função $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$ foram devidamente obtidos, com suas probabilidades de significância (*p-valor*) correspondentes. Estes resultados são apresentados na Tabela (6.17) e Tabela (6.18), respectivamente. As Tabelas (6.17 e 6.18) indicam uma ordem diferente do modelo de série temporal *fuzzy*, em comparação com a determinação realizada para alfa-nível = 0,25 e alfa-nível = 0,50. Isto é, a função de autocorrelação estendida $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$, identificou de forma estatisticamente significativa os parâmetros, tal que, a restrição $p > 0$ e $q > 0$ seja atendida, os modelos: ARMA(3; 2), ARMA(4; 2), ARMA(5; 2), ARMA(4; 5). As combinações paramétricas citadas acima são estatisticamente diferentes de zero, ao adotar um nível de significância de 10%, visto que, o *p-valor* = 0,099 < 10%, para o modelo ARMA(4;5), por exemplo. Portanto, dentre todos os resultados apresentados na Tabela (6.18), se deve rejeitar a hipótese de que estes parâmetros: $p = 3$ e $q = 2$ (ARMA(3;2)), $p = 4$ e $q = 2$ (ARMA(4;2)), $p = 5$ e $q = 2$ (ARMA(5;2)), como também $p = 4$ e $q = 5$ (ARMA(4;5)), sejam iguais a zero. Este fato ratifica a utilização dos modelos identificados pela rotina proposta neste trabalho, quando o valor do alfa-nível é 0,75.

Figura 6.6 - Conjuntos *fuzzy* construídos pela definição de pertinência triangular, onde o eixo *Y* indica a pertinência ao conjunto *fuzzy* A_j e o eixo *X* o preço em dólar da energia elétrica, com alfa-nível igual a 0,75 para ativar o suporte dos respectivos conjuntos *fuzzy*.

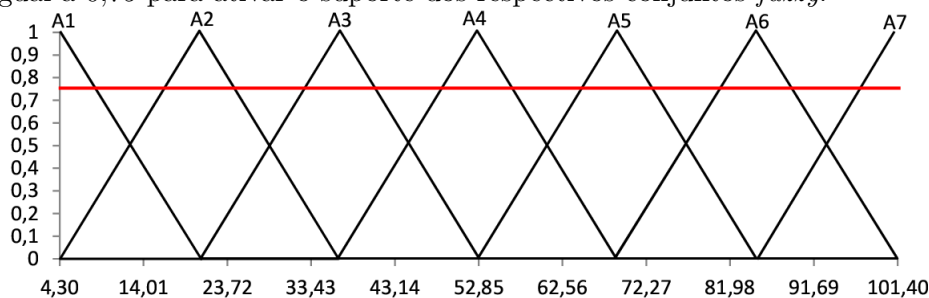


Tabela 6.17 - Resultados da função $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$ para o preço da energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,616	0,114	0,049	-0,055	-0,105	0,272
AR(1)	-0,534	0,240	0,137	0,027	-0,081	0,254
AR(2)	-0,370	-0,216	-0,285	0,146	-0,129	0,286
AR(3)	-0,476	0,064	-0,407	-0,346	-0,004	0,281
AR(4)	-0,278	-0,180	-0,536	-0,142	0,147	0,277
AR(5)	-0,196	-0,120	0,340	-0,178	-0,119	-0,010

Tabela 6.18 - Probabilidade de significância para os resultados da Tabela (6.17).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,000*	0,555	0,801	0,778	0,590	0,167
AR(1)	0,000*	0,247	0,408	0,902	0,655	0,172
AR(2)	0,012*	0,157	0,267	0,540	0,508	0,105
AR(3)	0,002*	0,718	0,026*	0,122	0,985	0,117
AR(4)	0,068*	0,254	0,001*	0,428	0,419	0,099*
AR(5)	0,2042	0,448	0,051*	0,331	0,515	0,956

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Passo 7: Gerar previsões pelo modelo de série temporal identificado no passo 6 (b(1;2;3));

Em virtude da existência de cinco modelos significativos para alfa-nível igual a 0,25, 0,50 e 0,75, foi ajustado o modelo de maior ordem dentre os possíveis candidatos, ou seja, um modelo ARMA (4;5). Este fato se justifica, devido o procedimento de ajuste do modelo de previsão que é utilizado neste estudo, necessariamente buscar uma estimativa paramétrica otimizada, a partir dos valores máximos determinados para p e q , onde o modelo de previsão retorne o menor erro possível. Neste contexto, o método de estimação dos parâmetros utilizado possui capacidade de testar, todas as combinações possíveis entre os parâmetros $0 < p \leq 5$ e $0 < q \leq 5$. Logo, todas as combinações de ordens inferiores as identificadas pela função $W_{i,k,j}^{CA}$ com 0,25 e 0,75, são contempladas integralmente.

Tabela 6.19 - Série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de *fuzzificação* realizado a partir de um número *fuzzy* triangular, relação *fuzzy* lógica e a série dos preços em dólar da energia elétrica com alfa-nível de 0,75.

Horário observado	Preço em dólar	Intervalo correspondente	Série temporal <i>fuzzy</i>	Relação <i>fuzzy</i> lógica	Série $\alpha = 0,75$
00:00	23,72	[4,30; 20,47; 36,66]	A_2	-	0,005293
00:30	23,24	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000969
01:00	23,42	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000506
01:30	23,28	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002428
02:00	22,03	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002745
02:30	20,11	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001884
03:00	18,86	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002127
03:30	17,42	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005121
04:00	16,43	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000557
04:30	18,41	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003609
05:00	18,00	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000769
05:30	20,36	[4,30; 16,43; 28,56]	A_2	$A_2 \rightarrow A_2$	0,002098
06:00	23,05	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000827
06:30	26,93	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000711
07:00	26,48	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,004905
07:30	26,88	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001562
08:00	34,16	[20,47; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,002053
08:30	37,76	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,003623
09:00	34,07	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,005540
09:30	34,61	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000680
10:00	33,58	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,004218
10:30	33,60	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000442
11:00	30,13	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,001575
11:30	25,07	[4,30; 20,47; 36,66]	A_2	$A_3 \rightarrow A_2$	0,001590
12:00	21,98	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002022
12:30	24,26	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005681
13:00	25,58	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,007139
13:30	25,83	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001078
14:00	24,66	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,004295
14:30	23,83	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000151
15:00	25,99	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,000508
15:30	24,71	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,002475
16:00	26,05	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,011388
16:30	27,10	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,003298
17:00	32,23	[20,47; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,004880
17:30	42,06	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,004379
18:00	77,08	[69,03; 85,21; 101,40]	A_6	$A_3 \rightarrow A_6$	0,000268
18:30	89,24	[69,03; 85,21; 101,40]	A_6	$A_6 \rightarrow A_6$	0,001309
19:00	71,77	[52,83; 69,03; 85,21]	A_5	$A_6 \rightarrow A_5$	0,000578
19:30	56,21	[36,66; 52,83; 69,03]	A_4	$A_5 \rightarrow A_4$	0,002067
20:00	46,05	[36,66; 52,83; 69,03]	A_4	$A_4 \rightarrow A_4$	0,002413
20:30	45,25	[36,66; 52,83; 69,03]	A_4	$A_4 \rightarrow A_4$	0,000000
21:00	40,35	[20,47; 36,66; 52,84]	A_3	$A_4 \rightarrow A_3$	0,003647
21:30	35,33	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,001829
22:00	29,42	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,001151
22:30	38,17	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,000654
23:00	36,02	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,001148
23:30	40,85	[20,47; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,004268

Posteriormente ao processo de *fuzzificação* realizado com a inclusão de um *alfa-nível*, então realizada para os dados de preço da energia elétrica em New South Wales na Austrália, (Tabelas 6.13, 6.14, 6.19), vários pares de relações *fuzzy* lógicas foram obtidos aos valores da série temporal, consequentemente implicando na concepção de grupos

de relações *fuzzy* lógicas, equivalentes aos três valores de *alfa-nível* adotados, como é plenamente possível verificar junto a Tabela (6.20).

Pelos resultados da Tabela (6.20), quanto aos grupos de relações *fuzzy* lógicas, se percebe que quando o *alfa-nível* foi igual a 0,50 ou 0,75, o processo de *fuzzificação* estabeleceu as mesmas transições entre os conjuntos *fuzzy* que foram construídos e conseqüentemente ativados. Neste contexto, com valor do *alfa-nível* de 0,50 ou 0,75, os mesmos suportes foram ativados em relação ao preço da energia elétrica, após a partição do universo de discurso em 7 conjuntos *fuzzy*.

É importante ressaltar, que devido a grande concentração dos dados em torno do valor mediano (26,68), para a série temporal dos preços da energia elétrica em New South Wales, na Austrália, os conjuntos *fuzzy* A_1 e A_7 , localizados nas extremidades do universo de discurso, os quais também são representados pelo plano cartesiano das Figuras (6.4); (6.5) e (6.6), não foram ativados pela pertinência máxima obtida, em relação aos dados de energia elétrica com *alfa-nível* 0,25, 0,50 ou 0,75, no processo de *fuzzificação*, fato relevante para acurácia do modelo identificado.

Mediante os resultados das Tabelas (6.21) e (6.22), após várias simulações para o ajuste do modelo identificado pela função de autocorrelação estendida *fuzzy*, onde a acurácia do modelo de previsão retornasse o menor erro possível, se chegou ao modelo FSARIMA(2; 1; 2)(1; 0; 1)₁₂, ou seja, incorporado a classe de modelos identificados pelo método proposto.

Tabela 6.20 - Grupos de relações *fuzzy* lógicas ativadas com *alfa-nível* de 0,25; 0,50 e 0,75.

Grupos	Relação <i>fuzzy</i> ($\alpha = 0,25$)	Relação <i>fuzzy</i> ($\alpha = 0,50$)	Relação <i>fuzzy</i> ($\alpha = 0,75$)
Grupo 1	$A_2 \rightarrow A_2, A_2 \rightarrow A_3$	$A_2 \rightarrow A_2, A_2 \rightarrow A_3$	$A_2 \rightarrow A_2, A_2 \rightarrow A_3$
Grupo 2	$A_3 \rightarrow A_2, A_3, A_5$	$A_3 \rightarrow A_2, A_3, A_6$	$A_3 \rightarrow A_2, A_3, A_6$
Grupo 3	$A_4 \rightarrow A_3, A_4 \rightarrow A_4$	$A_4 \rightarrow A_3, A_4 \rightarrow A_4$	$A_4 \rightarrow A_3, A_4 \rightarrow A_4$
Grupo 4	$A_5 \rightarrow A_4, A_5 \rightarrow A_6$	$A_5 \rightarrow A_4$	$A_5 \rightarrow A_4$
Grupo 5	$A_6 \rightarrow A_5$	$A_6 \rightarrow A_5, A_6 \rightarrow A_6$	$A_6 \rightarrow A_5, A_6 \rightarrow A_6$

O modelo de previsão *fuzzy* ajustado, possui parâmetros significativos (estatisticamente diferentes de zero, ao nível de significância $\alpha = 10\%$), como observado na Tabela (6.21), logo, os parâmetros são pertinentes ao modelo de série temporal *fuzzy* sazonal que foi identificado, o que implica em manter no modelo de previsão *fuzzy*, todos estes parâmetros que foram testados.

Tabela 6.21 - Teste de significância dos parâmetros do modelo FSARIMA(2; 1; 2)(1; 0; 1)₁₂.

Modelo	Coefficiente	Erro padrão	Estatística t	p -valor
AR(1)	1,1964	0,1433	8,35	0,000*
AR(2)	-0,4914	0,1429	-3,44	0,001*
SAR(12)	0,9010	0,3519	2,56	0,014*
MA(1)	0,8462	0,0012	733,27	0,000*
MA(2)	0,1820	0,0222	8,20	0,000*
SMA(12)	-0,7240	0,2172	-3,33	0,002*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Tabela 6.22 - Teste Q (Ljung-Box) para o modelo FSARIMA(2; 1; 2)(1; 0; 1)₁₂.

Fonte de variação	Lag	Qui-quadrado	Grau de liberdade	p -valor
Teste Q "Ljung-Box"	12,000	9,900	6,000	0,128
⋮	24,000	29,700	18,000	0,040*
⋮	36,000	45,200	30,000	0,037*
⋮	48,000	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

A série dos preços da energia elétrica em New South Wales na Austrália, representa uma série temporal com sazonalidade estocástica, devido ao p -valor do teste no período sazonal $s = 12$, que foi utilizado no ajuste do modelo identificado ser igual a $0,128 > \alpha = 10\%$, como demonstra a estatística de teste Q na Tabela (6.22). Neste contexto, justifica-se a utilização da metodologia de série temporal para prever o preço da energia elétrica em New South Wales na Austrália, mediante o modelo FSARIMA(2; 1; 2)(1; 0; 1)₁₂. Baseado nos resultados obtidos com o teste Q, no período da série temporal dos preços da energia elétrica utilizado para ajustar o modelo, quarenta e sete previsões foram realizadas objetivando validar este modelo (ver, Tabela 6.23).

Passo 8: Comparar as previsões obtidas no passo 7, com dados reais da série e medir o MASE.

Buscando reafirmar o bom ajuste do modelo FSARIMA(2; 1; 2)(1; 0; 1)₁₂, face aos resultados obtidos até este momento, quanto aos dados do preço de energia elétrica utilizados neste segundo estudo de caso, com a inclusão do *alfa-nível* no processo de *fuzzificação*, novos procedimentos foram realizados. Neste sentido, o valor obtido com a escala do erro médio absoluto do modelo mencionado acima, foi comparado ao dos três modelos de previsão clássicos (dados não *fuzzificados*) propostos por Dong et al. (2011). Desta forma, buscou-se medir a capacidade preditiva do modelo ajustado, mediante o método de iden-

tificação com *alfa-nível* (0,25; 0,5; 0,75) proposto neste trabalho, para então avaliar a acurácia do modelo FSARIMA(2; 1; 2)(1; 0; 1)₁₂, em relação ao preço da energia elétrica que é apresentado na Tabela (6.23).

Ao verificar o MASE de 0,71 na Tabela (6.23), o qual foi retornado pelo modelo FSARIMA(2; 1; 2)(1; 0; 1)₁₂, que foi proposto neste trabalho, se tem um indicativo de obtenção do menor erro de previsão, frente aos erros produzidos pelos três métodos propostos por Dong et al. (2011). Este fato sinaliza uma significativa capacidade de previsão do modelo *fuzzy*, com período de sazonalidade considerado em $s = 12$ ao modelo ajustado, sobretudo, em comparação aos três modelos desenvolvidos pelo autor mencionado acima, quanto ao preço da energia elétrica em New South Wales na Austrália, como se pode observar na Tabela (6.23), para o período considerado.

A eficiência na capacidade preditiva do modelo que está sendo proposto, com a incorporação do *alfa-nível* no processo de identificação, se torna mais evidente ao analisar a Figura (6.7), visto que, os valores das previsões do preço da energia elétrica em New South Wales na Austrália, foram plotados *versus* os valores reais desta mesma característica observável (variável). Dentro do período compreendido pela série temporal analisada, o bom ajuste dos dados ao modelo de previsão *fuzzy* sugerido ficou evidenciado graficamente, devido a localização dos valores de previsão do preço de energia elétrica, bem próximos dos reais observados, os quais auxiliaram na rotina de identificação do modelo de previsão *fuzzy* sazonal, que foi proposto a estes dados.

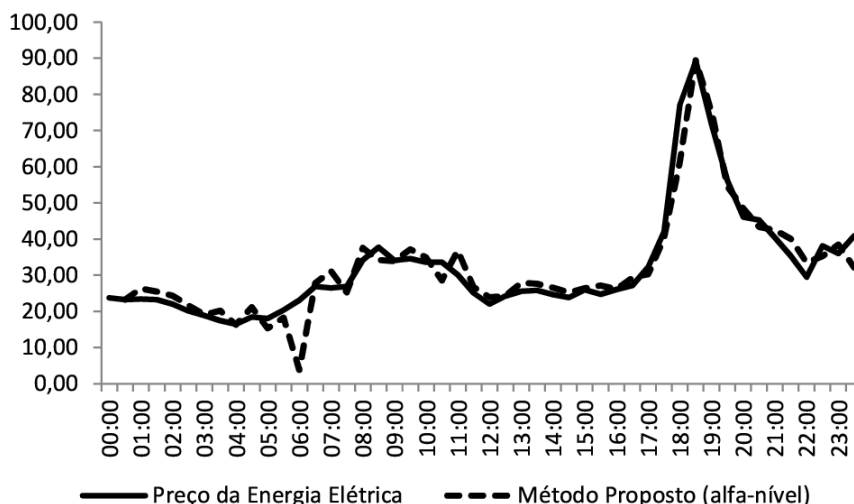
Com o intuito de comprovar novamente o bom ajuste dos dados de preço da energia elétrica, ao modelo FSARIMA(2;1;2)(1;0;1)₁₂ que foi identificado, se projetou o valor do preço da energia elétrica para um horizonte de 6 instantes de tempo, referentes ao dia seguinte do período utilizado para ajuste do modelo de previsão *fuzzy*, ou seja, para $t = 00:00; 00:30; \dots; 02:30$, tal como, observado na Tabela (6.24). As projeções apresentadas na Tabela (6.24) tiveram suas acurácias medidas, a partir do conceito de intervalo de confiança *fuzzy* (Equações ?? e ??), como estabelecido por Wu, (2009). A construção dos intervalos foi realizada com 99% de confiabilidade, onde se tornou possível constatar mais uma vez, que os dados ajustaram-se muito bem ao modelo identificado, face às projeções localizadas dentro dos intervalos. Neste contexto, como todos os valores projeta-

Tabela 6.23 - Comparação dos resultados de previsão pelo método proposto (preços com α -nível) em relação aos três métodos propostos em Dong et al. (2011).

Instante observado em horas	Preço da energia em dólar	Método SA-ARIMA (Dong et al. 2011)	Método EMD-SA-ARIMA (Dong et al. 2011)	Método SARIMA (Dong et al. 2011)	Método proposto: Treinamento	Método proposto: Validação
00:00	23,72	24,30	24,35	22,99	-	-
00:30	23,24	23,68	23,95	21,69	23,2227	-
01:00	23,42	22,55	22,85	22,39	26,2843	-
01:30	23,28	23,04	23,32	20,62	25,4621	-
02:00	22,03	21,55	21,75	16,17	24,3015	-
02:30	20,11	17,70	17,81	13,85	21,6535	-
03:00	18,86	15,71	15,75	12,30	19,1502	-
03:30	17,42	14,36	14,35	11,78	20,1731	-
04:00	16,43	13,91	13,84	11,48	16,2981	-
04:30	18,41	13,90	13,78	12,90	21,1430	-
05:00	18,00	14,99	14,80	16,04	15,2737	-
05:30	20,36	17,63	17,34	17,69	18,2856	-
06:00	23,05	19,33	18,94	26,70	23,7624	-
06:30	26,93	27,11	26,47	30,95	27,8372	-
07:00	26,48	30,23	29,41	30,29	31,1063	-
07:30	26,88	29,74	28,82	38,95	25,2378	-
08:00	34,16	37,73	36,43	42,32	37,5456	-
08:30	37,76	41,11	39,54	40,29	34,1711	-
09:00	34,07	39,24	37,61	40,25	33,9212	-
09:30	34,61	39,36	37,58	39,26	37,1136	-
10:00	33,58	38,67	36,78	39,08	34,8813	-
10:30	33,60	38,79	36,76	37,10	28,5448	-
11:00	30,13	36,91	34,84	34,04	36,8192	-
11:30	25,07	33,97	31,96	34,08	26,7411	-
12:00	21,98	33,61	31,50	33,17	23,8771	-
12:30	24,26	33,11	30,92	32,27	24,2737	-
13:00	25,58	32,53	30,27	31,12	27,8692	-
13:30	25,83	31,70	29,39	29,82	27,6187	-
14:00	24,66	30,64	28,30	28,76	26,5848	-
14:30	23,83	29,50	27,15	28,46	25,2088	-
15:00	25,99	29,55	27,10	28,11	26,3536	-
15:30	24,71	29,10	26,58	29,17	27,1288	-
16:00	26,05	30,35	27,63	32,13	26,1575	-
16:30	27,10	33,27	30,18	36,93	29,3302	-
17:00	32,23	38,24	34,56	60,08	30,2932	-
17:30	42,06	58,05	52,28	84,17	40,2550	-
18:00	77,08	84,13	75,51	76,65	61,4519	-
18:30	89,24	78,83	70,50	65,37	89,4908	-
19:00	71,77	67,17	59,85	51,50	75,2185	-
19:30	56,21	53,27	47,30	46,89	54,5628	-
20:00	46,05	48,00	42,47	43,82	48,4922	-
20:30	45,25	45,57	40,18	37,35	43,3356	-
21:00	40,35	39,71	34,88	34,61	-	42,4822
21:30	35,33	36,66	32,10	30,89	-	40,0179
22:00	29,42	32,50	28,36	35,12	-	33,5150
22:30	38,17	37,76	32,83	30,39	-	35,2881
23:00	36,02	33,08	28,66	32,97	-	38,4771
23:30	40,85	36,51	31,52	25,75	-	32,1158
MASE	-	1,02	0,99	1,76	-	0,71

dos encontram-se dentro dos intervalos, tem-se um indicativo de boa capacidade preditiva

Figura 6.7 - Série temporal do preço em dólar da energia elétrica em New South Wales na Austrália *versus* os valores das previsões pelo método proposto (preços com *alfa-nível*).



do modelo de previsão *fuzzy* proposto, em conjunto com o *alfa-nível* agregado ao processo de *fuzzificação* dos dados, consolidando a aplicabilidade deste modelo.

Tabela 6.24 - Resultados das projeções para o preço em dólar da energia elétrica em New South Wales na Austrália, correspondendo aos seis instantes de tempo posteriores ao período original da série histórica observada, mediante o método proposto (*alfa-nível*).

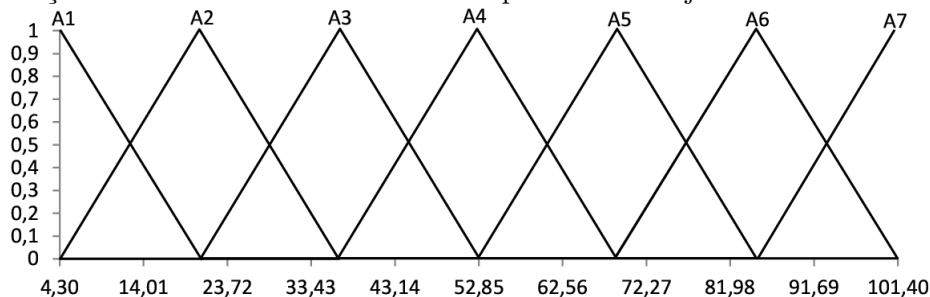
Horário da Projeção	Projeção para o preço de energia	Conjunto fuzzy	Suporte u_i	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
00:00	42,41	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
00:30	44,82	A_4	[36,66; 69,03]	[17,03 ; 64,37]	[41,30; 88,64]
01:00	42,92	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
01:30	39,37	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
02:00	35,32	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
02:30	33,27	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]

Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para identificar a ordem do modelo de série temporal *fuzzy*, sobre os dados históricos de interesse.

Passo 6(c): Para a função $\overline{W}_{i,k,j}^{RA}(h)$, se faz necessário proceder a seleção da função de autocorrelação amostral média aleatória, a partir de cada suporte dos conjuntos *fuzzy* ativados, e que foram construídos no passo 5, como se observa na Figura (6.8). Assim, em seguida deve-se aplicar a função de autocorrelação estendida *fuzzy*, sobre a série resultante deste processo de *fuzzificação* dos preços de energia elétrica. Neste sentido, adotando o passo 6(c), da rotina descrita na Seção 5.6, se obteve os valores de $\overline{W}_{i,k,j}^{RA}(h)$, com suas

probabilidades de significância (*p*-valor) correspondentes, que são apresentados na Tabela (6.27) e Tabela (6.28), respectivamente, possibilitando assim a identificação do modelo de série temporal *fuzzy*, que melhor capture toda a variabilidade inerente a série das *fac's* médias aleatórias que é apresentada na Tabela (6.26). O processo de *fuzzificação* que

Figura 6.8 - Conjuntos *fuzzy* pela definição de número *fuzzy* triangular, onde o eixo *Y* indica a pertinência ao conjunto *fuzzy* A_j e o eixo *X* o preço em dólar da energia elétrica, para função de autocorrelação amostral média aleatória nos suportes dos conjuntos ativados.



resultou nos valores da função de autocorrelação média aleatória (Tabela 6.26), indicaram grupos de relações *fuzzy* lógicas que não contemplaram os conjuntos *fuzzy* A_1 e A_7 , algo semelhante ao que ocorreu com os dados *defuzzificados* e com *alfa-nível*. Assim como discutido a priori, acredita-se que a justificativa deste fato seja a característica particular dos dados de energia elétrica, no período que foi selecionado para compor a série temporal, isto é, a maioria dos valores do preço de energia elétrica estão muito próximos do valor mediano.

Tabela 6.25 - Grupos de relações *fuzzy* lógicas ativadas para os dados da Tabela (6.26).

Grupo ativado	Relações <i>fuzzy</i> lógicas
Grupo 1	$A_2 \rightarrow A_2, A_2 \rightarrow A_3$
Grupo 2	$A_3 \rightarrow A_2, A_3 \rightarrow A_3, A_3 \rightarrow A_5$
Grupo 3	$A_4 \rightarrow A_3, A_4 \rightarrow A_4$
Grupo 4	$A_5 \rightarrow A_4, A_5 \rightarrow A_6$
Grupo 5	$A_6 \rightarrow A_5$

Analisando os resultados das Tabelas (6.27 e 6.28), se pode constatar que a função $\overline{W}_{j,k,j}^{RA}(h)$ identificou de forma significativa (estatisticamente diferente de zero para $\alpha = 10\%$), os seguintes modelos, onde $p > 0$ e $q > 0$: ARMA(3;1), ARMA(4;1), ARMA(3;2), ARMA(4;2) e ARMA(5;2), o que sugere ajustar um modelo de previsão *fuzzy* com os parâmetros $p = 5$ e $q = 2$, e, tal como já foi comentado neste estudo, o *software* utilizado para ajustar o modelo de previsão, possui capacidade suficiente quanto ao teste da

significância de todas as possíveis combinações de parâmetros, identificadas pela função de autocorrelação estendida *fuzzy* proposta, logo, considerando o maior grau identificado para p e q , as demais ordens inferiores serão contempladas.

Tabela 6.26 - Série histórica do preço em dólar da energia elétrica em New South Wales na Austrália, com o processo de *fuzzificação* realizado a partir de um número *fuzzy* triangular, relação *fuzzy* lógica e a série das funções de autocorrelação amostrais médias aleatórias.

Horário observado	Preço em dólar	Pertinência máxima	Intervalo correspondente	Série temporal <i>fuzzy</i>	Relação <i>fuzzy</i> lógica	<i>fac</i> média aleatória
00:00	23,72	0,799	[4,30; 20,47; 36,66]	A_2	-	0,028987
00:30	23,24	0,829	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,011621
01:00	23,42	0,818	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,014608
01:30	23,28	0,826	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,013511
02:00	22,03	0,904	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,009496
02:30	20,11	0,978	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,013524
03:00	18,86	0,900	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,005076
03:30	17,42	0,811	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,012368
04:00	16,43	0,750	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,033767
04:30	18,41	0,873	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,006643
05:00	18,00	0,847	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,019836
05:30	20,36	0,993	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,012370
06:00	23,05	0,841	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,017184
06:30	26,93	0,601	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,020580
07:00	26,48	0,629	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,023504
07:30	26,88	0,604	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,011813
08:00	34,16	0,845	[20,49; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,019687
08:30	37,76	0,932	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,006894
09:00	34,07	0,840	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,006417
09:30	34,61	0,873	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,015071
10:00	33,58	0,810	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,013471
10:30	33,60	0,811	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,004959
11:00	30,13	0,596	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,010497
11:30	25,07	0,716	[4,30; 20,47; 36,66]	A_2	$A_3 \rightarrow A_2$	0,009589
12:00	21,98	0,907	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,027913
12:30	24,26	0,766	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,022131
13:00	25,58	0,684	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,001968
13:30	25,83	0,669	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,008580
14:00	24,66	0,741	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,014329
14:30	23,83	0,792	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,011328
15:00	25,99	0,659	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,025505
15:30	24,71	0,738	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,011497
16:00	26,05	0,655	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,024272
16:30	27,10	0,590	[4,30; 20,47; 36,66]	A_2	$A_2 \rightarrow A_2$	0,020332
17:00	32,23	0,726	[20,49; 36,66; 52,84]	A_3	$A_2 \rightarrow A_3$	0,017565
17:30	42,06	0,666	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,019345
18:00	77,08	0,502	[52,85; 69,03; 85,21]	A_5	$A_3 \rightarrow A_5$	0,001456
18:30	89,24	0,751	[69,03; 85,21; 101,40]	A_6	$A_5 \rightarrow A_6$	0,016774
19:00	71,77	0,831	[52,85; 69,03; 85,21]	A_5	$A_6 \rightarrow A_5$	0,014923
19:30	56,21	0,792	[36,66; 52,84; 69,03]	A_4	$A_5 \rightarrow A_4$	0,004876
20:00	46,05	0,580	[36,66; 52,84; 69,03]	A_4	$A_4 \rightarrow A_4$	0,003172
20:30	45,25	0,531	[36,66; 52,84; 69,03]	A_4	$A_4 \rightarrow A_4$	0,016334
21:00	40,35	0,772	[20,49; 36,66; 52,84]	A_3	$A_4 \rightarrow A_3$	0,015835
21:30	35,33	0,918	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,026170
22:00	29,42	0,552	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,017494
22:30	38,17	0,907	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,020101
23:00	36,02	0,960	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,018196
23:30	40,85	0,741	[20,49; 36,66; 52,84]	A_3	$A_3 \rightarrow A_3$	0,008476

Nota: *fac* (Função de autocorrelação amostral).

Como a função de autocorrelação estendida *fuzzy* média aleatória, retornou uma identificação similar (ARMA(5;2)), a que foi obtida pela função de autocorrelação estendida *fuzzy* com alfa-nível (0,75), no passo 6 (b3), conseqüentemente o modelo de previsão *fuzzy* para o preço em dólar da energia elétrica, mediante a identificação pela função $\overline{W}_{i,k,j}^{RA}(h)$ resultará no mesmo modelo, ou seja, um modelo FSARIMA(2;1;2)(1;0;1)₁₂. Assim como ocorreu na identificação com *alfa-nível* igual a 0,75, o modelo com parâmetros significativos para os preços de energia elétrica, que estão sendo utilizados neste exemplo, e que resultaram em uma identificação de um processo ARMA(5;2), isto indicou uma convergência ao modelo de previsão FSARIMA(2;1;2)(1;0;1)₁₂, invariavelmente como já havia sido ajustado anteriormente (passo 6.b3), o que ratifica a identificação processada mediante o *alfa-nível* na rotina de *fuzzificação* dos preços de energia elétrica.

Tabela 6.27 - Resultados da função $\overline{W}_{i,k,j}^{RA}(h)$ para o preço da energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,465	0,008	0,067	-0,178	0,182	0,044
AR(1)	-0,477	0,041	-0,007	-0,106	0,154	0,048
AR(2)	-0,312	-0,223	-0,155	-0,017	0,038	0,026
AR(3)	-0,324	-0,515	-0,377	-0,116	0,061	0,032
AR(4)	-0,238	0,323	-0,406	-0,185	0,086	0,012
AR(5)	0,366	-0,087	-0,441	0,104	-0,109	0,112

Tabela 6.28 - Probabilidades de significância para os resultados da Tabela (6.27).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,001*	0,966	0,701	0,309	0,308	0,808
AR(1)	0,001*	0,823	0,967	0,527	0,428	0,788
AR(2)	0,036*	0,151	0,318	0,920	0,822	0,876
AR(3)	0,032*	0,001*	0,016*	0,531	0,758	0,855
AR(4)	0,119	0,091*	0,008*	0,359	0,666	0,950
AR(5)	0,018*	0,575	0,006*	0,586	0,578	0,574

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6 (c); Após a função de autocorrelação estendida ter determinado a ordem de $p = 5$ e $q = 2$, o modelo ajustado teve os seus parâmetros testados em termos de significância estatística ($\alpha = 1\%$). Na Tabela (6.29), se pode verificar que $p = 3, 4$ e 5 , apresentam *p-valor* = 0,570, 0,590 e 0,340, respectivamente, isto é, todos superiores a 1% (nível de significância). Este fato indica que o modelo de previsão ajustado, ao realizar uma previsão para o preço de energia elétrica em New South Wales na Austrália, não irá considerar estes parâmetros, pois, os mesmos são estatisticamente iguais à zero. No entanto, os demais parâmetros

ajustados do modelo FSARIMA(2;1;2)(1;0;1)₂₄, necessariamente devem ser preservados no modelo de previsão, devido $p\text{-valor} < 1\%$, para todos estes parâmetros, inclusive os termos autoregressivo SAR(24) e de média móvel SMA(24).

Tabela 6.29 - Teste de significância dos parâmetros do modelo FSARIMA(2; 1; 2)(1; 0; 1)₂₄.

Modelo	Coefficiente	Erro padrão	Estatística t	$p\text{-valor}$
AR(1)	1,0253	0,1850	5,54	0,000*
AR(2)	-1,2082	0,2624	-4,60	0,000*
SAR(24)	0,9391	0,2167	4,33	0,000*
MA(1)	0,5556	0,1875	2,96	0,005*
MA(2)	-1,0389	0,0262	-39,60	0,000*
SMA(24)	-0,7101	0,2072	-3,43	0,001*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 1\%$.

Na Tabela (6.30) é possível perceber que o teste de *Ljung-Box* indicou um componente de sazonalidade $s = 24$, que não é significante ao nível de 1%, isto é, os erros do modelo são ruído branco. Vale ressaltar que os termos $s = 12$ e 36 , são estatisticamente diferentes de zero ($< 1\%$), o que não representa um ruído branco. Testes computacionais otimizaram os parâmetros do modelo de previsão ajustado em função do menor MASE, onde $s = 24$ retornou o menor erro.

Tabela 6.30 - Teste Q (Ljung-Box) para o modelo FSARIMA(2; 1; 2)(1; 0; 1)₂₄.

Fonte de variação	Lag	Qui-quadrado	Grau de liberdade	$p\text{-valor}$
Teste Q “Ljung-Box”	12,000	9,300	3,000	0,026*
⋮	24,000	25,300	15,000	0,047*
⋮	36,000	46,600	27,000	0,011*
⋮	48,000	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 1\%$.

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

Como forma de avaliar a acurácia do método proposto para prever o preço de energia elétrica em New South Wales na Austrália, na Tabela (6.31) foi realizada uma comparação entre as previsões produzidas pelo modelo FSARIMA(2;1;2)(1;0;1)₂₄, em relação a três métodos de previsão propostos em Dong et al. (2011). Pelo critério de decisão adotado neste estudo, isto é, o menor valor do MASE, consultando a Tabela (6.31) é possível perceber que o método proposto com a fac média aleatória, retornou um erro inferior em comparação aos outros métodos propostos em Dong et al. (2011). Este fato indica uma acurácia alcançada pelo modelo FSARIMA(2;1;2)(1;0;1)₂₄, superior aos demais métodos,

Tabela 6.31 - Comparação dos resultados de previsão pelo método proposto (função de auto-correlação média aleatória) em relação aos três métodos propostos em Dong et al. (2011).

Instante observado em horas	Preço da energia em dólar	Método SA-ARIMA (Dong et al. 2011)	Método EMD-SA-ARIMA (Dong et al. 2011)	Método SARIMA (Dong et al. 2011)	Método proposto: Treinamento	Método proposto: Validação
00:00	23,72	24,30	24,35	22,99	-	-
00:30	23,24	23,68	23,95	21,69	21,5324	-
01:00	23,42	22,55	22,85	22,39	28,0511	-
01:30	23,28	23,04	23,32	20,62	24,0193	-
02:00	22,03	21,55	21,75	16,17	17,3235	-
02:30	20,11	17,70	17,81	13,85	17,0674	-
03:00	18,86	15,71	15,75	12,30	20,2738	-
03:30	17,42	14,36	14,35	11,78	22,7035	-
04:00	16,43	13,91	13,84	11,48	13,8303	-
04:30	18,41	13,90	13,78	12,90	14,0799	-
05:00	18,00	14,99	14,80	16,04	12,2880	-
05:30	20,36	17,63	17,34	17,69	21,0565	-
06:00	23,05	19,33	18,94	26,70	08,5416	-
06:30	26,93	27,11	26,47	30,95	29,1102	-
07:00	26,48	30,23	29,41	30,29	35,3622	-
07:30	26,88	29,74	28,82	38,95	26,1785	-
08:00	34,16	37,73	36,43	42,32	42,3406	-
08:30	37,76	41,11	39,54	40,29	41,3952	-
09:00	34,07	39,24	37,61	40,25	34,0139	-
09:30	34,61	39,36	37,58	39,26	36,4286	-
10:00	33,58	38,67	36,78	39,08	33,5590	-
10:30	33,60	38,79	36,76	37,10	32,7549	-
11:00	30,13	36,91	34,84	34,04	31,2333	-
11:30	25,07	33,97	31,96	34,08	26,2890	-
12:00	21,98	33,61	31,50	33,17	22,0825	-
12:30	24,26	33,11	30,92	32,27	22,3555	-
13:00	25,58	32,53	30,27	31,12	28,1812	-
13:30	25,83	31,70	29,39	29,82	26,5000	-
14:00	24,66	30,64	28,30	28,76	22,1887	-
14:30	23,83	29,50	27,15	28,46	22,1117	-
15:00	25,99	29,55	27,10	28,11	26,2447	-
15:30	24,71	29,10	26,58	29,17	27,6892	-
16:00	26,05	30,35	27,63	32,13	24,2262	-
16:30	27,10	33,27	30,18	36,93	25,0211	-
17:00	32,23	38,24	34,56	60,08	28,4974	-
17:30	42,06	58,05	52,28	84,17	41,4175	-
18:00	77,08	84,13	75,51	76,65	64,2237	-
18:30	89,24	78,83	70,50	65,37	91,1164	-
19:00	71,77	67,17	59,85	51,50	80,5248	-
19:30	56,21	53,27	47,30	46,89	56,3761	-
20:00	46,05	48,00	42,47	43,82	51,0200	-
20:30	45,25	45,57	40,18	37,35	46,4195	-
21:00	40,35	39,71	34,88	34,61	-	41,3265
21:30	35,33	36,66	32,10	30,89	-	37,2013
22:00	29,42	32,50	28,36	35,12	-	30,8666
22:30	38,17	37,76	32,83	30,39	-	31,5975
23:00	36,02	33,08	28,66	32,97	-	43,0667
23:30	40,85	36,51	31,52	25,75	-	31,2007
MASE	-	1,02	0,99	1,76	-	0,79

e ainda, o ajuste otimizado deste modelo foi realizado satisfatoriamente, pois, segundo Hyndman e Koehler (2006) um valor do MASE < 1 , indica que os dados utilizados no

ajuste do modelo de previsão, aderiram de forma eficiente ao modelo de previsão proposto. Logo, o modelo de previsão proposto neste estudo pode ser efetivamente utilizado.

Para ratificar a boa avaliação quanto à capacidade preditiva do modelo ajustado, tal como, discutido anteriormente, foram geradas previsões para um horizonte de seis passos à frente, e, analisados os resultados pelo conceito de intervalo de confiança *fuzzy* (Wu, 2009). Consultando a Tabela (6.32), se torna possível perceber que os valores das projeções para o preço de energia elétrica, estão localizados em pelo menos um dos intervalos de confiança *fuzzy* (inferiores ou/e superiores). Como os intervalos foram construídos com um altíssimo grau de confiabilidade, no caso 99%, a hipótese de um bom ajuste do modelo de previsão *fuzzy* proposto não deve ser refutada. Neste contexto, se pode garantir que o método proposto para prever o preço da energia elétrica em New South Wales na Austrália, contribui de forma significativa em um cenário de previsão de curto prazo, como uma ferramenta para auxiliar a tomada de decisão por um especialista, interessado em cotações futuras do preço da energia elétrica na região supracitada.

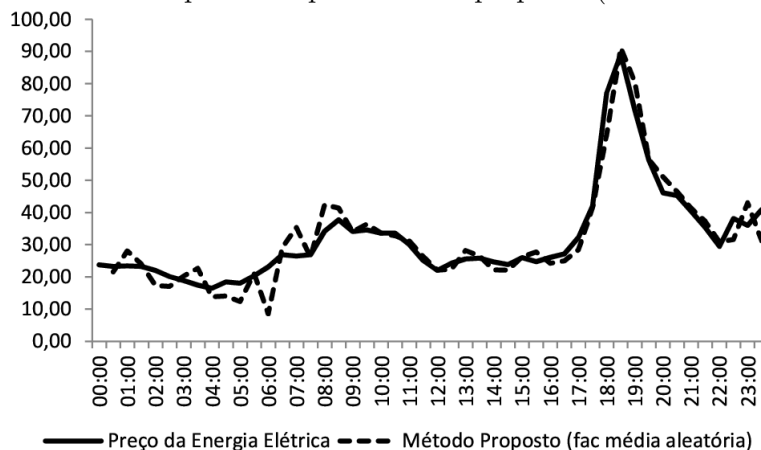
Tabela 6.32 - Resultados das projeções para o preço em dólar da energia elétrica em New South Wales na Austrália, correspondendo aos seis instantes de tempo posteriores ao período original da série histórica observada, mediante o método proposto (fac média aleatória).

Horário da Projeção	Projeção para o preço de energia	Conjunto <i>fuzzy</i>	Suporte u_i	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
00:00	35,43	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
00:30	36,53	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
01:00	37,44	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
01:30	41,34	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
02:00	41,98	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]
02:30	38,57	A_3	[20,49; 52,84]	[17,03 ; 64,37]	[41,30; 88,64]

Na Figura (6.9), o modelo de previsão *fuzzy* obtido pelo método proposto neste estudo, tem seus resultados comparados em relação aos valores reais dos preços de energia elétrica, compreendidos no espaço de trajetória utilizado para estimar o modelo FSARIMA(2;1;2)(1;0;1)₂₄. Em vários pontos na trajetória da série temporal dos preços de energia elétrica, o valor real coincide com o valor de previsão, como mostra a Figura (6.9), tal que, o padrão estrutural da série original é acompanhado ao longo de grande parte da trajetória, inclusive em um ponto de sazonalidade as 19h00min., quando se supõem a chegada dos consumidores residências em seus domicílios após um dia de trabalho, aumentando assim a demanda de energia elétrica, e, portanto elevando o preço da energia elétrica nesta região. Este fato ra-

tífica que o modelo ajustado conseguiu capturar de forma efetiva, a dependência inerente aos dados utilizados para estimação do modelo de previsão *fuzzy*, o que necessariamente potencializa a capacidade preditiva do método proposto neste estudo, em relação ao preço da energia elétrica em New South Wales na Austrália.

Figura 6.9 - Série temporal do preço em dólar da energia elétrica em New South Wales na Austrália *versus* os valores das previsões pelo método proposto (fac média aleatória).



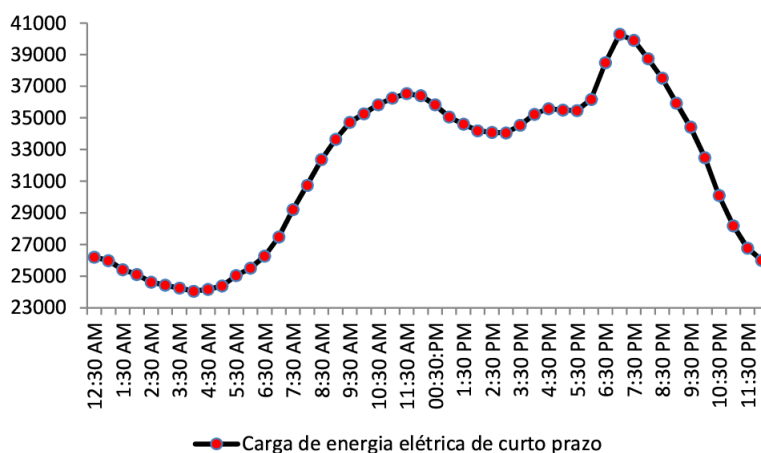
6.1.2 - Estudo de caso 2: Previsão fuzzy para a carga de energia elétrica de curto prazo.

A carga de energia elétrica que está disponível Sadaei et al. (2014) é utilizada neste estudo de caso. Mediante estes dados o desempenho do modelo de previsão será avaliado.

Passo 1: Coletar os dados de interesse que formarão a série temporal.

Passo 2: Construir o gráfico da série temporal da carga de energia elétrica de curto prazo.

Figura 6.10 - Série temporal da carga de energia elétrica de curto prazo, durante 48 instantes de tempo no período de 12:30 AM a 11:30 PM coletados a cada meia hora.



Passo 3: Definir o universo de discurso. Mediante os dados: $Z_{(Min)} = 24050$ e $Z_{(Max)} = 40292$, e ainda, $D_1 = 2710$ e $D_2 = 2708$. Portanto, o universo de discurso é igual a $U = [21340; 43000]$.

Passo 4: Particionar o universo de discurso em “ m ” conjuntos *fuzzy*.

Pela rotina de identificação na Figura (6.13), como a trajetória $T > 36$ ($T = 48$), o valor de m é igual a: 5, 7 ou 9. Segundo Miller (1994), o valor de $m = 7$ é o indicado e será adotado.

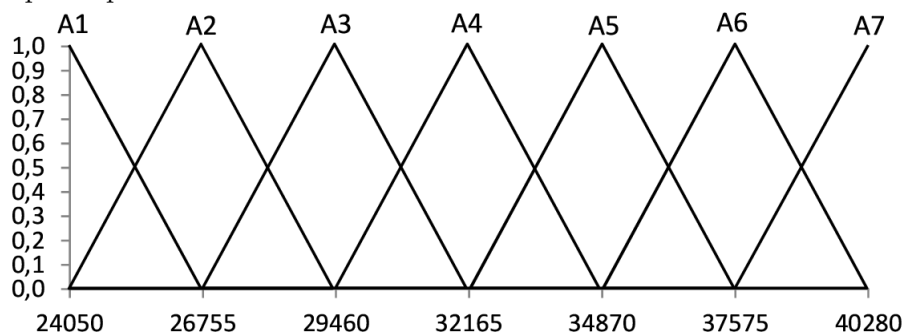
Passo 5: Construir os conjuntos *fuzzy* correspondentes a quantidade estabelecida no Passo 4 e *fuzzificar* a série histórica a partir da definição de número *fuzzy* triangular.

Seja A_j a representação do j -ésimo conjunto *fuzzy*, onde $j = 1; 2; \dots; 6; 7$. Logo, caso A_j seja um número *fuzzy* triangular, se chega que, $u_j = [a_j; b_j; c_j]$. Desta forma, os intervalos são:

$$\begin{aligned} u_1 &= [21340; 24050; 26760], u_2 = [24050; 26760; 29460], u_3 = [26760; 29460; 32170], \\ u_4 &= [29460; 32170; 34880], u_5 = [32170; 34880; 37590], u_6 = [34880; 37590; 40290] \text{ e} \\ u_7 &= [37590; 40290; 43000]. \end{aligned}$$

Cada um dos sete intervalos resultantes de $m = 7$, em relação aos conjuntos A_j (número *fuzzy* triangular), representam a base de um dos sete triângulos construídos, como mostra a Figura (6.11). Pela Equação (6.1), os valores da Tabela (6.33), referentes a pertinência máxima em relação aos conjuntos *fuzzy* A_j , se tornaram possíveis de obter. Relações *fuzzy* lógicas simples e relações agrupadas, com a série temporal dos valores *defuzzificados*, também são encontrados na Tabela (6.33). O método centróide foi adotado na *defuzzificação* da série temporal *fuzzy*.

Figura 6.11 - Conjuntos *fuzzy* para a carga de energia elétrica de curto prazo, durante 48 instantes de tempo no período de 12:30 AM a 11:30 PM coletados a cada meia hora.



Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para identificar a ordem do modelo de série temporal *fuzzy*, sobre os dados históricos de interesse.

(a) Para a função $W_{i,k,j}^{DFZ}$, se deve *defuzzificar* a série histórica mediante o método centróide, e em seguida aplicar a função de autocorrelação estendida *fuzzy* à série *defuzzificada*.

Tabela 6.33 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de *fuzzificação* mediante a pertinência máxima pelo número *fuzzy* triangular, lógica *fuzzy* simples e agrupada, série dos valores *defuzzificados* pelo método centroide em cada intervalo u_j associado.

Horário observado	Carga atual (MW)	Pertinência máxima	Intervalos <i>fuzzy</i> (u_j)	Conjuntos <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	Carga <i>defuzzificada</i>
12:30 AM	26203	0,794	[24050; 29460]	A_2	-	$A_1 \rightarrow A_1, A_2$	26769
01:00 AM	25982	0,713	[24050; 29460]	A_2	$A_2 \rightarrow A_2$	$A_2 \rightarrow A_1, A_2, A_3$	26815
01:30 AM	25411	0,502	[24050; 29460]	A_2	$A_2 \rightarrow A_2$	$A_3 \rightarrow A_2, A_3, A_4$	26719
02:00 AM	25101	0,612	[21340; 26760]	A_1	$A_2 \rightarrow A_1$	$A_4 \rightarrow A_3, A_5$	24074
02:30 AM	24621	0,789	[21340; 26760]	A_1	$A_1 \rightarrow A_1$	$A_5 \rightarrow A_4, A_5, A_6$	24005
03:00 AM	24427	0,861	[21340; 26760]	A_1	$A_1 \rightarrow A_1$	$A_6 \rightarrow A_5, A_6, A_7$	24048
03:30 AM	24247	0,927	[21340; 26760]	A_1	$A_1 \rightarrow A_1$	$A_7 \rightarrow A_6, A_7$	24046
04:00 AM	24050	1,000	[21340; 26760]	A_1	$A_1 \rightarrow A_1$		24103
04:30 AM	24156	0,961	[21340; 26760]	A_1	$A_1 \rightarrow A_1$		24097
05:00 AM	24376	0,880	[21340; 26760]	A_1	$A_1 \rightarrow A_1$		24092
05:30 AM	25035	0,637	[21340; 26760]	A_1	$A_1 \rightarrow A_1$		24077
06:00 AM	25508	0,538	[24050; 29460]	A_2	$A_1 \rightarrow A_2$		26781
06:30 AM	26259	0,815	[24050; 29460]	A_2	$A_2 \rightarrow A_2$		26758
07:00 AM	27471	0,737	[24050; 29460]	A_2	$A_2 \rightarrow A_2$		26748
07:30 AM	29213	0,909	[26760; 32170]	A_3	$A_2 \rightarrow A_3$		29457
08:00 AM	30742	0,527	[26760; 32170]	A_3	$A_3 \rightarrow A_3$		29478
08:30 AM	32369	0,927	[29460; 34880]	A_4	$A_3 \rightarrow A_4$		32190
09:00 AM	33649	0,546	[32170; 37590]	A_5	$A_4 \rightarrow A_5$		34924
09:30 AM	34724	0,942	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34835
10:00 AM	35273	0,855	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34872
10:30 AM	35838	0,646	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34879
11:00 AM	36256	0,508	[34880; 40290]	A_6	$A_5 \rightarrow A_6$		37518
11:30 AM	36542	0,613	[34880; 40290]	A_6	$A_6 \rightarrow A_6$		37617
12:00 AM	36406	0,563	[34880; 40290]	A_6	$A_6 \rightarrow A_6$		37501
00:30 PM	35836	0,647	[32170; 37590]	A_5	$A_6 \rightarrow A_5$		34890
01:00 PM	35056	0,935	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34908
01:30 PM	34602	0,897	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34905
02:00 PM	34191	0,746	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34839
02:30 PM	34084	0,706	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34918
03:00 PM	34054	0,695	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34895
03:30 PM	34535	0,873	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34850
04:00 PM	35237	0,868	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34883
04:30 PM	35592	0,737	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34878
05:00 PM	35492	0,774	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34884
05:30 PM	35470	0,782	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34919
06:00 PM	36177	0,521	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34920
06:30 PM	38499	0,663	[34880; 40290]	A_6	$A_5 \rightarrow A_6$		37585
07:00 PM	40292	0,999	[37590; 43000]	A_7	$A_6 \rightarrow A_7$		40251
07:30 PM	39902	0,856	[37590; 43000]	A_7	$A_7 \rightarrow A_7$		40340
08:00 PM	38744	0,573	[34880; 40290]	A_6	$A_7 \rightarrow A_6$		37620
08:30 PM	37514	0,972	[34880; 40290]	A_6	$A_6 \rightarrow A_6$		37584
09:00 PM	35931	0,612	[32170; 37590]	A_5	$A_6 \rightarrow A_5$		34882
09:30 PM	34419	0,830	[32170; 37590]	A_5	$A_5 \rightarrow A_5$		34870
10:00 PM	32490	0,882	[29460; 34880]	A_4	$A_5 \rightarrow A_4$		32170
10:30 PM	30096	0,765	[26760; 32170]	A_3	$A_4 \rightarrow A_3$		29475
11:00 PM	28188	0,529	[26760; 32170]	A_3	$A_3 \rightarrow A_3$		29480
11:30 PM	26776	0,994	[24050; 29460]	A_2	$A_3 \rightarrow A_2$		26768
12:00 PM	25997	0,718	[24050; 29460]	A_2	$A_2 \rightarrow A_2$		26738

Antes de iniciar o processo de identificação do modelo de série temporal, mediante uma função de autocorrelação, a comprovação da suposição de estacionariedade da série, se faz extremamente necessária. Neste contexto, o teste da raiz unitária de Dickey-Fuller, atende plenamente a esta condição. A Tabela (6.34) indica que a série temporal da carga de energia elétrica de curto prazo é estacionária, pois, $p\text{-valor} = 0,001 < \alpha = 1\%$. Ratificando o resultado obtido anteriormente, ao diferenciar a série temporal em 1 *lag*, isto

é, quando $n = 46$, passa a ser $n = 45$, como verificado na Tabela (6.34) e Tabela (6.35), respectivamente, o resultado da estatística de teste apresenta $p\text{-valor} = 0,000 < \alpha = 1\%$. Portanto, a série temporal é estacionária e a função de identificação para o modelo, neste caso, a função de autocorrelação estendida pode ser utilizada.

Após comprovar a estacionariedade da série temporal da carga de energia elétrica, a função de identificação foi utilizada. Consultando as Tabelas (6.36) e (6.37), se pode perceber que a função $W_{i,k,j}^{DF_Z}$ identificou de forma significativa, um modelo ARMA(5; 2), como de maior termo para p e q . Isto é, quando $p = 5$ e $q = 2$, o $p\text{-valor} = 0,028 < \alpha = 10\%$, logo, a função de identificação retornou um modelo de série temporal, com parâmetros de ordem razoavelmente elevada.

Tabela 6.34 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	Tau	P-valor
1	OLS	1,082	-4,93836	46	-4,440	0,001*
2	COV	1,082	1,23105	46	-4,440	0,001*
3	COV	1,082	0,19588	46	-4,440	0,001*
4	COV	1,082	-0,09454	46	-4,440	0,001*

Tabela 6.35 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	Tau	$p\text{-valor}$
1	OLS	1,556	-0,059544	45	-6,51192	0,000*
2	COV	1,556	0,034705	45	-6,51192	0,000*
3	COV	1,556	0,002701	45	-6,51192	0,000*
4	COV	1,556	-0,001579	45	-6,51192	0,000*

Nota: * Valor significativo para $\alpha = 1\%$.

Tabela 6.36 - Resultados da função $W_{i,k,j}^{DF_Z}$ para a carga de energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,225	0,213	0,150	0,069	0,217	0,003
AR(1)	-0,492	0,048	0,040	-0,036	0,218	0,034
AR(2)	-0,372	0,171	0,027	-0,017	0,198	-0,046
AR(3)	-0,011	-0,079	-0,236	0,029	0,156	-0,005
AR(4)	0,012	-0,326	-0,301	0,039	0,149	-0,018
AR(5)	0,396	0,435	0,377	0,141	0,232	-0,066

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6 (a);

Os resultados da Tabela (6.37) indicam um modelo com os termos $p \leq 5$ e $q \leq 2$. Neste contexto, todas as combinações de parâmetros iguais e inferiores à identificação, foram testadas e otimizadas computacionalmente. Após a identificação do modelo de série temporal *fuzzy* com o auxílio da função $W_{i,k,j}^{DF_Z}$, o modelo FSARIMA(4,1,1)(1,0,1)₃₆ foi

Tabela 6.37 - Probabilidade de significância para os resultados apresentados na Tabela 6.36.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,122	0,165	0,345	0,670	0,184	0,984
AR(1)	0,001*	0,784	0,811	0,815	0,172	0,863
AR(2)	0,013*	0,391	0,871	0,921	0,254	0,781
AR(3)	0,943	0,619	0,123	0,866	0,370	0,978
AR(4)	0,940	0,033*	0,053*	0,811	0,407	0,926
AR(5)	0,010*	0,017*	0,028*	0,445	0,201	0,765

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

ajustado mediante o menor erro de previsão retornado. É importante destacar, que o parâmetros AR(1) com $p\text{-valor} = 0,259$, não foi significativo, isto é, este parâmetro foi excluído do modelo quando se realizaram as previsões, pois é estatisticamente igual a zero, ao nível de significância $\alpha = 1\%$ (ver, Tabela 6.38). Como se trata de um modelo com o termo de sazonalidade integrado, o teste de Ljung-Box é necessário. A Tabela (6.40) indica que o termo sazonal de ordem $s = 36$ não é significativo (ruído branco), pois, o $p\text{-valor} = 0,037 < \alpha = 1\%$. Assim, o termo de sazonalidade será mantido no modelo, devido corroborar para o aumento da acurácia das previsões para carga de energia elétrica.

Tabela 6.38 - Testes das estimativas dos parâmetros ao modelo FSARIMA(4,1,1)(1,0,1)₃₆ ajustado a carga de energia elétrica.

Modelo	Coefficiente	Erro padrão	Estatística t	P-valor
AR(2)	0,646	0,125	5,150	0,000*
AR(3)	0,865	0,072	11,940	0,000*
AR(4)	-0,633	0,131	-4,850	0,000*
AR(5)	-0,516	0,172	-2,990	0,005*
SAR(36)	1,000	0,021	47,910	0,000*
MA(1)	0,990	0,049	20,070	0,000*
SMA(36)	-0,712	0,244	-2,910	0,006*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 1\%$.

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

Deve-se avaliar a capacidade preditiva do modelo FSARIMA(4,1,1)(1,0,1)₃₆, o qual foi identificado pela função de autocorrelação estendida *fuzzy*. Os resultados das previsões realizadas pelo modelo, a partir do método proposto aos dados *defuzzificados*, foram comparados a outros cinco métodos de previsão *fuzzy*, como mostra a Tabela (6.39). Utilizando a métrica de desempenho MASE, o modelo 1 que está sendo proposto alcançou valor absoluto de 0,87. Portanto, dentre os métodos propostos por Sadaei et al. (2014), apenas o método OTS-SVR, com valor do MASE igual a 1,09, demonstrou acurácia inferior ao método com dados *defuzzificados*, que foi desenvolvido neste trabalho.

Tabela 6.39 - Comparação das previsões pelo método proposto (dados *defuzzificados*) em relação aos cinco métodos propostos por Sadaei et al. (2014) para a carga de energia elétrica.

Instante de tempo observado (horas)	Carga de energia elétrica (MW)	Método OTS-SVR (2014)	Método PSO-ARIMA (2014)	Método LS-ARIMA (2014)	Método REWFTS-Model 1 (2014)	Método Sadaei et al. (2014)	Método proposto: Treinamento	Método proposto: Validação
12:30 AM	26.203	25.459	27.360	26.270	26.606	26.542	-	-
01:00 AM	25.982	24.730	25.296	26.370	25.572	26.205	26.766,4	-
01:30 AM	25.411	24.470	25.032	24.875	24.906	25.179	26.806,7	-
02:00 AM	25.101	26.415	26.038	24.773	24.609	24.784	23.933,5	-
02:30 AM	24.621	23.451	23.943	24.031	24.650	24.579	24.303,6	-
03:00 AM	24.427	24.915	23.388	24.034	24.646	24.011	24.837,8	-
03:30 AM	24.247	23.429	23.596	24.350	24.643	24.020	23.396,0	-
04:00 AM	24.050	24.570	24.648	23.685	24.646	24.101	24.070,2	-
04:30 AM	24.156	23.311	25.129	23.723	24.644	24.583	24.043,0	-
05:00 AM	24.376	23.122	24.993	24.201	24.647	24.581	23.359,4	-
05:30 AM	25.035	24.585	24.413	24.599	24.612	25.349	23.780,8	-
06:00 AM	25.508	25.980	24.879	25.897	24.908	25.749	26.806,6	-
06:30 AM	26.259	25.050	26.132	26.408	26.603	26.520	26.784,2	-
07:00 AM	27.471	26.270	28.551	27.941	26.999	27.069	26.747,5	-
07:30 AM	29.213	29.430	28.093	28.544	28.606	29.492	29.473,0	-
08:00 AM	30.742	30.232	31.279	30.944	30.049	30.426	29.488,9	-
08:30 AM	32.369	32.699	32.494	32.062	32.724	32.086	32.179,5	-
09:00 AM	33.649	32.958	32.953	33.107	34.047	33.981	34.935,2	-
09:30 AM	34.724	33.870	34.375	34.201	34.981	34.351	34.834,5	-
10:00 AM	35.273	35.720	36.354	35.850	35.911	35.749	34.857,2	-
10:30 AM	35.838	35.206	36.327	36.317	36.217	36.148	34.890,3	-
11:00 AM	36.256	36.909	36.152	36.523	36.600	35.960	37.519,2	-
11:30 AM	36.542	35.409	37.220	36.898	36.181	36.839	37.618,3	-
12:00 AM	36.406	38.190	36.606	36.121	36.878	36.815	37.527,6	-
00:30 PM	35.836	34.987	36.550	35.243	35.518	36.147	34.895,7	-
01:00 PM	35.056	36.056	35.761	34.839	35.470	34.780	34.924,0	-
01:30 PM	34.602	33.527	34.986	34.315	34.979	34.349	34.944,2	-
02:00 PM	34.191	35.055	34.141	33.658	35.469	33.725	34.858,0	-
02:30 PM	34.084	35.147	33.672	34.320	35.466	33.801	34.949,7	-
03:00 PM	34.054	32.976	33.152	34.778	35.472	34.028	34.935,5	-
03:30 PM	34.535	35.384	34.966	35.214	34.978	34.889	34.865,7	-
04:00 PM	35.237	35.107	34.231	35.654	35.917	35.180	34.913,9	-
04:30 PM	35.592	35.128	35.959	35.292	35.917	35.123	34.908,1	-
05:00 PM	35.492	34.653	36.596	35.707	35.908	35.222	34.886,4	-
05:30 PM	35.470	36.569	35.286	36.073	35.914	35.100	34.944,6	-
06:00 PM	36.177	35.416	36.924	35.461	35.900	36.538	34.923,4	-
06:30 PM	38.499	37.859	38.540	38.301	38.122	38.760	37.591,7	-
07:00 PM	40.292	41.663	41.030	39.676	39.831	39.896	40.240,1	-
07:30 PM	39.902	40.521	39.133	40.412	40.286	40.213	40.388,2	-
08:00 PM	38.744	37.910	38.437	38.297	38.969	38.186	37.505,3	-
08:30 PM	37.514	36.608	37.703	37.152	37.070	37.217	37.784,3	-
09:00 PM	35.931	37.363	35.696	35.104	35.346	35.579	35.415,9	-
09:30 PM	34.419	33.510	34.926	34.211	34.980	34.310	-	34.374,4
10:00 PM	32.490	31.445	32.043	32.970	32.842	32.208	-	32.206,9
10:30 PM	30.096	28.739	30.303	30.165	30.633	30.388	-	29.307,2
11:00 PM	28.188	29.300	28.515	27.693	27.564	27.917	-	29.186,4
11:30 PM	26.776	27.732	25.788	27.249	27.010	26.379	-	26.172,7
12:00 PM	25.997	25.161	26.790	26.365	25.570	26.200	-	27.553,0
MASE		1,09	0,71	0,51	0,60	0,37	-	0,87

Desta forma, o desempenho do método 1, proposto mediante os dados *defuzzificados* para previsão da carga de energia elétrica, pode ser considerável razoável, em comparação aos métodos de previsão propostos por Sadaei et al. (2014).

Tabela 6.40 - Teste Box-Pierce modificado (Ljung-Box) pela estatística Qui-quadrado ao modelo FSARIMA(4,1,1)(1,0,1)₃₆ ajustado para carga de energia elétrica.

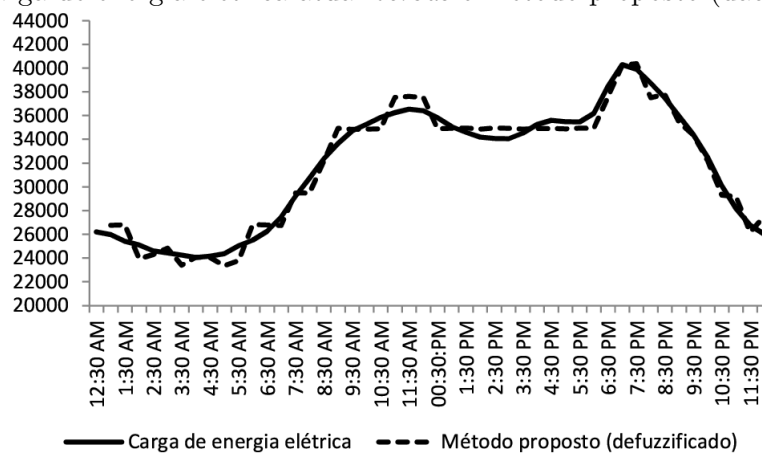
Lag	Chi-square	Degree freedom	P-value
12	11,400	04	0,022*
24	11,500	16	0,774
36	42,800	28	0,037
48	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 1\%$.

Adicionalmente, o conceito de intervalo de confiança *fuzzy*, foi utilizado para avaliar seis projeções da carga de energia elétrica, como mostra a Tabela (6.41). As seis projeções realizadas pelo modelo FSARIMA(4,1,1)(1,0,1)₃₆, retornaram valores dentro dos intervalos inferiores e/ou superiores, que foram construídos com 99% de confiança. Este fato indica que o modelo foi muito bem identificado, e ainda, que os dados *defuzzificados* se ajustaram perfeitamente ao modelo de previsão.

Tabela 6.41 - Intervalo de confiança *fuzzy* para as 6 projeções de carga de energia elétrica pelo método proposto (dados *defuzzificados*).

Horário projetado	Projeção da carga de energia elétrica (MW)	Conjunto <i>fuzzy</i>	Interval <i>fuzzy</i> (u_j)	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
12:30 PM	28.372,00	A3	[26.760; 32.170]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
01:00 AM	25.991,70	A2	[24.050; 29.460]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
01:30 AM	28.626,50	A3	[26.760; 32.170]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
02:00 AM	31.660,80	A4	[29.460; 34.880]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
02:30 AM	33.249,50	A4	[29.460; 34.880]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
03:00 AM	38.271,30	A6	[34.880; 40.290]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]

Figura 6.12 - Carga de energia elétrica atual *versus* o método proposto (dados *defuzzificados*).

Na Figura (6.12) é possível perceber que os valores das previsões realizadas pelo modelo proposto, na grande maioria acompanham de forma muito próxima, os pontos reais na trajetória estudada, indicando um bom ajuste dos dados ao modelo proposto.

Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para identificar o modelo de série temporal *fuzzy*, sobre os dados históricos de interesse.

Passo 6(b): (b1) Pela função $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$, obter a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy* A_i . Depois, aplicar a função de autocorrelação estendida.

Tabela 6.42 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de *fuzzificação* mediante a definição de número *fuzzy* triangular, lógica *fuzzy* simples e agrupada, série da carga de energia elétrica com α -nível = 0,25 a partir de cada intervalo u_j associado.

Horário observado	Carga atual (MW)	Intervalos <i>fuzzy</i> (u_j)	Conjuntos <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	Série $\alpha = 0,25$
12:30 AM	26.203	[24.050; 29.460]	A_2	-	$A_1 \rightarrow A_1, A_2$	0,000661
01:00 AM	25.982	[21.340; 26.760]	A_1	$A_2 \rightarrow A_1$	$A_2 \rightarrow A_1, A_2, A_3$	0,004405
01:30 AM	25.411	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_3 \rightarrow A_2, A_3, A_4$	0,000556
02:00 AM	25.101	[24.050; 29.460]	A_2	$A_1 \rightarrow A_2$	$A_4 \rightarrow A_3, A_4, A_5$	0,001490
02:30 AM	24.621	[21.340; 26.760]	A_1	$A_2 \rightarrow A_1$	$A_5 \rightarrow A_4, A_5, A_6$	0,000826
03:00 AM	24.427	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_6 \rightarrow A_5, A_6, A_7$	0,000393
03:30 AM	24.247	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_7 \rightarrow A_6, A_7$	0,003132
04:00 AM	24.050	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,002051
04:30 AM	24.156	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,004222
05:00 AM	24.376	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,000352
05:30 AM	25.035	[24.050; 29.460]	A_2	$A_1 \rightarrow A_2$		0,000196
06:00 AM	25.508	[21.340; 26.760]	A_1	$A_2 \rightarrow A_1$		0,002403
06:30 AM	26.259	[24.050; 29.460]	A_2	$A_1 \rightarrow A_2$		0,006171
07:00 AM	27.471	[26.760; 32.170]	A_3	$A_2 \rightarrow A_3$		0,004200
07:30 AM	29.213	[26.760; 32.170]	A_3	$A_3 \rightarrow A_3$		0,001517
08:00 AM	30.742	[29.460; 34.880]	A_4	$A_3 \rightarrow A_4$		0,005730
08:30 AM	32.369	[29.460; 34.880]	A_4	$A_4 \rightarrow A_4$		0,000314
09:00 AM	33.649	[29.460; 34.880]	A_4	$A_4 \rightarrow A_4$		0,002553
09:30 AM	34.724	[32.170; 37.590]	A_5	$A_4 \rightarrow A_5$		0,007873
10:00 AM	35.273	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003829
10:30 AM	35.838	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,000579
11:00 AM	36.256	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,003958
11:30 AM	36.542	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,000869
12:00 AM	36.406	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,002056
00:30 PM	35.836	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003787
01:00 PM	35.056	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000598
01:30 PM	34.602	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001446
02:00 PM	34.191	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001266
02:30 PM	34.084	[29.460; 34.880]	A_4	$A_5 \rightarrow A_4$		0,002856
03:00 PM	34.054	[29.460; 34.880]	A_4	$A_4 \rightarrow A_4$		0,002787
03:30 PM	34.535	[32.170; 37.590]	A_5	$A_4 \rightarrow A_5$		0,000273
04:00 PM	35.237	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001760
04:30 PM	35.592	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003517
05:00 PM	35.492	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003521
05:30 PM	35.470	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,007777
06:00 PM	36.177	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,010407
06:30 PM	38.499	[34.880; 40.290]	A_7	$A_6 \rightarrow A_7$		0,001195
07:00 PM	40.292	[37.590; 43.000]	A_7	$A_7 \rightarrow A_7$		0,002221
07:30 PM	39.902	[37.590; 43.000]	A_7	$A_7 \rightarrow A_7$		0,001189
08:00 PM	38.744	[37.590; 43.000]	A_7	$A_7 \rightarrow A_7$		0,003481
08:30 PM	37.514	[34.880; 40.290]	A_6	$A_7 \rightarrow A_6$		0,001331
09:00 PM	35.931	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,006770
09:30 PM	34.419	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,004512
10:00 PM	32.490	[29.460; 34.880]	A_4	$A_5 \rightarrow A_4$		0,001218
10:30 PM	30.096	[26.760; 32.170]	A_3	$A_4 \rightarrow A_3$		0,002044
11:00 PM	28.188	[24.050; 29.460]	A_2	$A_3 \rightarrow A_2$		0,003077
11:30 PM	26.776	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,005167
12:00 PM	25.997	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,005521

Figura 6.13 - Conjuntos construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_i e o eixo X a carga de energia elétrica, com α -nível = 0,25 para ativar o suporte dos conjuntos *fuzzy* correspondentes.

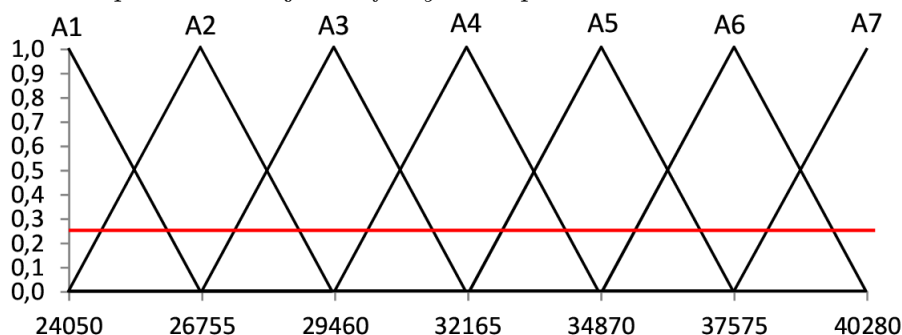


Tabela 6.43 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$ para a carga de energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,424	-0,084	-0,003	-0,100	0,048	0,133
AR(1)	-0,487	-0,083	0,006	-0,101	0,019	0,070
AR(2)	-0,500	-0,096	0,112	-0,087	-0,001	0,028
AR(3)	-0,488	-0,543	-0,286	-0,033	-0,001	0,137
AR(4)	-0,546	0,231	-0,088	0,025	-0,057	0,079
AR(5)	-0,459	0,252	-0,153	-0,315	-0,032	0,126

Tabela 6.44 - Probabilidade de significância dos resultados da Tabela (6.43).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,004*	0,620	0,985	0,559	0,780	0,439
AR(1)	0,001*	0,628	0,973	0,560	0,918	0,718
AR(2)	0,001*	0,582	0,537	0,677	0,998	0,901
AR(3)	0,001*	0,000*	0,075*	0,881	0,997	0,508
AR(4)	0,000*	0,285	0,715	0,917	0,830	0,786
AR(5)	0,003*	0,251	0,537	0,112	0,890	0,688

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Verificando os resultados das Tabelas (6.43 e 6.44), a função de autocorrelação estendida *fuzzy* quando α -nível é igual a 0,25, identificou de maneira significativa a 10%, desde que $p > 0$ e $q > 0$, um modelo ARMA(3; 2). Portanto, todos os modelos de série temporal com combinações paramétricas, desde que, satisfaçam a condição $0 < p \leq 3$ e $0 < q \leq 2$, serão avaliados na otimização computacional do processo de modelagem estatística, que irá ajustar o modelo de previsão *fuzzy*, com o menor erro de previsão possível para carga de energia elétrica.

Na Tabela (6.42) é possível perceber que o limite mínimo α para ativação dos conjuntos *fuzzy* A_i , que neste caso foi de 0,25 (ver, Figura 6.13), não impediu durante o processo de *fuzzificação* uma transição completa, entre os suportes correspondentes as partições de

U , as quais foram implementadas para a carga de energia elétrica. A série temporal com α -nível = 0,25, que foi obtida nesta rotina possibilitará identificar o modelo por $W_{i,k,j}^{CA}$.

Tabela 6.45 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de *fuzzificação* mediante a definição de número *fuzzy* triangular, lógica *fuzzy* simples e agrupada, série da carga de energia elétrica com α -nível = 0,50 a partir de cada intervalo u_j associado.

Horário observado	Carga atual (MW)	Intervalos <i>fuzzy</i> (u_j)	Conjuntos <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	Série $\alpha = 0,50$
12:30 AM	26.203	[24.050; 29.460]	A_2	-	$A_1 \rightarrow A_1, A_2$	0,003730
01:00 AM	25.982	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$	$A_2 \rightarrow A_1, A_2, A_3$	0,000000
01:30 AM	25.411	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$	$A_3 \rightarrow A_2, A_3, A_4$	0,001686
02:00 AM	25.101	[21.340; 26.760]	A_1	$A_2 \rightarrow A_1$	$A_4 \rightarrow A_3, A_5$	0,002970
02:30 AM	24.621	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_5 \rightarrow A_4, A_5, A_6$	0,006454
03:00 AM	24.427	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_6 \rightarrow A_5, A_6, A_7$	0,004289
03:30 AM	24.247	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_7 \rightarrow A_6, A_7$	0,003260
04:00 AM	24.050	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,001795
04:30 AM	24.156	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,007245
05:00 AM	24.376	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,007434
05:30 AM	25.035	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,000225
06:00 AM	25.508	[24.050; 29.460]	A_2	$A_1 \rightarrow A_2$		0,002385
06:30 AM	26.259	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,004927
07:00 AM	27.471	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,002591
07:30 AM	29.213	[26.760; 32.170]	A_3	$A_2 \rightarrow A_3$		0,004232
08:00 AM	30.742	[26.760; 32.170]	A_3	$A_3 \rightarrow A_3$		0,003625
08:30 AM	32.369	[29.460; 34.880]	A_4	$A_3 \rightarrow A_4$		0,001413
09:00 AM	33.649	[32.170; 37.590]	A_5	$A_4 \rightarrow A_5$		0,000162
09:30 AM	34.724	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003630
10:00 AM	35.273	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000340
10:30 AM	35.838	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,005977
11:00 AM	36.256	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,003514
11:30 AM	36.542	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,000505
12:00 AM	36.406	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,000836
00:30 PM	35.836	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,002061
01:00 PM	35.056	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001447
01:30 PM	34.602	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,006994
02:00 PM	34.191	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,002016
02:30 PM	34.084	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000923
03:00 PM	34.054	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001892
03:30 PM	34.535	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,011722
04:00 PM	35.237	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000115
04:30 PM	35.592	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,002921
05:00 PM	35.492	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,006327
05:30 PM	35.470	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000671
06:00 PM	36.177	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003053
06:30 PM	38.499	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,000994
07:00 PM	40.292	[37.590; 43.000]	A_7	$A_6 \rightarrow A_7$		0,005532
07:30 PM	39.902	[37.590; 43.000]	A_7	$A_7 \rightarrow A_7$		0,003235
08:00 PM	38.744	[34.880; 40.290]	A_6	$A_7 \rightarrow A_6$		0,002271
08:30 PM	37.514	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,006429
09:00 PM	35.931	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,001556
09:30 PM	34.419	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,005016
10:00 PM	32.490	[29.460; 34.880]	A_4	$A_5 \rightarrow A_4$		0,001001
10:30 PM	30.096	[26.760; 32.170]	A_3	$A_4 \rightarrow A_3$		0,000488
11:00 PM	28.188	[26.760; 32.170]	A_3	$A_3 \rightarrow A_3$		0,003163
11:30 PM	26.776	[24.050; 29.460]	A_2	$A_3 \rightarrow A_2$		0,004958
12:00 PM	25.997	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,003042

Passo 6(b): (b2) Para a função $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$, se deve obter a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy*, com alfa-nível pré-estabelecido e em seguida aplicar a função de autocorrelação estendida *fuzzy* na série.

Tal como, já havia sido notado no processo de *fuzzificação* da série temporal da carga de energia elétrica, quando α -nível é igual a 0,25. Na Tabela (6.45) onde o limite mínimo para ativação dos conjuntos *fuzzy* passa a ser 0,50, a transição dos conjuntos sobre U também é completa. O aumento do nível mínimo para ativação dos conjuntos, não influenciou na ocorrência dos mesmos, garantindo assim uma homogeneidade na distribuição dos suportes sobre U , como se pode verificar também na Figura (6.14). A partir da série obtida com α -nível = 0,50, o processo de identificação do modelo de série temporal será realizado pela função de autocorrelação estendida.

Figura 6.14 - Conjuntos construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_j e o eixo X a carga de energia elétrica, com α -nível = 0,50 para ativar o suporte dos conjuntos *fuzzy* correspondentes.

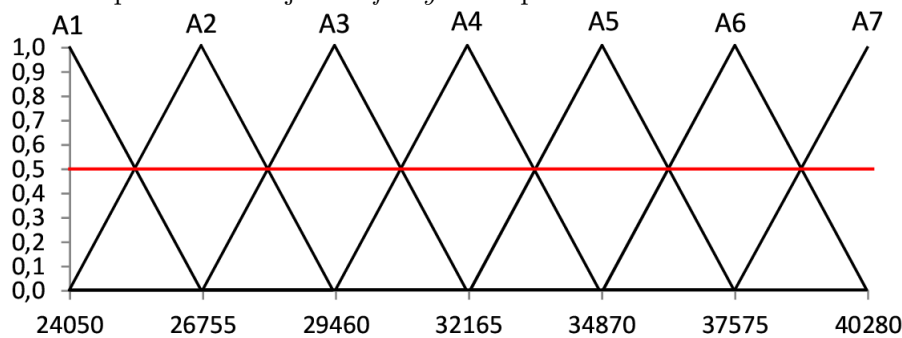


Tabela 6.46 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,50; b)$ para a carga de energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,405	0,009	-0,164	0,174	-0,026	-0,114
AR(1)	-0,379	-0,027	-0,125	0,172	-0,022	-0,035
AR(2)	-0,452	-0,442	-0,366	0,029	-0,075	-0,011
AR(3)	-0,136	0,035	-0,393	-0,085	-0,106	-0,228
AR(4)	0,090	-0,011	-0,171	-0,094	0,033	-0,214
AR(5)	0,023	-0,171	-0,136	-0,313	-0,045	-0,079

Tabela 6.47 - Probabilidade de significância dos resultados da Tabela (6.46).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,006*	0,956	0,328	0,309	0,882	0,515
AR(1)	0,010*	0,870	0,481	0,301	0,909	0,868
AR(2)	0,002*	0,003*	0,020*	0,865	0,765	0,965
AR(3)	0,368	0,820	0,012*	0,643	0,618	0,196
AR(4)	0,554	0,946	0,460	0,678	0,882	0,260
AR(5)	0,880	0,308	0,592	0,086*	0,826	0,743

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

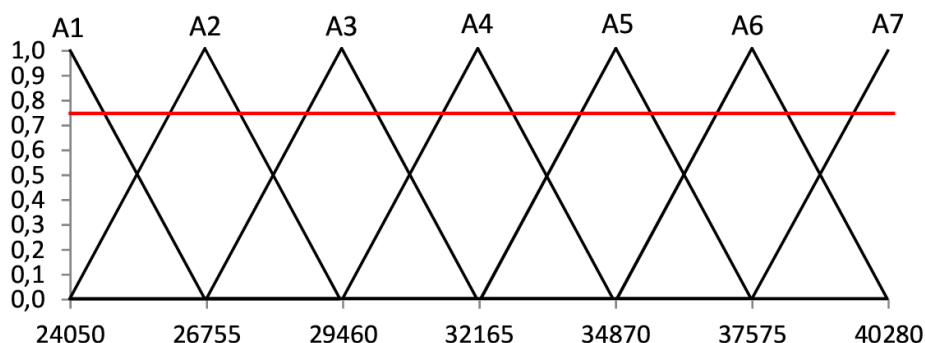
De acordo com os resultados da Tabela (6.46 e 6.47), o modelo de maior ordem paramétrica identificado pela função de autocorrelação estendida *fuzzy*, tal que, $p > 0$ e $q > 0$, é um ARMA(5; 3). O aumento do limite mínimo para ativação dos conjuntos

fuzzy, neste caso, com α -nível = 0,50, em relação a identificação anterior (α -nível = 0,25), alterou a sequencia de ativação dos conjuntos verificada com um nível mais baixo. Este fato, influenciou diretamente no processo de identificação do modelo de série temporal, como pode ser comprovado com os resultados da Tabela (6.47). Neste contexto, a Tabela (6.44) indica um modelo ARMA(3; 2), com um α -nível = 0,25, ao contrário da Tabela (6.47), que identificou um modelo ARMA(5;3), porém, a um limite mínimo de ativação dos conjuntos *fuzzy* sobre o universo de discurso com α -nível = 0,50.

Passo 6(b): (b3) Para a função $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$, se deve obter a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy*, com alfa-nível pré-estabelecido e em seguida aplicar a função de autocorrelação estendida *fuzzy* na série.

No processo de *fuzzificação* com α -nível igual a 0,75, para série temporal da carga de energia elétrica, que é apresentado na Tabela (6.48). Novamente se observa que todos os conjuntos *fuzzy* construídos sobre o universo de U , foram ativados mesmo aumentando o nível mínimo para aderência dos valores da série sobre estes conjuntos. Neste contexto, a homogeneidade na distribuição dos conjunto A_j em relação a U , como se comprova também na Figura (6.15), foi plenamente preservada comparando a α -nível = 0,25 e 0,50.

Figura 6.15 - Conjuntos construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_j e o eixo X a carga de energia elétrica, com α -nível = 0,75 para ativar o suporte dos conjuntos *fuzzy* correspondentes.



Os resultados das Tabelas (6.49 e 6.50), indicam a identificação pela função de autocorrelação estendida *fuzzy*, adotando no processo de *fuzzificação* um α -nível de 0,75. Neste processo, o modelo identificado foi um ARMA(5; 3) para uma significância estatística de 10%. O ajuste do modelo de previsão *fuzzy* irá considerar todas as combinações paramétricas, que satisfaçam a restrição $0 < p \leq 5$ e $0 < q \leq 3$, durante a otimização computacional no processo de modelagem.

Tabela 6.48 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de *fuzzificação* mediante a definição de número *fuzzy* triangular, lógica *fuzzy* simples e agrupada, série da carga de energia elétrica com α -nível = 0,75 a partir de cada intervalo u_j associado.

Horário observado	Carga atual (MW)	Intervalos <i>fuzzy</i> (u_j)	Conjuntos <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	Série $\alpha = 0,75$
12:30 AM	26.203	[24.050; 29.460]	A_2	-	$A_1 \rightarrow A_1, A_2$	0,002930
01:00 AM	25.982	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$	$A_2 \rightarrow A_1, A_2, A_3$	0,000163
01:30 AM	25.411	[24.050; 29.460]	A_1	$A_2 \rightarrow A_1$	$A_3 \rightarrow A_2, A_3, A_4$	0,000713
02:00 AM	25.101	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_4 \rightarrow A_3, A_5$	0,001350
02:30 AM	24.621	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_5 \rightarrow A_4, A_5, A_6$	0,002715
03:00 AM	24.427	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_6 \rightarrow A_5, A_6, A_7$	0,002935
03:30 AM	24.247	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_7 \rightarrow A_6, A_7$	0,004778
04:00 AM	24.050	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,005250
04:30 AM	24.156	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,001478
05:00 AM	24.376	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,003679
05:30 AM	25.035	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,003095
06:00 AM	25.508	[24.050; 29.460]	A_2	$A_1 \rightarrow A_2$		0,000118
06:30 AM	26.259	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,000323
07:00 AM	27.471	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,003998
07:30 AM	29.213	[26.760; 32.170]	A_3	$A_2 \rightarrow A_3$		0,000881
08:00 AM	30.742	[26.760; 32.170]	A_3	$A_3 \rightarrow A_3$		0,003933
08:30 AM	32.369	[29.460; 34.880]	A_4	$A_3 \rightarrow A_4$		0,004552
09:00 AM	33.649	[32.170; 37.590]	A_5	$A_4 \rightarrow A_5$		0,004944
09:30 AM	34.724	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,006955
10:00 AM	35.273	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,002403
10:30 AM	35.838	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001698
11:00 AM	36.256	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,002282
11:30 AM	36.542	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,003390
12:00 AM	36.406	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,001787
00:30 PM	35.836	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,000953
01:00 PM	35.056	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003707
01:30 PM	34.602	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,004357
02:00 PM	34.191	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,008017
02:30 PM	34.084	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,004388
03:00 PM	34.054	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003022
03:30 PM	34.535	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,005452
04:00 PM	35.237	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000503
04:30 PM	35.592	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,004561
05:00 PM	35.492	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001682
05:30 PM	35.470	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003858
06:00 PM	36.177	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003884
06:30 PM	38.499	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,002343
07:00 PM	40.292	[37.590; 43.000]	A_7	$A_6 \rightarrow A_7$		0,002761
07:30 PM	39.902	[37.590; 43.000]	A_7	$A_7 \rightarrow A_7$		0,000352
08:00 PM	38.744	[34.880; 40.290]	A_6	$A_7 \rightarrow A_6$		0,001443
08:30 PM	37.514	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,004613
09:00 PM	35.931	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,004099
09:30 PM	34.419	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000411
10:00 PM	32.490	[29.460; 34.880]	A_4	$A_5 \rightarrow A_4$		0,000718
10:30 PM	30.096	[26.760; 32.170]	A_3	$A_4 \rightarrow A_3$		0,000453
11:00 PM	28.188	[26.760; 32.170]	A_3	$A_3 \rightarrow A_3$		0,004048
11:30 PM	26.776	[24.050; 29.460]	A_2	$A_3 \rightarrow A_2$		0,004027
12:00 PM	25.997	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,006451

Todavia, como a identificação obtida com α -nível = 0,25, retornou um modelo ARMA(3; 2), este está plenamente contido no modelo ARMA(5; 3) identificado com α -nível = 0,50 e 0,75. Neste último caso, uma convergência para mesma identificação pode ser percebida, visto que, o modelo resultante deste processo é exatamente o mesmo nas duas situações

($\alpha = 0,50$ e $0,75$). O modelo de previsão *fuzzy* ajustado a carga é um ARMA(5; 3).

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6(b);

A rotina de otimização do modelo de previsão *fuzzy* para a carga de energia elétrica, que buscou-se o menor de erro de previsão, ajustou um modelo FSARIMA(5;1;1)(1;0;1)₃₆.

Tabela 6.49 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$ para a carga de energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,361	-0,022	-0,007	-0,142	0,090	-0,133
AR(1)	-0,404	0,106	0,023	-0,145	-0,019	-0,040
AR(2)	-0,457	0,028	0,034	-0,089	0,037	-0,022
AR(3)	-0,401	-0,452	-0,204	-0,092	0,050	-0,049
AR(4)	-0,395	-0,537	0,173	-0,047	0,166	-0,068
AR(5)	-0,413	-0,535	-0,149	-0,447	-0,121	-0,142

Tabela 6.50 - Probabilidade de significância dos resultados da Tabela (6.49).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,013*	0,892	0,966	0,386	0,591	0,428
AR(1)	0,006*	0,584	0,896	0,361	0,908	0,813
AR(2)	0,002*	0,875	0,846	0,597	0,828	0,899
AR(3)	0,008*	0,003*	0,282	0,598	0,774	0,804
AR(4)	0,010*	0,000*	0,372	0,828	0,332	0,720
AR(5)	0,008*	0,000*	0,449	0,007*	0,489	0,462

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Tabela 6.51 - Testes das estimativas dos parâmetros ao modelo FSARIMA(5,1,1)(1,0,1)₃₆ ajustado a carga de energia elétrica.

Modelo	Coefficiente	Erro padrão	Estatística t	P-valor
AR(1)	1,357	0,213	6,370	0,000*
AR(2)	-1,992	0,185	-10,760	0,000*
AR(3)	1,790	0,312	5,740	0,000*
AR(4)	-1,367	0,192	-7,120	0,000*
AR(5)	0,669	0,194	3,440	0,001*
SAR(36)	-0,948	0,129	-7,340	0,000*
MA(2)	-0,017	0,032	-31,910	0,000*
SMA(36)	0,754	0,223	3,380	0,002*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Tabela 6.52 - Teste de *Box-Pierce* modificado (*Ljung-Box*) pela estatística Qui-quadrado ao modelo FSARIMA(5,1,1)(1,0,1)₃₆ ajustado para carga de energia elétrica.

Lag	Qui-quadrado	Grau de liberdade	P-valor
12	6,300	02	0,042
24	8,800	14	0,841
36	40,400	26	0,036

Nota: Nível de significância adotado igual a $\alpha = 1\%$.

A significância estatística dos parâmetros do modelo ajustado foi devidamente avaliada

na Tabela (6.51), onde foi possível perceber, os termo de média móvel $q = 1$ e $q = 3$, isto é, MA(1) e MA(3), respectivamente, como não sendo significativos. Este fato é justificado, devido o p -valor = 0,880 (MA(1)), e ainda, p -valor = 0,707 (MA(3)), logo, ambos valores maiores que o nível de significância do teste de 10%. Portanto, o modelo de previsão ajustado não deve considerar estes parâmetros, quando gerar previsões para a carga de energia elétrica. Quanto ao teste *Box-Pierce* (*Ljung-Box*) da estatística Qui-quadrado na Tabela (6.52), em relação ao termo de sazonalidade adotado. A ordem do coeficiente $s = 36$ é estatisticamente significativa, devido o p -valor = 0,036 > 1%. Então, este componente sazonal deve ser mantido no modelo de previsão *fuzzy* para a carga de energia elétrica, mesmo porque, a exclusão do termo $s = 36$ da estrutura do modelo FASARIMA(5;1;1)(1;0;1)₃₆, implicará na redução significativa da acurácia das previsões.

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

Considera-se a medida de desempenho MASE como critério de avaliação das previsões. Ao analisar a Tabela (6.53), há fortes evidências de uma excelente identificação do modelo de previsão, com o auxílio da função de autocorrelação estendida *fuzzy*, mediante a incorporação do critério de ativação dos conjuntos *fuzzy* segundo um α -nível. Esta afirmação é plenamente justificada, devido o valor do MASE para o modelo proposto ser igual a 0,18, que é inferior aos cinco métodos apresentados em Sadaei et al. (2014). O método de previsão que mais se aproximou da acurácia alcançada com o método proposto neste trabalho, como se pode observar na Tabela (6.53), retornou um valor do MASE = 0,37. Como este resultado representa mais que o dobro do erro de previsão alcançado pelo modelo FSARIMA(5;1;1)(1;0;1)₃₆, se pode concluir que o processo de modelagem para previsão *fuzzy* da carga de energia elétrica foi muito bem realizado.

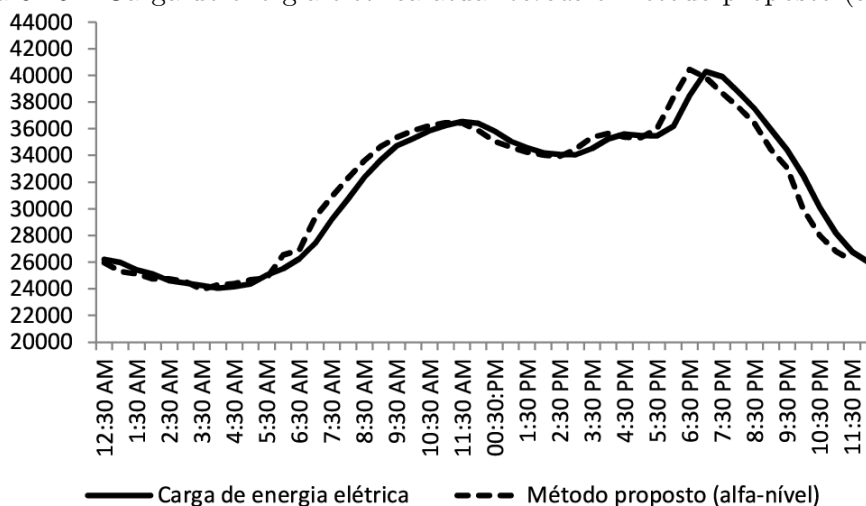
A metodologia de intervalo de confiança *fuzzy*, que foi desenvolvida por Wu (2009), é utilizada de forma complementar na avaliação da capacidade preditiva do modelo proposto neste estudo. Seis projeções para a carga de energia elétrica foram realizadas, a partir do modelo FSARIMA(5,1,1)(1,0,1)₃₆, onde os resultados apresentados na Tabela (6.54), estão todos compreendidos dentro dos intervalos *fuzzy* construídos com 99% de confiança. Estes resultados comprovam uma boa capacidade preditiva do modelo de previsão para a carga de energia elétrica.

Tabela 6.53 - Comparação dos resultados das previsões mediante o método proposto (α -nível) em relação aos cinco métodos propostos em Sadaei et al. (2014) para a carga de energia elétrica.

Instante de tempo observado (horas)	Carga de energia elétrica (MW)	Método OTS-SVR (2014)	Método PSO-ARIMA (2014)	Método LS-ARIMA (2014)	Método REWFTS-Model 1 (2014)	Método Sadaei et al. (2014)	Método proposto: Treinamento	Método proposto: Validação
12:30 AM	26.203	25.459	27.360	26.270	26.606	26.542	-	-
01:00 AM	25.982	24.730	25.296	26.370	25.572	26.205	25.947,6	-
01:30 AM	25.411	24.470	25.032	24.875	24.906	25.179	25.306,5	-
02:00 AM	25.101	26.415	26.038	24.773	24.609	24.784	25.107,6	-
02:30 AM	24.621	23.451	23.943	24.031	24.650	24.579	24.752,4	-
03:00 AM	24.427	24.915	23.388	24.034	24.646	24.011	24.738,1	-
03:30 AM	24.247	23.429	23.596	24.350	24.643	24.020	24.551,0	-
04:00 AM	24.050	24.570	24.648	23.685	24.646	24.101	23.933,2	-
04:30 AM	24.156	23.311	25.129	23.723	24.644	24.583	24.304,1	-
05:00 AM	24.376	23.122	24.993	24.201	24.647	24.581	24.398,4	-
05:30 AM	25.035	24.585	24.413	24.599	24.612	25.349	24.671,6	-
06:00 AM	25.508	25.980	24.879	25.897	24.908	25.749	24.822,4	-
06:30 AM	26.259	25.050	26.132	26.408	26.603	26.520	26.537,2	-
07:00 AM	27.471	26.270	28.551	27.941	26.999	27.069	26.893,0	-
07:30 AM	29.213	29.430	28.093	28.544	28.606	29.492	29.416,1	-
08:00 AM	30.742	30.232	31.279	30.944	30.049	30.426	30.922,8	-
08:30 AM	32.369	32.699	32.494	32.062	32.724	32.086	32.309,0	-
09:00 AM	33.649	32.958	32.953	33.107	34.047	33.981	33.613,1	-
09:30 AM	34.724	33.870	34.375	34.201	34.981	34.351	34.660,8	-
10:00 AM	35.273	35.720	36.354	35.850	35.911	35.749	35.348,7	-
10:30 AM	35.838	35.206	36.327	36.317	36.217	36.148	35.871,4	-
11:00 AM	36.256	36.909	36.152	36.523	36.600	35.960	36.240,8	-
11:30 AM	36.542	35.409	37.220	36.898	36.181	36.839	36.461,7	-
12:00 AM	36.406	38.190	36.606	36.121	36.878	36.815	36.429,8	-
00:30 PM	35.836	34.987	36.550	35.243	35.518	36.147	35.880,9	-
01:00 PM	35.056	36.056	35.761	34.839	35.470	34.780	35.073,3	-
01:30 PM	34.602	33.527	34.986	34.315	34.979	34.349	34.620,2	-
02:00 PM	34.191	35.055	34.141	33.658	35.469	33.725	34.260,8	-
02:30 PM	34.084	35.147	33.672	34.320	35.466	33.801	34.026,1	-
03:00 PM	34.054	32.976	33.152	34.778	35.472	34.028	33.910,0	-
03:30 PM	34.535	35.384	34.966	35.214	34.978	34.889	34.470,6	-
04:00 PM	35.237	35.107	34.231	35.654	35.917	35.180	35.353,8	-
04:30 PM	35.592	35.128	35.959	35.292	35.917	35.123	35.640,1	-
05:00 PM	35.492	34.653	36.596	35.707	35.908	35.222	35.379,7	-
05:30 PM	35.470	36.569	35.286	36.073	35.914	35.100	35.333,6	-
06:00 PM	36.177	35.416	36.924	35.461	35.900	36.538	36.005,4	-
06:30 PM	38.499	37.859	38.540	38.301	38.122	38.760	38.360,2	-
07:00 PM	40.292	41.663	41.030	39.676	39.831	39.896	40.438,2	-
07:30 PM	39.902	40.521	39.133	40.412	40.286	40.213	39.844,8	-
08:00 PM	38.744	37.910	38.437	38.297	38.969	38.186	38.664,3	-
08:30 PM	37.514	36.608	37.703	37.152	37.070	37.217	37.617,0	-
09:00 PM	35.931	37.363	35.696	35.104	35.346	35.579	36.409,5	-
09:30 PM	34.419	33.510	34.926	34.211	34.980	34.310	-	34.431,8
10:00 PM	32.490	31.445	32.043	32.970	32.842	32.208	-	33.064,0
10:30 PM	30.096	28.739	30.303	30.165	30.633	30.388	-	29.951,2
11:00 PM	28.188	29.300	28.515	27.693	27.564	27.917	-	28.007,4
11:30 PM	26.776	27.732	25.788	27.249	27.010	26.379	-	26.815,7
12:00 PM	25.997	25.161	26.790	26.365	25.570	26.200	-	26.116,2
MASE		1,09	0,71	0,51	0,60	0,37	-	0,18

Tabela 6.54 - Intervalo de confiança *fuzzy* para as 6 projeções de carga de energia elétrica pelo método proposto (α -nível).

Horário projetado	Projeção da carga de energia elétrica (MW)	Conjunto <i>fuzzy</i>	Interval <i>fuzzy</i> (u_j)	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
12:30 PM	33.064,00	A4	[29.460; 34.880]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
01:00 AM	31.327,00	A4	[29.460; 34.880]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
01:30 AM	29.564,60	A3	[26.760; 32.170]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
02:00 AM	27.894,40	A2	[24.050; 29.460]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
02:30 AM	26.818,40	A2	[24.050; 29.460]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
03:00 AM	25.549,90	A2	[24.050; 29.460]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]

Figura 6.16 - Carga de energia elétrica atual *versus* o método proposto (α -nível).

Reafirmando as impressões obtidas até o momento com o ajuste do modelo de previsão *fuzzy* estimado. A Figura (6.16), indica um padrão de comportamento de trajetória para a carga de energia elétrica, bem semelhante ao da série de valores das previsões realizadas.

Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para identificar o modelo de série temporal *fuzzy*, sobre os dados históricos de interesse.

Passo 6(c): Para utilizar a função de identificação $\overline{W}_{i,k,j}^{RA}(h)$, se deve selecionar o valor da função de autocorrelação amostral (fac) média aleatória, no suporte do conjunto *fuzzy* ativado pela pertinência máxima em relação a A_j , mediante os intervalos u_j construídos no passo 5.

Analisando a Tabela (6.55), é possível perceber que a pertinência máxima em relação aos conjuntos A_i foi utilizada. Este procedimento permitiu que a série de valores da função fac média aleatória fosse obtida, a partir do suporte de cada conjunto *fuzzy* ativado, com maior aderência do valor da série temporal no instante i , ao conjunto A_j (ver, Figura 6.17). É importante ressaltar que, todos os conjuntos *fuzzy* foram ativados neste processo

Tabela 6.55 - Série histórica da carga de energia elétrica em 25/9/2005, com processo de *fuzzificação* mediante a pertinência máxima do número *fuzzy* triangular, lógica *fuzzy* simples e agrupada, série da função de autocorrelação amostral média aleatória em cada intervalo u_j .

Horário observado	Carga atual (MW)	Pertinência máxima	Intervalos <i>fuzzy</i> (u_j)	Conjuntos <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	fac média aleatória
12:30 AM	26.203	0,794	[24.050; 29.460]	A_2	-	$A_1 \rightarrow A_1, A_2$	0,013
01:00 AM	25.982	0,713	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$	$A_2 \rightarrow A_1, A_2, A_3$	0,017
01:30 AM	25.411	0,502	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$	$A_3 \rightarrow A_2, A_3, A_4$	0,008
02:00 AM	25.101	0,612	[21.340; 26.760]	A_1	$A_2 \rightarrow A_1$	$A_4 \rightarrow A_3, A_5$	0,004
02:30 AM	24.621	0,789	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_5 \rightarrow A_4, A_5, A_6$	0,003
03:00 AM	24.427	0,861	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_6 \rightarrow A_5, A_6, A_7$	0,001
03:30 AM	24.247	0,927	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$	$A_7 \rightarrow A_6, A_7$	0,001
04:00 AM	24.050	1,000	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,001
04:30 AM	24.156	0,961	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,003
05:00 AM	24.376	0,880	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,003
05:30 AM	25.035	0,637	[21.340; 26.760]	A_1	$A_1 \rightarrow A_1$		0,007
06:00 AM	25.508	0,538	[24.050; 29.460]	A_2	$A_1 \rightarrow A_2$		0,015
06:30 AM	26.259	0,815	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,001
07:00 AM	27.471	0,737	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,006
07:30 AM	29.213	0,909	[26.760; 32.170]	A_3	$A_2 \rightarrow A_3$		0,001
08:00 AM	30.742	0,527	[26.760; 32.170]	A_3	$A_3 \rightarrow A_3$		0,001
08:30 AM	32.369	0,927	[29.460; 34.880]	A_4	$A_3 \rightarrow A_4$		0,000
09:00 AM	33.649	0,546	[32.170; 37.590]	A_5	$A_4 \rightarrow A_5$		0,002
09:30 AM	34.724	0,942	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,005
10:00 AM	35.273	0,855	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,003
10:30 AM	35.838	0,646	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001
11:00 AM	36.256	0,508	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,004
11:30 AM	36.542	0,613	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,002
12:00 AM	36.406	0,563	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,001
00:30 PM	35.836	0,647	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,002
01:00 PM	35.056	0,935	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001
01:30 PM	34.602	0,897	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001
02:00 PM	34.191	0,746	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001
02:30 PM	34.084	0,706	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,002
03:00 PM	34.054	0,695	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,001
03:30 PM	34.535	0,873	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000
04:00 PM	35.237	0,868	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,000
04:30 PM	35.592	0,737	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,004
05:00 PM	35.492	0,774	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,008
05:30 PM	35.470	0,782	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,005
06:00 PM	36.177	0,521	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,004
06:30 PM	38.499	0,663	[34.880; 40.290]	A_6	$A_5 \rightarrow A_6$		0,008
07:00 PM	40.292	0,999	[37.590; 43.000]	A_7	$A_6 \rightarrow A_7$		0,008
07:30 PM	39.902	0,856	[37.590; 43.000]	A_7	$A_7 \rightarrow A_7$		0,007
08:00 PM	38.744	0,573	[34.880; 40.290]	A_6	$A_7 \rightarrow A_6$		0,000
08:30 PM	37.514	0,972	[34.880; 40.290]	A_6	$A_6 \rightarrow A_6$		0,003
09:00 PM	35.931	0,612	[32.170; 37.590]	A_5	$A_6 \rightarrow A_5$		0,002
09:30 PM	34.419	0,830	[32.170; 37.590]	A_5	$A_5 \rightarrow A_5$		0,004
10:00 PM	32.490	0,882	[29.460; 34.880]	A_4	$A_5 \rightarrow A_4$		0,001
10:30 PM	30.096	0,765	[26.760; 32.170]	A_3	$A_4 \rightarrow A_3$		0,001
11:00 PM	28.188	0,529	[26.760; 32.170]	A_3	$A_3 \rightarrow A_3$		0,001
11:30 PM	26.776	0,994	[24.050; 29.460]	A_2	$A_3 \rightarrow A_2$		0,005
12:00 PM	25.997	0,718	[24.050; 29.460]	A_2	$A_2 \rightarrow A_2$		0,002

de *fuzzificação*, o que garante uma transição completa entre estes conjuntos, ou seja, todas as partições do universo de discurso U foram ativadas neste processo. Este fato indica uma divisão de U realizada de forma eficiente, isto é, compreendendo toda a trajetória da série temporal da carga de energia elétrica estudada.

A função de autocorrelação estendida média aleatória foi aplicada, sobre a série de dados resultante do processo de *fuzzificação* da Tabela (6.55). Os resultados desta rotina de identificação que são apresentados nas Tabelas (6.56 e 6.57), indicam um modelo ARMA(5; 3), como o modelo de maior ordem para os parâmetros p e q , com um nível de significância estatística de 10%. Utiliza-se uma plataforma de otimização paramétrica pontual, auxiliando o ajuste do modelo de previsão *fuzzy*. Todos os modelos que satisfaçam à restrição: $0 < p \leq 5$ e $0 < q \leq 3$, durante o processo de modelagem serão considerados e avaliados computacionalmente.

Figura 6.17 - Conjuntos *fuzzy* para a carga de energia elétrica de curto prazo, durante 48 instantes de tempo no período de 12:30 AM a 11:30 PM coletados a cada meia hora.

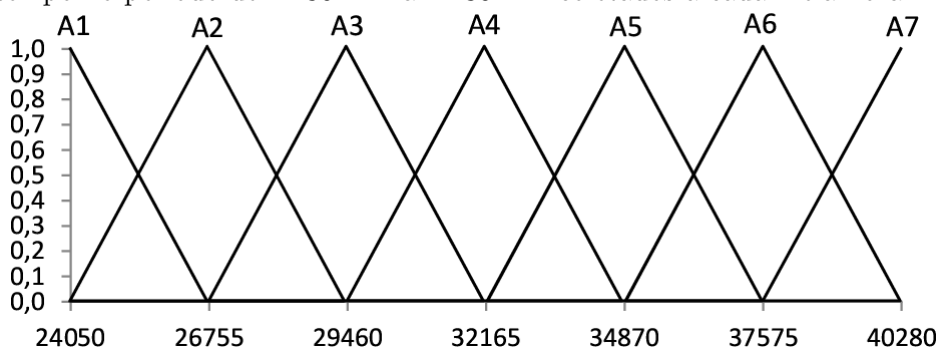


Tabela 6.56 - Resultados da $\overline{W}_{i,k,j}^{RA}(h; b)$ para a carga de energia elétrica.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,405	0,009	-0,164	0,174	-0,026	-0,114
AR(1)	-0,379	-0,027	-0,125	0,172	-0,022	-0,035
AR(2)	-0,452	-0,442	-0,366	0,029	-0,075	-0,011
AR(3)	-0,136	0,035	-0,393	-0,085	-0,106	-0,228
AR(4)	0,090	-0,011	-0,171	-0,094	0,033	-0,214
AR(5)	0,023	-0,171	-0,136	-0,313	-0,045	-0,079

Tabela 6.57 - Probabilidade de significância dos resultados da Tabela 6.56.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,006*	0,956	0,328	0,309	0,882	0,515
AR(1)	0,010*	0,870	0,481	0,301	0,909	0,868
AR(2)	0,002*	0,003*	0,020*	0,865	0,765	0,965
AR(3)	0,368	0,820	0,012*	0,643	0,618	0,196
AR(4)	0,554	0,946	0,460	0,678	0,882	0,260
AR(5)	0,880	0,308	0,592	0,086*	0,826	0,743

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

O critério do menor erro de previsão retornado pelo processo de ajuste, neste caso será considerado para escolha do modelo. Desta maneira, acredita-se que o modelo de previsão

será muito melhor estruturado, como também alcançará maior capacidade preditiva em relação à carga de energia elétrica.

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6(c);

Durante a rotina computacional de busca do valor ótimo, para o ajuste dos parâmetros do modelo de previsão *fuzzy*, o processo indicou o modelo FSARIMA(3;1;1)(1;0;1)₃₆. Na Tabela (6.58), a significância estatística dos parâmetros do modelo ajustado, foi testada a um nível de 10%. Neste contexto, a estatística de teste *t-Student*, demonstra os parâmetros AR(2) e AR(3), como sendo não significativos (*p-valor* = 0,316 e 0,201, respectivamente). Devido estes dois resultados serem maiores que 10%, o modelo de previsão *fuzzy* ajustado irá excluir estes dois parâmetros, pois, os mesmos são estatisticamente iguais à zero ao nível de significância de 10%.

Tabela 6.58 - Testes das estimativas dos parâmetros ao modelo FSARIMA(3,1,1)(1,0,1)₃₆ ajustado a carga de energia elétrica.

Modelo	Coefficiente	Erro padrão	Estatística <i>t</i>	P-valor
AR(1)	0,914	0,140	6,550	0,000*
AR(4)	-0,948	0,129	-7,340	0,000*
AR(5)	0,416	0,146	2,840	0,007*
SAR(36)	1,000	0,001	2211,880	0,000*
MA(1)	-1,048	0,001	-1054,340	0,000*
SMA(36)	-0,716	0,231	-3,100	0,004*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Um teste Qui-quadrado é necessário para avaliar a significância do termo $s = 36$ (Tabela (6.59)). O resultado do teste indica ao nível de 1%, uma não significância estatística, pois, $p\text{-valor} = 0,032 > 1\%$. Portanto, o componente de sazonalidade $s = 36$, obrigatoriamente deve ser mantido no modelo de previsão *fuzzy*, em face do erro retornado necessariamente ser um ruído branco, o que contribui para previsão da carga de energia elétrica.

Tabela 6.59 - Teste de *Box-Pierce* modificado (*Ljung-Box*) pela estatística Qui-quadrado ao modelo FSARIMA(3,1,1)(1,0,1)₃₆ ajustado para carga de energia elétrica.

Lag	Qui-quadrado	Grau de liberdade	P-valor
12	17,300	4	0,002*
24	24,400	16	0,080
36	43,400	28	0,032
48	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 1\%$.

Na Tabela (6.60), as evidências de uma excelente identificação do modelo de previsão *fuzzy*, são fortemente ratificadas. A medida de desempenho MASE, para este método pro-

Tabela 6.60 - Comparação dos resultados das previsões do método proposto (fac média aleatória) com os cinco métodos propostos em Sadaei et al. (2014) para a carga de energia elétrica.

Instante de tempo observado (horas)	Carga de energia elétrica (MW)	Método OTS-SVR (2014)	Método PSO-ARIMA (2014)	Método LS-ARIMA (2014)	Método REWFTS-Model 1 (2014)	Método Sadaei et al. (2014)	Método proposto: Treinamento	Método proposto: Validação
12:30 AM	26.203	25.459	27.360	26.270	26.606	26.542	-	-
01:00 AM	25.982	24.730	25.296	26.370	25.572	26.205	26.059,6	-
01:30 AM	25.411	24.470	25.032	24.875	24.906	25.179	25.669,2	-
02:00 AM	25.101	26.415	26.038	24.773	24.609	24.784	24.917,6	-
02:30 AM	24.621	23.451	23.943	24.031	24.650	24.579	24.231,1	-
03:00 AM	24.427	24.915	23.388	24.034	24.646	24.011	24.370,1	-
03:30 AM	24.247	23.429	23.596	24.350	24.643	24.020	24.474,4	-
04:00 AM	24.050	24.570	24.648	23.685	24.646	24.101	24.234,6	-
04:30 AM	24.156	23.311	25.129	23.723	24.644	24.583	24.255,8	-
05:00 AM	24.376	23.122	24.993	24.201	24.647	24.581	24.476,1	-
05:30 AM	25.035	24.585	24.413	24.599	24.612	25.349	24.876,7	-
06:00 AM	25.508	25.980	24.879	25.897	24.908	25.749	25.522,2	-
06:30 AM	26.259	25.050	26.132	26.408	26.603	26.520	26.232,1	-
07:00 AM	27.471	26.270	28.551	27.941	26.999	27.069	27.530,0	-
07:30 AM	29.213	29.430	28.093	28.544	28.606	29.492	29.240,9	-
08:00 AM	30.742	30.232	31.279	30.944	30.049	30.426	30.748,5	-
08:30 AM	32.369	32.699	32.494	32.062	32.724	32.086	32.442,3	-
09:00 AM	33.649	32.958	32.953	33.107	34.047	33.981	33.623,0	-
09:30 AM	34.724	33.870	34.375	34.201	34.981	34.351	34.708,9	-
10:00 AM	35.273	35.720	36.354	35.850	35.911	35.749	35.283,9	-
10:30 AM	35.838	35.206	36.327	36.317	36.217	36.148	35.726,2	-
11:00 AM	36.256	36.909	36.152	36.523	36.600	35.960	36.239,8	-
11:30 AM	36.542	35.409	37.220	36.898	36.181	36.839	36.510,1	-
12:00 AM	36.406	38.190	36.606	36.121	36.878	36.815	36.317,8	-
00:30 PM	35.836	34.987	36.550	35.243	35.518	36.147	35.920,8	-
01:00 PM	35.056	36.056	35.761	34.839	35.470	34.780	35.061,9	-
01:30 PM	34.602	33.527	34.986	34.315	34.979	34.349	34.639,7	-
02:00 PM	34.191	35.055	34.141	33.658	35.469	33.725	34.366,7	-
02:30 PM	34.084	35.147	33.672	34.320	35.466	33.801	34.068,8	-
03:00 PM	34.054	32.976	33.152	34.778	35.472	34.028	34.133,9	-
03:30 PM	34.535	35.384	34.966	35.214	34.978	34.889	34.598,5	-
04:00 PM	35.237	35.107	34.231	35.654	35.917	35.180	35.061,2	-
04:30 PM	35.592	35.128	35.959	35.292	35.917	35.123	35.598,3	-
05:00 PM	35.492	34.653	36.596	35.707	35.908	35.222	35.340,7	-
05:30 PM	35.470	36.569	35.286	36.073	35.914	35.100	35.232,0	-
06:00 PM	36.177	35.416	36.924	35.461	35.900	36.538	36.226,1	-
06:30 PM	38.499	37.859	38.540	38.301	38.122	38.760	38.312,9	-
07:00 PM	40.292	41.663	41.030	39.676	39.831	39.896	40.268,0	-
07:30 PM	39.902	40.521	39.133	40.412	40.286	40.213	40.065,0	-
08:00 PM	38.744	37.910	38.437	38.297	38.969	38.186	38.501,1	-
08:30 PM	37.514	36.608	37.703	37.152	37.070	37.217	37.436,5	-
09:00 PM	35.931	37.363	35.696	35.104	35.346	35.579	36.076,0	-
09:30 PM	34.419	33.510	34.926	34.211	34.980	34.310	-	34.647,0
10:00 PM	32.490	31.445	32.043	32.970	32.842	32.208	-	33.045,3
10:30 PM	30.096	28.739	30.303	30.165	30.633	30.388	-	30.189,7
11:00 PM	28.188	29.300	28.515	27.693	27.564	27.917	-	28.358,6
11:30 PM	26.776	27.732	25.788	27.249	27.010	26.379	-	26.851,1
12:00 PM	25.997	25.161	26.790	26.365	25.570	26.200	-	25.806,5
MASE		1,09	0,71	0,51	0,60	0,37	-	0,14

posto assumiu um valor igual a 0,14, em relação a resultados bem superiores dos demais métodos avaliados. O método de previsão que mais se aproximou do desempenho, obtido pelo modelo FSARIMA(3;1;1)(1;0;1)₃₆, foi o método de Sadaei et al. (2014), com valor

do $MASE = 0,37$. No entanto, este valor do $MASE = 0.37$, é aproximadamente 2,5 vezes superior, ao que foi obtido pelo método proposto. Vale ressaltar ainda, que o valor mínimo assumido pelo MASE é zero, por se tratar de uma métrica com erro absoluto. Portanto, o erro de previsão obtido pelo modelo $FSARIMA(3;1;1)(1;0;1)_{36}$, que foi identificado pelo algoritmo está bem próximo deste limite.

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

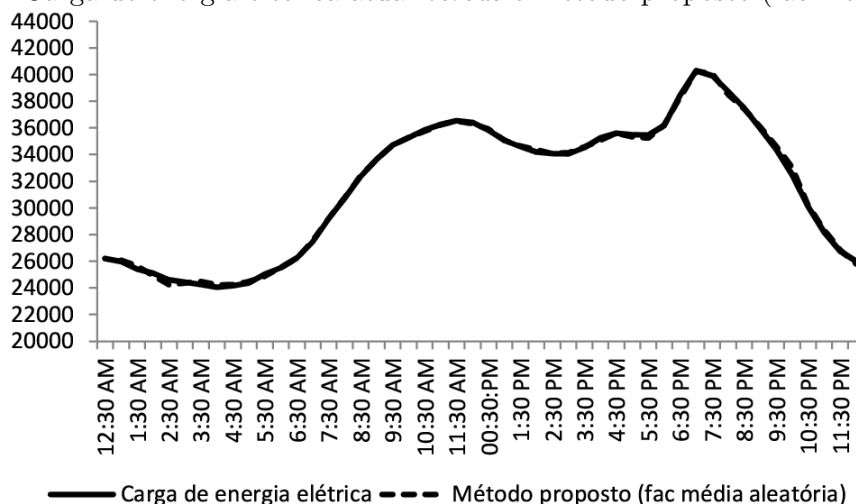
Quanto a acurácia das projeções realizadas pelo modelo $FSARIMA(3;1;1)(1;0;1)_{36}$. Na Tabela (6.61), é possível observar seis projeções para carga de energia elétrica, referentes aos instantes de tempo 12:30 PM a 03:00 AM, a partir da série temporal original.

Os intervalos de confiança *fuzzy* foram construídos, com um altíssimo grau de confiabilidade (99%). Desta forma, como todas as projeções produzidas pelo modelo $FSARIMA(3;1;1)(1;0;1)_{36}$, pertencem a pelo menos um dos intervalos de confiança *fuzzy*, se pode aceitar que o modelo de previsão está bem identificado.

Tabela 6.61 - Intervalo de confiança *fuzzy* para as 6 projeções de carga de energia elétrica pelo método proposto (fac média aleatória).

Horário projetado	Projeção da carga de energia elétrica (MW)	Conjunto <i>fuzzy</i>	Interval <i>fuzzy</i> (u_j)	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
12:30 PM	26.859,60	A2	[24.050; 29.460]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
01:00 AM	28.641,50	A3	[26.760; 32.170]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
01:30 AM	31.736,40	A4	[29.460; 34.880]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
02:00 AM	34.892,10	A5	[32.170; 37.590]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
02:30 AM	37.585,80	A6	[34.880; 40.290]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]
03:00 AM	39.733,10	A7	[37.590; 43.000]	[23.760,12; 35.168,45]	[29.175,91; 40.581,23]

Figura 6.18 - Carga de energia elétrica atual *versus* o método proposto (fac média aleatória).



A Figura (6.18) reflete graficamente, o comportamento temporal dos valores atuais *versus*

as previsões pelo método proposto. Nesta Figura (6.18), claramente existe uma sobreposição dos valores atuais da série temporal de carga de energia elétrica, em relação aos valores de previsão pelo modelo FSARIMA(3;1;1)(1;0;1)₃₆.

6.1.3 - Estudo de caso 3: Previsão fuzzy para a quantidade de matrículas na Universidade do Alabama nos Estados Unidos

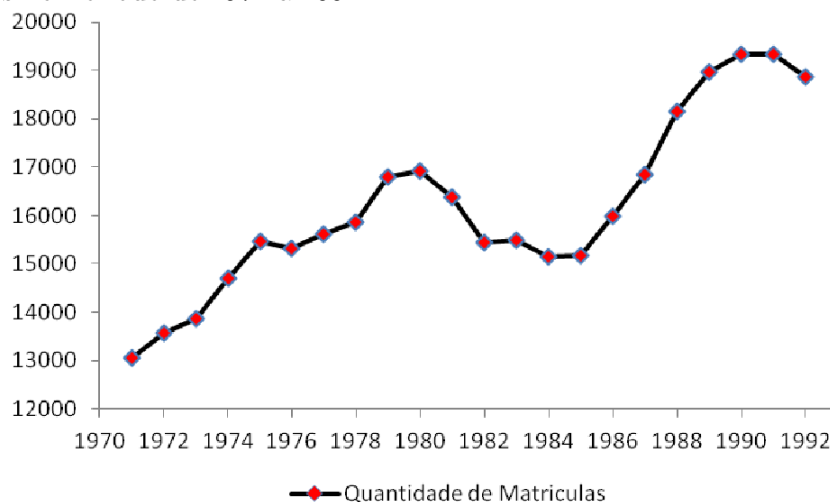
Para avaliar novamente o desempenho das ferramentas propostas neste trabalho, quanto à capacidade na identificação de processos autoregressivos. Será utilizada a série de dados históricos referente à quantidade de matrículas na Universidade do Alabama nos Estados Unidos. Estes dados inicialmente foram utilizados por Song e Chisson (1993a), posteriormente utilizados em Wong et al. (2010), Egrioglu et al. (2010, 2011), Chen and Tanuwijaya (2011), Carvalho Jr. et al. (2012), Gangwar and Kumar (2012), Qiu et al. (2012), Shah (2012), Jasim et al. (2012), Wang et al. (2013), Egrioglu et al. (2013), Feng et al. (2014), Wang et al. (2014a), dentre outros. Seguindo a rotina dada no Capítulo 4 (Seção 4.6), os passos serão implementados neste exemplo.

Passo 1: Coletar os dados de interesse que formarão a série temporal.

Como a base de dados referente a quantidade de matrículas na Universidade do Alabama está disponível na literatura, se pode iniciar a rotina descrita na Seção 4.6, pelo Passo 2.

Passo 2: Construir o gráfico da série temporal referente a quantidade de matrículas na Universidade do Alabama nos Estados Unidos, durante o período de 1971 a 1992.

Figura 6.19 - Série Temporal da Quantidade de Matrículas na Universidade do Alabama nos Estados Unidos no Período de 1971 a 1992.



Passo 3: Definir o Universo de Discurso.

O universo de discurso é definido segundo a Equação (4.42), para que desta forma, seja possível garantir que o valor máximo da série histórica (Z_i) representado por $Z_{(Max)}$ e o valor mínimo da série histórica representado por $Z_{(Min)}$, pertençam ao intervalo numérico representado por U .

Portanto, para a série histórica da quantidade de matrículas, se tem que o resultado é $\sigma = 1.045$, a partir da Equação (4.43). Então, como $Z_{(Min)} = 13.055$ e $Z_{(Max)} = 19.337$, o que leva a Equação (4.42) assumir os valores, $U = [12.010; 20.380]$.

Passo 4: Particionar o Universo de Discurso em “ m ” Conjuntos *Fuzzy*.

Para que seja possível particionar U , se faz necessário definir a amplitude (comprimento) “ l ” de cada intervalo que será construído, tal como, descrito na Equação (4.44).

Como o objetivo inicial é comparar os resultados obtidos por Song (2003), com o das ferramentas de identificação propostas neste trabalho, serão utilizados sete conjuntos *fuzzy*, ou seja, $m = 7$, tal como, o autor citado acima utilizou na rotina de *fuzzificação* da série de matrículas. Porém, neste estudo a função de pertinência será calculada a partir da definição de número *fuzzy* triangular, ao invés da função trivalente utilizada por Song (2003). Assim, para $m = 7$, se tem $l = 2.090$.

Passo 5: Construir os conjuntos *fuzzy* correspondentes a quantidade (m) estabelecida no passo 4 e *fuzzificar* a série histórica a partir da definição de número *fuzzy* triangular.

Seja A_j a representação do j -ésimo conjunto *fuzzy*, tal que, $j = 1; 2; \dots; 6; 7$. Logo, caso A_j seja um número *fuzzy* triangular, tal como descrito na Definição 3.5, se chega que, $u_j = [a_j; b_j; c_j]$.

$$A_j = \begin{cases} 0, & \text{se } Z_i \leq a_j; \\ \frac{Z_i - a_j}{b_j - a_j}, & \text{se } a_j < Z_i \leq b_j; \\ \frac{c_j - Z_i}{c_j - b_j}, & \text{se } b_j < Z_i \leq c_j. \\ 0, & \text{se } Z_i > c_j, \end{cases}$$

Neste contexto, se tem sete intervalos, onde devido a função de pertinência a ser calculada, para cada valor da quantidade de matrículas em relação aos conjuntos A_j , provém da definição de número *fuzzy* triangular, cada intervalo representa a base de um dos sete triângulos que serão construídos, como mostra a Figura (6.20). Portanto, como os intervalos são representados pelos pontos a_j , b_j e c_j , onde $a_j < b_j < c_j$, para $j = 1; 2; \dots; 7$. Então, os intervalos são dados por:

$$u_1 = [12.010; 13.060; 14.100], \quad u_2 = [13.060; 14.100; 15.150], \quad u_3 = [14.105; 15.150; 16.195], \\ u_4 = [15.150; 16.200; 17.240], \quad u_5 = [16.200; 17.240; 18.290], \quad u_6 = [17.245; 18.290; 19.335] \text{ e} \\ u_7 = [18.290; 19.340; 20.380].$$

Torna-se importante destacar, que a ativação de cada conjunto *fuzzy* é realizada pela pertinência máxima, do valor da quantidade de matrícula em relação ao conjunto *fuzzy* A_j , pois um mesmo valor de matrícula pode está contido em mais de um intervalo u_j , ou seja, uma mesma quantidade de matrículas pode pertencer a dois conjuntos *fuzzy*, onde o conjunto *fuzzy* ativado no processo de *fuzzificação* será aquele onde a quantidade de matrículas observadas, apresentar maior valor de pertinência. Por exemplo, a quantidade de matrículas 13.055 apresenta pertinência ao conjunto *fuzzy* A_1 igual a 0,995 e 0,005 em relação ao conjunto *fuzzy* A_2 na Figura (6.20). Assim, o valor 13.055 irá ativar pela pertinência máxima o conjunto *fuzzy* A_1 , na rotina de *fuzzificação*. Com raciocínio semelhante, o valor da quantidade de matrículas igual a 13.563, coincidentemente também está contido nos dois conjuntos *fuzzy* A_1 e A_2 , como se pode observa na Figura (6.20), onde a pertinência máxima obtida é de 0,516 em relação ao conjunto *fuzzy* A_1 e 0,484 em relação ao conjunto A_2 , portanto, pela pertinência máxima obtida o valor de matrícula 13.563 irá ativar o conjunto *fuzzy* A_1 . Neste contexto, a série temporal *fuzzy* apresentada na Tabela (6.62) é construída, com base em cada valor da série histórica para as quantidades de matrículas observadas entre 1971 e 1992.

Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para determinar a ordem do modelo de série temporal *fuzzy*, em relação aos dados observados de uma variável de interesse.

(a) Para a função $W_{i,k,j}^{DFZ}$, se deve *defuzzificar* a série histórica mediante o método centroide, e em seguida aplicar a função de autocorrelação estendida *fuzzy* nesta série *defuzzificada* obtida.

Figura 6.20 - Conjuntos *Fuzzy* Pela Definição de Número *Fuzzy* Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, onde o Eixo Y Representa a Pertinência ao Conjunto *Fuzzy* A_i e o Eixo X Indica a Quantidade de Matrículas.

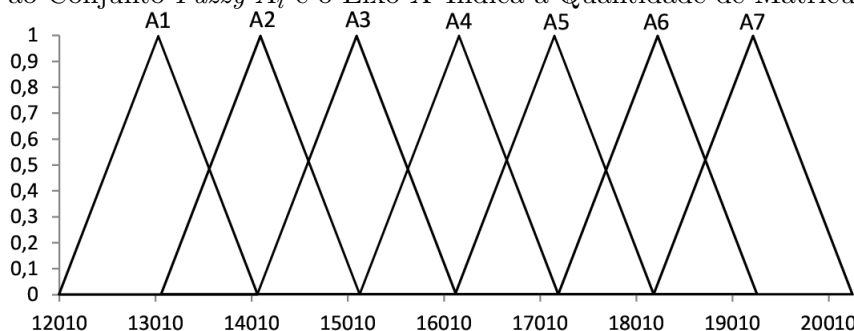


Tabela 6.62 - Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de *Fuzzificação* Realizado a partir de Um Número *Fuzzy* Triangular, Relação *Fuzzy* Lógica e os Valores de Matrículas *Defuzzificadas*.

Ano	Valor de Matrícula	Pertinência Máxima	Intervalo Correspondente	Série Temporal <i>Fuzzy</i>	Relação <i>Fuzzy</i> Lógica	Matrículas <i>Defuzzificadas</i>
1971	13.055	0,995	[12.010; 14.100]	A_1	-	13.009
1972	13.563	0,516	[12.010; 14.100]	A_1	$A_1 \rightarrow A_1$	13.074
1973	13.867	0,776	[13.060; 15.150]	A_2	$A_1 \rightarrow A_2$	14.084
1974	14.696	0,566	[14.105; 16.195]	A_3	$A_2 \rightarrow A_3$	15.153
1975	15.460	0,703	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	15.143
1976	15.311	0,846	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	15.148
1977	15.603	0,566	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	15.151
1978	15.861	0,677	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	16.170
1979	16.807	0,584	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	17.271
1980	16.919	0,691	[16.200; 18.290]	A_5	$A_5 \rightarrow A_5$	17.244
1981	16.388	0,819	[15.150; 17.240]	A_4	$A_5 \rightarrow A_4$	16.180
1982	15.433	0,729	[14.105; 16.195]	A_3	$A_4 \rightarrow A_3$	15.171
1983	15.497	0,668	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	15.148
1984	15.145	0,995	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	15.151
1985	15.163	0,988	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	15.140
1986	15.984	0,794	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	16.205
1987	16.859	0,634	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	17.237
1988	18.150	0,866	[17.245; 19.335]	A_6	$A_5 \rightarrow A_6$	18.271
1989	18.970	0,648	[18.290; 20.380]	A_7	$A_6 \rightarrow A_7$	19.327
1990	19.328	0,989	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	19.284
1991	19.337	0,997	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	19.375
1992	18.876	0,558	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	19.361

Inicialmente se constrói os conjuntos *fuzzy*, a partir da definição de número *fuzzy* triangular, tal como, se observa na Figura (6.20), para posteriormente obter a pertinência máxima de cada valor das matrículas ao conjunto *fuzzy* correspondente, e em seguida obter a série temporal *fuzzy* e os valores *defuzzificados* dados na Tabela (6.62) mediante o método centróide, pois este método não considera apenas os valores máximos para o

calculado, e ainda, por dividir a área de saída *fuzzy* em duas partes iguais, como também devido a função de pertinência triangular apresentar sobreposição, como se observa na Figura (6.20), para justificar a utilização do método centróide.

Seguindo o passo 6 (a), da rotina descrita na Seção 5.6, se obteve os valores de $W_{i,k,j}^{DFZ}$, com suas probabilidades de significância (*P-Valor*) correspondentes, que são apresentados na Tabela (6.64) e Tabela (6.65).

Como no processo de *fuzzificação* realizado sobre os dados de matrículas (Tabela 6.62), resultam em relações *fuzzy* para pares de valores da série de dados observados, também foram estabelecidos os grupos destas relações *fuzzy* lógicas, como mostra a Tabela (6.63). O grupo de relações *fuzzy* lógicas indica a transição conjunta, e não apenas individual, de cada suporte, isto é, dos conjuntos *fuzzy* ativados, quando cada valor da série de matrículas é considerado no processo de *fuzzificação*.

Tabela 6.63 - Grupos de Relações *Fuzzy* Lógicas Ativadas Para os Dados da Tabela (6.62).

Grupo Ativado	Relações <i>Fuzzy</i> Lógicas
Grupo 1	$A_1 \rightarrow A_1, A_1 \rightarrow A_2$
Grupo 2	$A_2 \rightarrow A_3$
Grupo 3	$A_3 \rightarrow A_3, A_3 \rightarrow A_4$
Grupo 4	$A_4 \rightarrow A_3, A_4 \rightarrow A_5$
Grupo 5	$A_5 \rightarrow A_4, A_5 \rightarrow A_5, A_5 \rightarrow A_6$
Grupo 6	$A_6 \rightarrow A_7$
Grupo 7	$A_7 \rightarrow A_7$

Caso se rejeite a hipótese $H_0 : (p; q) = 0$, para um modelo ARMA($p; q$), este é considerado significativo, ou seja, os parâmetros p e q são estatisticamente diferentes de zero à um nível de significância α pré-estabelecido, e portanto, o modelo deve ser considerado satisfatório para representar a série temporal, o que resulta em gerar previsões sobre a característica de interesse.

Mediante os resultados das Tabelas (6.64) e (6.65), se pode concluir que o modelo identificado pela função $W_{i,k,j}^{DFZ}$ é um ARMA(4; 1), onde o valor da função para estes parâmetros é -0,692 (Ver, Tabela 6.64) e a probabilidade de significância desta estimação paramétrica pontual é 0,007, como se observa na Tabela (6.65), ou seja, a um nível de significância de 10% se tem que, $P\text{-valor} = 0,007 < \alpha = 0,10$, o que leva a rejeitar a hipótese de que estes parâmetros sejam iguais a zero. No entanto, apesar da função de autocorrelação estendida ter identificado outro modelo, um ARMA(0; 0), com *P-valor*

= 0,021 (significativo para $\alpha = 0,10$), devido este não indicar valor numérico para os parâmetros $p > 0$ e $q > 0$, no modelo ARMA(p ; q), se deve considerar apenas o segundo modelo identificado, pela função $W_{i,k,j}^{DFZ}$, que é o ARMA(4; 1).

Tabela 6.64 - Resultados da $W_{i,k,j}^{DFZ}$ para as Matrículas na Universidade do Alabama.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,504	-0,066	-0,287	-0,260	-0,113	-0,214
AR(1)	0,277	0,016	-0,256	-0,135	0,230	-0,018
AR(2)	0,034	-0,263	-0,074	-0,135	0,164	0,010
AR(3)	-0,023	-0,361	-0,146	-0,137	0,141	-0,040
AR(4)	0,258	-0,692	-0,386	-0,147	0,204	-0,427
AR(5)	0,262	-0,337	-0,413	-0,100	-0,434	-

Tabela 6.65 - Probabilidades de Significância dos Resultados da $W_{i,k,j}^{DFZ}$ (Tabela 6.64).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,021*	0,805	0,286	0,358	0,700	0,470
AR(1)	0,216	0,956	0,340	0,608	0,387	0,948
AR(2)	0,883	0,262	0,773	0,619	0,579	0,973
AR(3)	0,923	0,129	0,600	0,669	0,654	0,885
AR(4)	0,287	0,007*	0,198	0,658	0,525	0,247
AR(5)	0,296	0,236	0,181	0,754	0,226	-

Nota: * Valor significativo para $\alpha = 10\%$.

Então, as Tabelas (6.64) e (6.65) sugerem um modelo de série temporal *fuzzy*, com o componente autoregressivo de ordem 4 e de média móvel de ordem 1, ou seja, um modelo ARMA(4; 1), onde este resultado difere do obtido por Song (2003), pois este identificou um modelo de baixa ordem (no caso um MA(2)) para esta mesma base de dados, que está sendo utilizada neste exemplo, isto pode ser explicado devido Song (2003) ter utilizado a função de autocorrelação amostral, a qual possui uma capacidade de identificação limitada, pois considera apenas o componente de média móvel e não abrange o componente autoregressivo de um modelo ARMA, algo que é capturado integralmente pelo algoritmo da função de autocorrelação estendida *fuzzy*. A função identifica eficientemente o termo AR e MA simultaneamente, como observado nas Tabelas (6.64) e (6.65), pelos resultados obtidos pela função $W_{i,k,j}^{DFZ}$ e *p-valor*, respectivamente.

É necessário testar a partir de então, a estacionariedade da série para atender aos pressupostos de uma série temporal. Os resultados apresentados na Tabela (6.66), quanto ao teste da raiz unitária de *Dickey-Fuller* aumentado (ver, Hylleberg, 1995), para a série temporal das matrículas sem diferenciação ($d = 0$), demonstram uma não-estacionariedade desta série, pois para o modelo utilizado o valor da estatística de teste *tau* é de -1,553 re-

tornando um P-valor igual a 0,487, assim, não se rejeitou a hipótese nula de raiz unitária, o que implicou numa série não estacionária. Porém, ao aplicar a diferença de primeira ordem ($d = 1$) na série da quantidade de matrículas, esta passou a apresentar comportamento estacionário, ao longo do período compreendido pela amostra observada, pois, o valor de *tau* reduziu para -3,08365, com probabilidade de significância igual a 0,045044 (ver, Tabela 6.67), o que implicou em rejeitar a hipótese nula de raiz unitária, logo, a um nível de significância de 10% a série temporal das matrículas na Universidade do Alabama é estacionária, permitindo assim a utilização da metodologia de série temporal.

Tabela 6.66 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	<i>Tau</i>	P-valor
1	OLS	488136,00	2673,19	20	-1,55300	0,48700
2	COV	488136,00	2575891,32	20	-1,55300	0,48700
3	COV	488136,00	-160,24	20	-1,55300	0,48700
4	COV	488136,00	120,89	20	-1,55300	0,48700

Tabela 6.67 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	<i>Tau</i>	P-valor
1	OLS	158527,44	554,39	19	-3,08365	0,045044*
2	COV	158527,44	40293,13	19	-3,08365	0,045044*
3	COV	158527,44	-60,63	19	-3,08365	0,045044*
4	COV	158527,44	32,75	19	-3,08365	0,045044*

Nota: * Valor significativo para $\alpha = 10\%$.

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6 (a);

Após identificar a ordem do modelo de série temporal *fuzzy*, com o auxílio da função $W_{i,k,j}^{DFZ}$, no passo 6(a) da rotina estabelecida na Seção 4.6, se deve seguir ao passo 7, ou seja, gerar as previsões para as matrículas na Universidade do Alabama no período de 1971 a 1992. A partir da Tabela (6.65), se tem uma combinação de parâmetros significativos (estatisticamente diferentes de zero, ao nível de significância $\alpha = 10\%$), para o modelo de série temporal *fuzzy*, onde após a utilização do modelo com o termo autoregressivos e de médias móveis, sendo os parâmetros $p = 4$ e $q = 1$, se chegou a um modelo FSARIMA($p; d; q$)($P; D; Q$)_s, ou seja, um modelo SARIMA *fuzzy*, tal como, definido por Tseng e Tzeng (2002), visto que, a série dos dados *defuzzificados* das matrículas na Universidade do Alabama, após várias simulações em busca do melhor modelo de série temporal, a partir da identificação sugerida pela função $W_{i,k,j}^{DFZ}$, os dados da série se ajustam perfeitamente ao modelo FSARIMA(4; 1; 1)(1; 0; 1)₁₃, modelo o qual foi testada a significância de seus parâmetros a partir de um *software* estatístico, sendo que, este modelo apresentou

um MASE igual a 0,226 (ver, Tabela, 6.70), o qual é inferior a outros métodos de previsão *fuzzy* em série temporal, que são encontrados na literatura para a mesma base de dados utilizada neste trabalho, indicando assim, que o processo de identificação proposto neste trabalho a partir da função $W_{i,k,j}^{DFZ}$ é eficiente, pois contribui de forma significativa para a diminuição do erro de previsão pelo modelo sugerido, com o auxílio da função de autocorrelação estendida *fuzzy*.

A Tabela (6.68) apresenta os coeficientes do modelo ajustado FSARIMA(4; 1; 1)(1; 0; 1)₁₃, assim como os testes de significância dos mesmos, mediante a estatística de teste oriunda da função densidade de probabilidade *t-Student*, com as respectivas probabilidades de significância.

Tabela 6.68 - Teste de Significância dos Parâmetros do Modelo FSARIMA(4; 1; 1)(1; 0; 1)₁₃.

Modelo	Coefficiente	Erro Padrão	Estatística <i>t</i>	<i>P-Valor</i>
AR(1)	-0,388	0,093	-4,190	0,001*
AR(2)	-0,453	0,113	-4,020	0,001*
AR(3)	-0,391	0,118	-3,320	0,005*
AR(4)	-1,002	0,096	-10,480	0,000*
SAR(13)	0,999	0,044	22,460	0,000*
MA(1)	-0,981	0,296	-3,320	0,005*
SMA(13)	-0,668	0,363	-1,840	0,087*

Nota: * Todos os coeficientes são significativos (estatisticamente diferente de zero) para $\alpha = 10\%$.

Pelos resultados da Tabela (6.68), se pode concluir que o modelo ajustado é significativo, visto que, os testes para os coeficientes do modelo FSARIMA(4; 1; 1)(1; 0; 1)₁₃, a partir da estatística de teste *t-Student* são significativos, a um nível de significância de 10%, inclusive o coeficiente para o componente sazonal ($s = 13$) do termo de média móvel com parâmetro 1, pois neste caso, o $P-Valor = 0,087 < \alpha = 0,10$, logo, este coeficiente é significativo, ou seja, é estatisticamente diferente de zero, pois as hipóteses que estão sendo testadas neste caso específico são as de que, $H_0 : \phi(B)(SMA) = 0$ versus $H_1 : \phi(B)(SMA) \neq 0$. Para os demais coeficientes, a hipótese de igualdade dos mesmos à zero, sempre será rejeitada para $\alpha = 10\%$, visto que, o *P-valor* sempre será igual a 0,000 (AR(4) e SAR(12)), ou então 0,005 (MA(1)), portanto, menor do que $\alpha = 0,10$, levando a manter estes termos no modelo de previsão ajustado à série temporal *fuzzy*.

Na Tabela (6.69), tem-se a estatística de teste *Q* “*Ljung-Box*”, para avaliar se os dados e erros do modelo de previsão *fuzzy* utilizados são independentes, o que impossibilitaria a utilização da teoria de série temporal, caso a suposição de independência entre os dados

não seja refutada, mas, a independência dos erros deve ser garantida para que se tenha um ruído branco na série.

Tabela 6.69 - Teste Q (Ljung-Box) Para o Modelo FSARIMA(4; 1; 1)(1; 0; 1)₁₃.

Fonte de Variação	Lag	Qui-Quadrado	Grau de Liberdade	P-Valor
Teste Q "Ljung-Box"	12,000	18,200	5,000	0,003*
⋮	24,000	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 1\%$.

Em relação ao teste Q *Ljung-Box* na Tabela (6.69), este testa as hipóteses $H_0 : \rho_1 = \rho_2 = \rho_3 = \dots = \rho_n$, versus $H_1 : \rho_1 \neq \rho_2 \neq \rho_3 \neq \dots \neq \rho_n$, ou seja, em H_0 : Os erros do modelo de série temporal são independentes versus H_1 : Os erros do modelo de série temporal não são independentes. Adotando um nível de significância igual a 10%, se deve rejeitar a hipótese nula, ou seja, como o $P - Valor = 0,003 < \alpha = 0,01$, portanto, há uma relação de dependência entre os dados da série temporal *fuzzy* observada, quando o componente sazonal $s \leq 12$, ou seja, com uma periodicidade equivalente a cada 13 observações continuamente, este modelo de previsão se torna capaz de estimar com maior eficiência o número de matrículas, o que conseqüentemente possibilita a utilização do modelo FSARIMA(4; 1; 1)(1; 0; 1)₁₃ que foi identificado, mediante a utilização da função $W_{t,k,j}^{DFZ}$, como uma ferramenta para previsão das matrículas na Universidade do Alabama, dentro do período compreendido por este estudo, como mostra-se na Tabela (6.70).

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

Os resultados apresentados na Tabela (6.70), ratificam o excelente ajuste do modelo FSARIMA (4; 1; 1)(1; 0; 1)₁₃, algo que também pode ser constatado na Figura (6.21), pois em comparação com outros modelos propostos na literatura para os dados da Tabela (6.62), o modelo proposto neste trabalho alcançou o menor MASE (0,23), o que indica, além de um ajuste significativo dos coeficientes dos parâmetros do modelo indicado pela função $W_{i,k,j}^{DFZ}$, como foi comprovado nas Tabelas (6.68) e (6.69), o erro de previsão gerado por este modelo foi menor, que os obtidos pelos demais métodos apresentados na Tabela (6.70), os quais também são referentes as quantidades de matrícula na Universidade do Alabama entre 1971 e 1992, utilizando um modelo de série temporal *fuzzy* para a previsão desta variável, porém, mediante diferentes métodos de estimação dos parâmetros do modelo de série temporal *fuzzy*, assim como, incorporando funções de pertinência para ativar os conjuntos *fuzzy* correspondentes, não necessariamente provenientes da definição de nú-

Tabela 6.70 - Comparação dos resultados de previsões pelo método proposto neste trabalho (função de autocorrelação *defuzzificada*), em relação a 14 métodos de previsão fuzzy que encontram-se publicados na literatura (Song e Chisson, 1993b - Chen e Tanuwijaya, 2011).

Ano	Número de Matrículas	Song and Chisson (1993b)	Sullivan and Woodall (1994)	Chen (1996)	Huarng (2001a)	Chen (2002)	Lee and Chou (2004)	Tsaur et al. (2005)	Método Proposto: Treinamento	Método Proposto: Validação
1971	13055	-	-	-	-	-	-	-	-	-
1972	13563	14000	13500	14000	14000	-	-	-	13070,8	-
1973	13867	14000	14500	14000	14000	-	-	-	14086,9	-
1974	14696	14000	14500	14000	14000	14500	14568	14000	15150,0	-
1975	15460	15500	15231	15500	15500	15500	15654	15500	14909,8	-
1976	15311	16000	15563	16000	15500	15500	15654	15500	14728,0	-
1977	15603	16000	15563	16000	16000	15500	15654	16000	14958,9	-
1978	15861	16000	15500	16000	16000	15500	15654	16000	16211,3	-
1979	16807	16000	15500	16000	16000	16500	16197	16000	16896,3	-
1980	16919	16813	16684	16833	17500	16500	17283	16500	17223,8	-
1981	16388	16813	16684	16833	16000	16500	17283	16500	16202,0	-
1982	15433	16789	15500	16833	16000	15500	16197	15500	15153,9	-
1983	15497	16000	15563	16000	16000	15500	15654	15500	15167,1	-
1984	15145	16000	15563	16000	15500	15500	15654	15500	15132,9	-
1985	15163	16000	15563	16000	16000	15500	15654	15500	15155,4	-
1986	15984	16000	15563	16000	16000	15500	15654	15500	16186,4	-
1987	16859	16000	15500	16000	16000	16500	16197	16500	-	17058,3
1988	18150	16813	16577	16833	17500	18500	17283	18500	-	17908,4
1989	18970	19000	19500	19000	19000	18500	18369	19000	-	19184,5
1990	19328	19000	19500	19000	19000	19500	19454	19000	-	19302,7
1991	19337	19000	19500	19000	19500	19500	19454	19000	-	19076,7
1992	18876	-	-	19000	19000	18500	-	-	-	19329,1
MASE	-	1,01	0,86	0,98	0,77	0,50	0,81	0,59	-	0,23

Ano	Número de Matrículas	Singh (2007a)	Singh (2007b)	Liu (2007)	Cheng et al. (2008a)	Singh (2009)	Wong et al. (2010)	Chen e Tanuwijaya (2011)	Método Proposto: Treinamento	Método Proposto: Validação
1971	13055	-	-	-	-	-	-	-	-	-
1972	13563	-	-	13500	13680,5	-	-	13512	13070,8	-
1973	13867	-	-	13800	13731,3	13492	13500	13998	14086,9	-
1974	14696	14500	14286	14700	13761,7	14389	14500	14658	15150,0	-
1975	15460	15358	15361	15600	15194,6	15379	15500	15341	14909,8	-
1976	15311	15500	15468	15400	15374,8	15466	15466	15501	14728,0	-
1977	15603	15500	15512	15750	15359,9	15364	15392	15501	14958,9	-
1978	15861	15500	15582	15400	16410,3	15534	15549	15501	16211,3	-
1979	16807	16500	16500	16800	16436,1	16408	16433	17065	16896,3	-
1980	16919	16500	16361	17100	17130,7	16604	16656	17159	17223,8	-
1981	16388	16500	16362	17100	17141,9	16810	16624	17159	16202,0	-
1982	15433	15581	15744	15300	15363,8	15561	15556	15341	15153,9	-
1983	15497	15500	15560	15750	15372,1	15524	15524	15501	15167,1	-
1984	15145	15500	15498	15400	15378,5	15497	15497	15501	15132,9	-
1985	15163	15500	15306	15300	15343,3	15343	15305	15501	15155,4	-
1986	15984	15500	15442	15750	15345,1	15189	15308	15501	16186,4	-
1987	16859	16402	16558	16800	16448,4	16402	16402	17065	-	17058,3
1988	18150	18500	17187	17100	17135,9	18555	18500	17159	-	17908,4
1989	18970	18500	18475	18900	18915,0	18533	18534	18832	-	19184,5
1990	19328	19471	19382	19200	18997,0	19390	19345	19333	-	19302,7
1991	19337	19500	19487	19050	19032,8	19394	19423	19083	-	19076,7
1992	18876	18651	18744	19050	19033,7	18500	18752	19083	-	19329,1
MASE	-	0,51	0,56	0,43	0,67	0,58	0,48	0,50	-	0,23

mero *fuzzy* triangular, tal como foi implementado neste método proposto.

Destaca-se ainda que o bom ajuste dos dados de matrícula ao modelo identificado pela função EACF *fuzzy*, fica mais evidenciado caso se compare também os seus resultados, com o de outros métodos de previsão *fuzzy*, obtidos e validados a partir de outras métricas de acurácia, diferentes do MASE apresentado na Tabela (6.70), tal como, a proposta de previsão estabelecida por Egrioglu et al. (2011). Nesta proposta foi utilizada a medida de acurácia conhecida por raiz do erro médio quadrático (RMSE), com valor obtido para as previsões igual a 258,19, que representa um valor maior ao alcançado pelo método proposto neste estudo que foi de 221,42. Desta forma, se comprova novamente um ajuste significativo dos dados de matrícula da Universidade do Alabama, ao modelo de previsão *fuzzy* FSARIMA(4; 1; 1)(1; 0; 1)₁₃ identificado pelo método proposto neste trabalho, em virtude do resultado obtido com o RMSE neste estudo de caso.

Tabela 6.71 - Comparação dos resultados de previsões pelo método proposto neste trabalho (função de autocorrelação *defuzzificada*), em relação a 7 métodos de previsão *fuzzy* que encontram-se publicados na literatura (Gangwar e Kumar, 2012 - Lu et al., 2015).

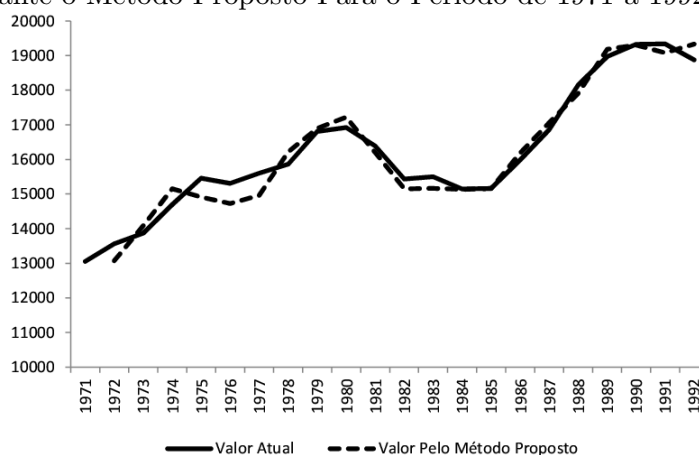
Ano	Número de Matrículas	Gangwar e Kumar (2012)	Chen and Kao (2013)	Wang et al. (2013)	Uslu et al. (2014)	Wang et al. (2014)	Lu et al. (2015) ($h = 7$)	Lu et al. (2015) ($h = 17$)	Método Proposto: (Treinamento)	Método Proposto: (Validação)
1971	13055	-	-	-	-	-	-	-	-	-
1972	13563	-	14347	13486	13650	13944	14279	13678	13070,8	-
1973	13867	13500	14347	14156	13650	13944	14279	13678	14086,9	-
1974	14696	14500	14347	15215	14836	13944	14279	14602	15150,0	-
1975	15460	15500	15550	15906	15332	15328	15392	15498	14909,8	-
1976	15311	15500	15550	15906	15447	15753	15392	15192	14728,0	-
1977	15603	15500	15550	15906	15447	15753	15392	15641	14958,9	-
1978	15861	15500	15550	15906	15447	15753	16467	15827	16211,3	-
1979	16807	-	16290	16559	16746	16279	16467	16744	16896,3	-
1980	16919	-	17169	16559	17075	17270	17161	17618	17223,8	-
1981	16388	16500	17169	16559	16380	17270	17161	16392	16202,0	-
1982	15433	15500	16290	16559	15457	16279	14916	15410	15153,9	-
1983	15497	15500	15550	15906	15447	15753	15392	15498	15167,1	-
1984	15145	15500	15550	15906	15447	15753	15392	15192	15132,9	-
1985	15163	15500	15550	15906	15332	15753	15392	15567	15155,4	-
1986	15984	-	15550	15906	16027	15753	15470	15567	16186,4	-
1987	16859	-	16290	16659	16746	16279	16467	16744	-	17058,3
1988	18150	18500	17179	16659	18211	17270	17161	17618	-	17908,4
1989	18970	18500	18907	19451	19059	19466	19257	19036	-	19184,5
1990	19328	19337	18907	18808	19059	18933	19257	19574	-	19302,7
1991	19337	19500	18907	18808	19059	18933	19257	19146	-	19076,7
1992	18876	18704	18907	18808	19059	18933	19257	19146	-	19329,1
MASE	-	0,40	0,79	0,88	0,29	0,85	0,72	0,35	-	0,23

Segundo Hyndman e Koehler (2006), quando o valor do MASE obtido é inferior a 1, se tem um indicativo forte que em média o erro de previsão retornado pelo modelo utilizado é baixo, o que contribui diretamente para um elevado desempenho da capacidade preditiva do modelo identificado. Neste contexto, o modelo de previsão encontrado pelo

método proposto neste trabalho para a quantidade de matrículas, se apresenta como uma ferramenta eficiente e com um considerável desempenho para prever a característica numérica estudada, face ao resultado de 0,23 do MASE (Tabela 6.71) retornado pelo modelo FSARIMA(4; 1; 1)(1; 0; 1)₁₃.

A geração de projeções por um modelo de série temporal *fuzzy*, ao qual a base de dados foi ajustada, se torna fundamental no processo de identificação, pois assim, o modelo identificado é validado, para conseqüentemente a partir das projeções realizadas, ser possível mensurar a capacidade preditiva do modelo utilizado, e então poder realizar projeções confiáveis, como se ver também em Kai e Wen-Gang (2009). Neste sentido, para comprovar a confiabilidade do modelo de previsão *fuzzy* identificado, seis projeções foram realizadas mediante este modelo de série temporal *fuzzy* identificado, onde se utilizou o conceito de intervalo de confiança *fuzzy*, tal como estabelecido por Wu (2009), com 95% de probabilidade de confiança para esta construção, com o intuito de ratificar uma razoável capacidade preditiva deste modelo, o que pode ser comprovado caso as projeções estejam localizadas dentro dos intervalos obtidos. Após a determinação do modelo de previsão via série temporal *fuzzy*, o qual foi obtido com auxílio da função de identificação proposta neste trabalho, se gerou projeções para 6 instantes de tempo t (1993-1998) subseqüentes a série histórica das matrículas na Universidade do Alabama, como mostra a Tabela (6.72), onde novamente constatou-se um excelente ajuste dos dados da Tabela (6.62), ao modelo identificado.

Figura 6.21 - Série Temporal Referente a Quantidade de Matrículas na Universidade do Alabama nos Estados Unidos no Período de 1971 a 1992 *Versus* a Série Temporal Referente aos Valores das Previsões Mediante o Método Proposto Para o Período de 1971 a 1992.



Pela definição de intervalo de confiança *fuzzy* dado por Wu (2009), se chega aos intervalos de confiança aproximados, mediante o teorema central do limite (TCL), com coeficiente de confiança dado por $1 - \alpha$, onde $0 \leq \alpha \leq 1$, seguindo as Equações (3.19) e (3.20).

Portanto, utilizando as Equações (3.19) e (3.20), para as seis projeções que foram realizadas em relação a quantidade de matrículas, a partir do modelo FSARIMA(4; 1; 1)(1; 0; 1)₁₃, se contruíram os intervalos de confiança *fuzzy* com 95% de confiança, para que desta forma seja possível verificar a capacidade preditiva do modelo identificado, em um horizonte de seis anos após o final da série em estudo. Então, para 95% de confiança, $Z_\alpha = 1,96$, onde $n = 7$, $\bar{x}_h^L = 15.151,43$, $\bar{x}_h^U = 17.241,43$ e $s_{hL} = s_{hU} = 2.260,93$, logo, para as seis projeções obtidas pelo método proposto neste trabalho, as quais estão contidas na Tabela 6.72, se tem os intervalos u_i e os intervalos *fuzzy* equivalentes para os vértices mínimo ($\tilde{\mu}_h^L$) e máximo ($\tilde{\mu}_h^U$), respectivamente.

Analisando os resultados da Tabela (6.72), para as projeções geradas aos instantes de tempo posteriores ao período utilizado para estimar o modelo SARIMA *fuzzy* que foi identificado, se percebe que estas projeções alcançaram um resultado razoável, pois se posicionaram dentro de um dos intervalos de confiança *fuzzy* com 95% de probabilidade, correspondentes aos conjuntos que foram construídos no processo de *fuzzificação* dos dados originais de matrículas, com exceção da projeção 19.109 para o número de matrículas no ano de 1995, que ficou fora do intervalo de confiança com 95% de probabilidade, mostrando com isso que, para este ano especificamente o modelo não conseguiu realizar uma projeção satisfatória, algo que pode ser explicado pela variabilidade que está inerente nestes dados de matrícula na Universidade do Alabama, especificamente no período que foi utilizado para determinar o modelo de série temporal *fuzzy* sazonal.

Tabela 6.72 - Resultados das Projeções Para a Quantidade de Matrículas na Universidade do Alabama nos Estados Unidos Pelo Método Proposto Entre os Anos de 1993 e 1998.

Ano	Projeção Via Método Proposto	Intervalo u_j	$I.C(\tilde{\mu}_h^L; 95\%)$	$I.C(\tilde{\mu}_h^U; 95\%)$
1993	18.140	[17.245; 19.335]	[13.485; 16.818]	[15.575; 18.908]
1994	18.642	[17.245; 19.335]	[13.485; 16.818]	[15.575; 18.908]
1995	19.109	[18.290; 20.380]	[13.485; 16.818]	[15.575; 18.908]
1996	18.286	[17.245; 19.335]	[13.485; 16.818]	[15.575; 18.908]
1997	17.943	[17.245; 19.335]	[13.485; 16.818]	[15.575; 18.908]
1998	17.807	[16.200; 18.290]	[13.485; 16.818]	[15.575; 18.908]

Passo 6: Escolher uma das três funções de autocorrelação estendida *fuzzy*, que estão sendo

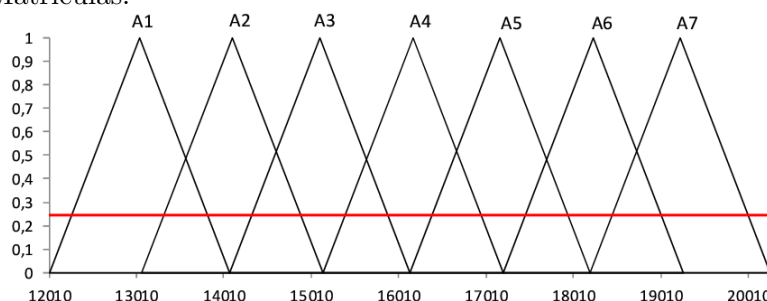
propostas neste trabalho na Seção 5.6, para identificar a ordem do modelo de série temporal *fuzzy*;

(b) Caso a função escolhida seja $W_{i,k,j}^{CA}(h; \alpha; b)$, a série temporal *crisp* deve ser submetida ao *alfa-nível*, com valor definido *a priori* pelo especialista, por exemplo, $\alpha = 0,25; 0,5$ ou $0,75$, para posteriormente realizar a seleção aleatória, a partir do suporte dos conjuntos *fuzzy* que foram construídos no passo 5, procedendo posteriormente com a aplicação da função de autocorrelação estendida *fuzzy*.

(b.1) Aplicar a função $W_{i,k,j}^{CA}(h; \alpha; b)$, a série temporal *crisp* submetida ao *alfa-nível* = $0,25$; para ativar os conjuntos *fuzzy* e selecionar no suporte de cada conjunto um elemento.

Como passo inicial se deve estabelecer o valor máximo de pertinência aos conjuntos *fuzzy* que foram construídos no passo 5. No caso do *alfa-nível* = $0,25$, a ativação dos conjuntos e conseqüente definição dos suportes de onde serão selecionadas as observações, para compor a série de dados se dará tal como, se observa na Figura (6.22), e ainda, os resultados do processo de *fuzzificação*, ou seja, intervalos correspondentes aos valores de matrículas, a série temporal *fuzzy*, relação lógica *fuzzy* e valores da série com *alfa-nível* = $0,25$, são dados na Tabela (6.73).

Figura 6.22 - Conjuntos *Fuzzy* Pela Definição de Número *Fuzzy* Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, Com Alfa-Nível Igual a $0,25$, onde o Eixo Y Representa a Pertinência ao Conjunto *Fuzzy* A_i e o Eixo X Indica a Quantidade de Matrículas.



Nas Tabelas (6.74) e (6.75), se tem os resultados da função de autocorrelação estendida *fuzzy* $W_{i,k,j}^{CA}(h; \alpha; b)$ e as probabilidades de significância, respectivamente, tal que, estes resultados foram obtidos durante a rotina de identificação da ordem na série temporal *fuzzy*, a partir da utilização da função de autocorrelação estendida, sobre a série resultante do processo de *fuzzificação* com *alfa-nível* = $0,25$, a qual pode ser observada diretamente na Tabela (6.73).

Analisando os resultados das Tabelas (6.74) e (6.75), se pode perceber que a função de autocorrelação estendida com o *alfa-nível* = 0,25, identificou de forma significativa (estatisticamente diferente de zero), um modelo ARMA(0; 0) e um modelo ARMA(1; 0), onde o valor de $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b) = -0,4729$ e $-0,4640$, respectivamente, com uma probabilidade de significância de 0,0302 e 0,0380, respectivamente, portanto, à um nível de significância $\alpha = 10\%$, onde estes ajustes são significativos, pois $\alpha = 0, 10 > P - Valor = 0,0302$ e $\alpha = 0, 10 > P - Valor = 0,0380$. Neste sentido, a função $W_{i,k,j}^{CA}(h; \alpha; b)$ identificou um modelo de série temporal *fuzzy*, sem o componente de médias móveis e autoregressivo, além de um outro modelo apenas com o componente autoregressivo de forma significativa, no caso um AR(1), resultado o qual difere do modelo MA(1) identificado pela função de autocorrelação amostral média com *alfa-nível*, apresentado por Song (2003), para esta mesma série de dados históricos.

Tabela 6.73 - Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de Fuzzificação Realizado A Partir de Um Número *Fuzzy* Triangular, Relação *Fuzzy* Lógica e os Valores de Matrículas Com Alfa-Nível de 0,25.

Ano	Valor de Matrícula	Intervalo Correspondente	Série Temporal <i>Fuzzy</i>	Relação <i>fuzzy</i> Lógica	Série Com $\alpha = 0, 25$
1971	13.055	[12.010; 14.100]	A_1	-	0,0035
1972	13.563	[13.060; 15.150]	A_2	$A_1 \rightarrow A_2$	0,0022
1973	13.867	[13.060; 15.150]	A_2	$A_2 \rightarrow A_2$	0,0041
1974	14.696	[14.105; 16.195]	A_3	$A_2 \rightarrow A_3$	0,0004
1975	15.460	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0050
1976	15.311	[15.150; 17.240]	A_4	$A_4 \rightarrow A_4$	0,0022
1977	15.603	[15.150; 17.240]	A_4	$A_4 \rightarrow A_4$	0,0009
1978	15.861	[14.105; 16.195]	A_3	$A_4 \rightarrow A_3$	0,0054
1979	16.807	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0033
1980	16.919	[15.150; 17.240]	A_4	$A_4 \rightarrow A_4$	0,0003
1981	16.388	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0001
1982	15.433	[15.150; 17.240]	A_4	$A_5 \rightarrow A_4$	0,0025
1983	15.497	[15.150; 17.240]	A_4	$A_4 \rightarrow A_4$	0,0008
1984	15.145	[14.105; 16.195]	A_3	$A_4 \rightarrow A_3$	0,0029
1985	15.163	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0022
1986	15.984	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0013
1987	16.859	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0036
1988	18.150	[17.245; 19.335]	A_6	$A_5 \rightarrow A_6$	0,0005
1989	18.970	[17.245; 19.335]	A_6	$A_6 \rightarrow A_6$	0,0013
1990	19.328	[18.290; 20.380]	A_7	$A_6 \rightarrow A_7$	0,0050
1991	19.337	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0005
1992	18.876	[17.245; 19.335]	A_6	$A_7 \rightarrow A_6$	0,0008

Logo, se pode acreditar que o fato da função $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b)$, ao estimar simultaneamente o componente autoregressivo e de médias móveis, em um modelo ARMA($p; q$),

Tabela 6.74 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b)$ para Matrículas na Universidade do Alabama.

<i>Lag</i>	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,4729	-0,1735	0,3084	-0,2043	-0,0057	0,0201
AR(1)	-0,4640	-0,1851	0,2099	-0,0936	-0,1855	0,0604
AR(2)	0,1821	-0,0145	0,2025	-0,1275	-0,1531	0,0657
AR(3)	-0,0774	-0,2159	-0,3527	-0,1797	-0,1880	-0,0986
AR(4)	-0,1473	-0,1125	-0,2842	-0,1953	-0,3209	-0,0699
AR(5)	0,3047	-0,0315	-0,2706	-0,3364	-0,1674	-

a torna mais eficiente e precisa, quanto ao processo de identificação da ordem de uma série temporal *fuzzy*, em relação ao método defendido por Song (2003), quando se utiliza a função de autocorrelação estendida, em uma série originada de um processo de *fuzzificação* mediante a definição de número *fuzzy* triangular, combinada a um *alfa-nível* = 0,25, para ativar os conjuntos *fuzzy* que indicarão os suportes correspondentes, para com isto, se tornar possível selecionar um valor específico dentro de cada intervalo u_i .

Tabela 6.75 - Probabilidades de Significância da $W_{i,k,j}^{CA}(h; \alpha = 0, 25; b)$ na Tabela (6.74).

<i>Lag</i>	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,0302*	0,5086	0,2496	0,4723	0,9844	0,9450
AR(1)	0,0380*	0,4867	0,4440	0,7390	0,5123	0,8281
AR(2)	0,4273	0,9495	0,4686	0,6552	0,5845	0,8335
AR(3)	0,7425	0,3613	0,1957	0,5176	0,5006	0,7339
AR(4)	0,5436	0,6984	0,3020	0,5552	0,2697	0,8357
AR(5)	0,2229	0,9179	0,3209	0,3130	0,5821	-

Nota: * Valor significativo para $\alpha = 10\%$.

A Tabela (6.76) apresenta os coeficientes do modelo ajustado FSARIMA(1; 1; 0)(1; 0; 0)₁₂, ou seja, um modelo autoregressivo sazonal *fuzzy*, o qual é semelhante ao modelo FSARIMA($p; d; q$)($P; D; Q$)_s, definido por Tseng e Tzeng (2002), porém, sem o termo de média móvel. Assim, na Tabela (6.76) também há os testes de significância dos coeficientes do modelo ajustado, mediante a estatística de teste oriunda da função densidade de probabilidade *t-Student*, com as respectivas probabilidades de significância (*P-Valor*). Na Tabela (6.77), tem-se a estatística de teste *Q* “*Ljung-Box*”, para verificar se os dados da série temporal *fuzzy* utilizada são independentes, o que impossibilitaria a utilização da teoria de série temporal, caso esta suposição não seja refutada, visto que, se deve garantir uma relação de dependência entre os dados que compõem a série temporal, construída com o auxílio do *alfa-nível* = 0,25, durante o processo de *fuzzificação*.

Tabela 6.76 - Teste de Significância dos Parâmetros do Modelo FSARIMA(1; 1; 0)(1; 0; 0)₁₂.

Modelo	Coefficiente	Erro Padrão	Estatística t	P -Valor
AR(1)	0,4807	0,1913	2,51	0,021*
SAR(12)	0,5679	0,2500	2,27	0,034*

Nota: * Valor significativo para $\alpha = 10\%$.

Tabela 6.77 - Teste Q Ljung-Box Para o Modelo FSARIMA(1; 1; 0)(1; 0; 0)₁₂.

Teste Q "Ljung-Box"	Lag	Qui-Quadrado	Grau de Liberdade	P -Valor
	12,000	13,300	10,000	0,205

Pelos resultados da Tabela (6.76), se pode concluir que o modelo ajustado é significativo, visto que, os testes para os coeficientes do modelo FSARIMA(1; 1; 0)(1; 0; 0)₁₂, a partir da estatística de teste t -Student são significativos, a um nível de significância de 10%, pois o $P - Valor = 0,021 < \alpha = 0,10$, tal como, $P - Valor = 0,034 < \alpha = 0,10$, logo, estes coeficientes são estatisticamente diferentes de zero, logo, a manutenção dos mesmos no modelo ajustado é válida.

Em relação ao teste Q Ljung-Box na Tabela (6.77), adotando um nível de significância igual a 10%, não se deve rejeitar a hipótese nula, ou seja, como o $P - Valor = 0,205 > \alpha = 0,10$, então, não há uma relação de dependência entre os dados da série temporal *fuzzy*, mediante o *alfa-nível* = 0,25, o que impossibilita a utilização do modelo FSARIMA(1; 1; 0)(1; 0; 0)₁₂, como ferramenta de previsão das matrículas na Universidade do Alabama, no período de 1971 a 1992.

Vale ressaltar que, como já foi discutido anteriormente, a não rejeição da hipótese H_0 no teste Q Ljung-Box compromete o ajuste do modelo identificado, pois a questão de dependência entre as observações que compõem uma série temporal, por si só, caracteriza uma condição de existência e também essencial para que esta metodologia possa ser aplicada, a uma característica numérica de interesse representada por uma série histórica, visto que, uma série temporal deve ser concebida como uma sequencia continua, em um intervalo de tempo ininterrupto, onde a observação $t - 2$ deve apresentar uma relação de dependência com a observação $t - 1$, onde por sua vez, t representa o instante de tempo observado na série temporal que se está analisando.

(b.2) Aplicar a função $W_{i,k,j}^{CA}(h; \alpha; b)$, a série temporal *crisp* submetida ao *alfa-nível* = 0,5; para ativar os conjuntos *fuzzy* e selecionar no suporte de cada conjunto um elemento; Caso *alfa-nível* = 0,50, a ativação dos conjuntos e a determinação dos suportes considera a Figura (6.23), Já a série de dados e os resultados da *fuzzificação* estão na Tabela (6.78).

Figura 6.23 - Conjuntos *Fuzzy* Pela Definição de Número *Fuzzy* Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, Com Alfa-Nível Igual a 0,50, onde o Eixo Y Representa a Pertinência ao Conjunto *Fuzzy* A_i e o Eixo X Indica a Quantidade de Matrículas.

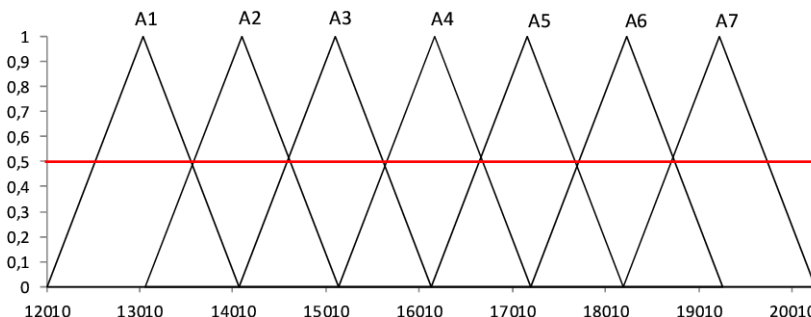


Tabela 6.78 - Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de Fuzzificação Realizado A Partir de Um Número *Fuzzy* Triangular, Relação *Fuzzy* Lógica e os Valores de Matrículas Com Alfa-Nível de 0,50.

Ano	Valor de Matrícula	Intervalo Correspondente	Série Temporal <i>Fuzzy</i>	Relação <i>fuzzy</i> Lógica	Série Com $\alpha = 0,50$
1971	13.055	[12.010; 14.100]	A_1	-	0,0047
1972	13.563	[12.010; 14.100]	A_1	$A_1 \rightarrow A_1$	0,0026
1973	13.867	[13.060; 15.150]	A_2	$A_1 \rightarrow A_2$	0,0038
1974	14.696	[14.105; 16.195]	A_3	$A_2 \rightarrow A_3$	0,0007
1975	15.460	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0026
1976	15.311	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0048
1977	15.603	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0003
1978	15.861	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0042
1979	16.807	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0060
1980	16.919	[16.200; 18.290]	A_5	$A_5 \rightarrow A_5$	0,0063
1981	16.388	[15.150; 17.240]	A_4	$A_5 \rightarrow A_4$	0,0046
1982	15.433	[14.105; 16.195]	A_3	$A_4 \rightarrow A_3$	0,0003
1983	15.497	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0044
1984	15.145	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0044
1985	15.163	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0004
1986	15.984	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0035
1987	16.859	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0036
1988	18.150	[17.245; 19.335]	A_6	$A_5 \rightarrow A_6$	0,0062
1989	18.970	[18.290; 20.380]	A_7	$A_6 \rightarrow A_7$	0,0027
1990	19.328	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0021
1991	19.337	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0053
1992	18.876	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0036

Nas Tabelas (6.79) e (6.80), se tem os resultados da função de autocorrelação estendida *fuzzy* $W_{i,k,j}^{CA}(h; \alpha; b)$ e as probabilidades de significância, respectivamente, tal que, estes resultados foram obtidos durante a rotina de identificação da ordem na série temporal *fuzzy*, a partir da utilização da função de autocorrelação estendida, sobre a série resultante do processo de *fuzzificação* com *alfa-nível* = 0,50, que pode ser observada na Tabela (6.78).

Tabela 6.79 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ para Matrículas na Universidade do Alabama.

<i>Lag</i>	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,4072	-0,2339	0,2068	-0,1140	0,3534	-0,5507
AR(1)	-0,5239	-0,5761	0,1636	0,0410	0,1548	-0,5834
AR(2)	-0,3538	-0,4133	0,1439	-0,0073	0,1704	-0,4839
AR(3)	-0,3431	-0,2870	0,1980	-0,0090	0,1701	-0,5468
AR(4)	0,6287	0,1227	0,1078	0,3527	0,5058	-0,5557
AR(5)	0,4620	-0,3563	-0,0528	-0,4082	0,3754	-

Tabela 6.80 - Probabilidades de Significância da $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ na Tabela (6.79).

<i>Lag</i>	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,0620*	0,3529	0,4298	0,6724	0,1938	0,0602*
AR(1)	0,0191*	0,0101*	0,5473	0,8770	0,5788	0,0543*
AR(2)	0,1231	0,0903*	0,6064	0,9790	0,5704	0,1088
AR(3)	0,1455	0,2523	0,4955	0,9762	0,5868	0,0992*
AR(4)	0,0095*	0,7164	0,7408	0,2695	0,1582	0,0758*
AR(5)	0,0646*	0,1621	0,8464	0,2122	0,2066	-

Nota: * Valor significativo para $\alpha = 10\%$.

Observando os resultados das Tabelas (6.79) e (6.80), se pode perceber que a função de autocorrelação estendida com o *alfa-nível* = 0,50, identificou de forma significativa (estatisticamente diferente de zero), cinco modelos ARMA($p; q$) considerando apenas os modelos onde $p > 0$ e $q > 0$, pois estes conseguem capturar de forma muito mais abrangente a variabilidade inerente aos dados observados. Neste sentido, a função de autocorrelação estendida $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ como mostram as Tabelas (6.79) e (6.80), os cinco modelos identificados de forma significativa ($P - Valor < \alpha = 0, 10$), ou seja, com valor da $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$ estatisticamente diferentes de zero, foram: ARMA(1; 1), ARMA(2; 1), ARMA(1; 5), ARMA(3; 5) e ARMA(4; 5). Portanto, em comparação com o resultado obtido por Song (2003), este autor identificou um modelo MA(1) mediante a função de autocorrelação amostral com *alfa-nível*, para a série do número de matrículas na Universidade do Alabama, resultado o qual é menos abrangente em termos de capacidade preditiva da série em estudo, em relação a um modelo ARMA($p; q$).

Dentre os modelo identificados pela função de autocorrelação estendida *fuzzy*, apenas o modelo ARIMA(1; 1; 1) apresentou significância estatística em seus parâmetros, a um nível de significância $\alpha = 10\%$, como se pode observa na Tabela (6.81), mediante os testes de significância dos coeficientes do modelo ajustado, pela estatística de teste *t-Student*, com as probabilidades de significância associadas, e ainda, o componente de sazonalidade para este modelo não se mostrou relevante para a composição do mesmo, por este motivo não houve inclusão do termo “s” correspondente. Na Tabela (6.82), a estatística de teste *Q “Ljung-Box”* para o modelo ARIMA(1; 1; 1), pode ser observada aos dados da série temporal *fuzzy*, onde se pode perceber que os dados da série utilizada não são independentes, logo, se garante uma relação de dependência entre os dados que compõem a série temporal, construída mediante um *alfa-nível* = 0,50.

Tabela 6.81 - Teste de Significância dos Parâmetros do Modelo ARIMA(1; 1; 1).

Modelo	Coefficiente	Erro Padrão	Estatística <i>t</i>	<i>P-Valor</i>
AR(1)	1,0702	0,4639	2,31	0,040*
MA(1)	1,9466	0,4802	4,05	0,002*

Nota: * Valor significativo para $\alpha = 10\%$.

Tabela 6.82 - Teste Q Ljung-Box Para o Modelo ARIMA(1; 1; 1).

Teste <i>Q “Ljung-Box”</i>	<i>Lag</i>	Qui-Quadrado	Grau de Liberdade	<i>P-Valor</i>
	12,000	17,700	3,000	0,001*

Nota: * Valor significativo para $\alpha = 10\%$.

Pelos resultados da Tabela (6.81), se pode concluir que o modelo ajustado é significativo, visto que, os testes para os coeficientes do modelo ARIMA(1; 1; 1), a partir da estatística de teste *t-Student* são significativos, a um nível de significância de 10%, pois o $P-Valor = 0,040 < \alpha = 0,10$, tal como, $P-Valor = 0,002 < \alpha = 0,10$, logo, estes coeficientes são estatisticamente diferentes de zero e portanto, devem ser mantidos no modelo identificado pela função de autocorrelação estendida *fuzzy*, com *alfa-nível* igual a 0,50.

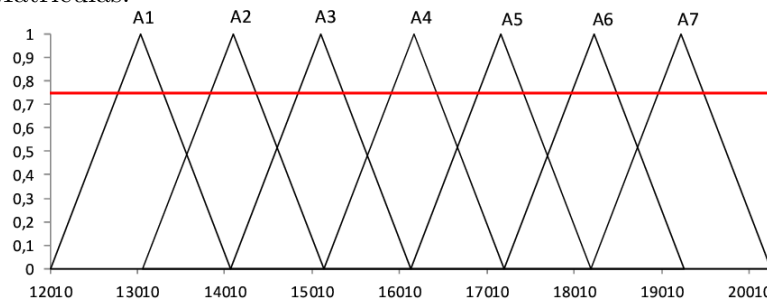
Em relação ao teste *Q Ljung-Box* na Tabela (6.82), adotando um nível de significância igual a 10%, se rejeita a hipótese nula de que, não há uma relação de dependência entre os dados da série temporal *fuzzy*, mediante o *alfa-nível* = 0,50, pois $P-Valor = 0,001 < \alpha = 0,10$, assim, o modelo ARIMA(1; 1; 1) pode ser utilizado como ferramenta de previsão das matrículas na Universidade do Alabama, dentro do período de 1971 a 1992, caso se tenha este interesse.

(b.3) Aplicar a função $W_{i,k,j}^{CA}(h; \alpha; b)$, a série temporal *crisp* submetida ao *alfa-nível* = 0,75; para ativar os conjuntos *fuzzy* e selecionar no suporte de cada conjunto um elemento.

Estabelecendo o valor máximo de pertinência aos conjuntos *fuzzy* que foram construídos no passo 5. Para o caso do *alfa-nível* = 0,75, a ativação dos conjuntos e consequente definição dos suportes de onde serão selecionadas as observações, para compor a série de dados, se dará tal como, se observa na Figura (6.24), e os resultados da *fuzzificação* estão na Tabela (6.83).

Mediante o processo de *fuzzificação* com a incorporação do *alfa-nível* realizado sobre os dados de matrículas (Tabelas 6.73, 6.78, 6.83), vários pares de relações *fuzzy* lógicas foram obtidos para os valores da série de dados, o que também resultou em grupos de relações *fuzzy* lógicas, para os três valores de *alfa-nível* utilizados, como mostra a Tabela (6.84). Pelos resultados dos grupos de relações *fuzzy* lógicas, se percebe que quando o *alfa-nível* foi igual a 0,50 ou 0,75, o processo de *fuzzificação* estabeleceu as mesmas transições entre os conjuntos *fuzzy* que foram construídos.

Figura 6.24 - Conjuntos *Fuzzy* Pela Definição de Número *Fuzzy* Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, Com Alfa-Nível Igual a 0,75, onde o Eixo Y Representa a Pertinência ao Conjunto *Fuzzy* A_i e o Eixo X Indica a Quantidade de Matrículas.



Portanto, o valor do *alfa-nível* igual a 0,50 ou 0,75 ativaram necessariamente os mesmos suportes associados aos dados de matrículas, após a partição do universo de discurso em 7 conjuntos *fuzzy*. Este fato possibilita afirmar que, para estes dados de matrículas especificamente, utilizar no processo de *fuzzificação* um valor de *alfa-nível* igual 0,50 ou 0,75 retornará aproximadamente os mesmos resultados, isto é, a ativação de conjuntos *fuzzy* idênticos e, portanto, basicamente será originada a mesma série temporal *fuzzificada*, a partir dos valores de *alfa-nível* supracitados.

Tabela 6.83 - Dados Históricos Referentes à Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Com o Processo de *Fuzzificação* Realizado A Partir de Um Número *Fuzzy* Triangular, Relação *Fuzzy* Lógica e os Valores de Matrículas Com Alfa-Nível de 0,75.

Ano	Valor de Matrícula	Intervalo Correspondente	Série Temporal <i>Fuzzy</i>	Relação <i>fuzzy</i> Lógica	Série Com $\alpha = 0,75$
1971	13.055	[12.010; 14.100]	A_1	-	0,0022
1972	13.563	[12.010; 14.100]	A_1	$A_1 \rightarrow A_1$	0,0012
1973	13.867	[13.060; 15.150]	A_2	$A_1 \rightarrow A_2$	0,0016
1974	14.696	[14.105; 16.195]	A_3	$A_2 \rightarrow A_3$	0,0046
1975	15.460	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0038
1976	15.311	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0028
1977	15.603	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0036
1978	15.861	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0084
1979	16.807	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0005
1980	16.919	[16.200; 18.290]	A_5	$A_5 \rightarrow A_5$	0,0012
1981	16.388	[15.150; 17.240]	A_4	$A_5 \rightarrow A_4$	0,0086
1982	15.433	[14.105; 16.195]	A_3	$A_4 \rightarrow A_3$	0,0022
1983	15.497	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0062
1984	15.145	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0018
1985	15.163	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0012
1986	15.984	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0044
1987	16.859	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0006
1988	18.150	[17.245; 19.335]	A_6	$A_5 \rightarrow A_6$	0,0009
1989	18.970	[18.290; 20.380]	A_7	$A_6 \rightarrow A_7$	0,0024
1990	19.328	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0039
1991	19.337	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0002
1992	18.876	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0007

Tabela 6.84 - Grupos de Relações *Fuzzy* Lógicas Ativadas Com Alfa-Nível de 0,25; 0,50 e 0,75.

Grupos	Relação <i>Fuzzy</i> ($\alpha = 0,25$)	Relação <i>Fuzzy</i> ($\alpha = 0,50$)	Relação <i>Fuzzy</i> ($\alpha = 0,75$)
Grupo 1	$A_1 \rightarrow A_2$	$A_1 \rightarrow A_1, A_1 \rightarrow A_2$	$A_1 \rightarrow A_1, A_1 \rightarrow A_2$
Grupo 2	$A_2 \rightarrow A_2, A_2 \rightarrow A_3$	$A_2 \rightarrow A_3$	$A_2 \rightarrow A_3$
Grupo 3	$A_3 \rightarrow A_3, A_3 \rightarrow A_4$	$A_3 \rightarrow A_3, A_3 \rightarrow A_4$	$A_3 \rightarrow A_3, A_3 \rightarrow A_4$
Grupo 4	$A_4 \rightarrow A_3, A_4, A_5$	$A_4 \rightarrow A_3, A_4 \rightarrow A_5$	$A_4 \rightarrow A_3, A_4 \rightarrow A_5$
Grupo 5	$A_5 \rightarrow A_4, A_5 \rightarrow A_6$	$A_5 \rightarrow A_4, A_5, A_6$	$A_5 \rightarrow A_4, A_5, A_6$
Grupo 6	$A_6 \rightarrow A_6, A_6 \rightarrow A_7$	$A_6 \rightarrow A_7$	$A_6 \rightarrow A_7$
Grupo 7	$A_7 \rightarrow A_6, A_7 \rightarrow A_7$	$A_7 \rightarrow A_7$	$A_7 \rightarrow A_7$

Nas Tabelas (6.85) e (6.86), se tem os resultados da função de autocorrelação estendida *fuzzy* $W_{i,k,j}^{CA}(h; \alpha; b)$ e as suas probabilidades de significância equivalentes, respectivamente, tal que, estes resultados foram obtidos durante a rotina de identificação do modelo de série temporal *fuzzy*, considerando $\alpha = 0,75$, a partir da série de dados (α -nível) observados na Tabela (6.83). Mediante as Tabelas (6.85) e (6.86), se pode perceber que a função

Tabela 6.85 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0, 75; b)$ para Matrículas na Universidade do Alabama.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,4344	-0,2524	0,2704	-0,0484	-0,0674	0,0033
AR(1)	-0,5171	-0,5058	0,2863	0,0017	0,0109	-0,0137
AR(2)	-0,5437	0,2198	0,1834	-0,3492	0,1862	-0,1286
AR(3)	0,1407	0,2156	0,1658	-0,5130	0,1673	-0,2754
AR(4)	0,1936	-0,1733	-0,0814	-0,4677	0,0431	-0,2076
AR(5)	0,2319	-0,3493	0,1800	-0,3147	0,1706	-

de autocorrelação estendida *fuzzy* $W_{t,k,j}^{CA}(h; \alpha = 0, 75; b)$, identificou de forma significativa cinco modelos ARMA($p; q$), ou seja, ARMA(0; 0), ARMA(1; 0), ARMA(2; 0), ARMA(1; 1) e ARMA(4; 3).

Tabela 6.86 - Probabilidades de Significância da $W_{i,k,j}^{CA}(h; \alpha = 0, 75; b)$ na Tabela (6.85).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,0465*	0,3245	0,3125	0,8630	0,8103	0,9905
AR(1)	0,0208*	0,0243*	0,3094	0,9951	0,9692	0,9617
AR(2)	0,0178*	0,4715	0,5611	0,2458	0,4661	0,6037
AR(3)	0,5505	0,5184	0,5826	0,1031	0,5170	0,2781
AR(4)	0,4248	0,5287	0,7642	0,0763*	0,8700	0,4753
AR(5)	0,3537	0,1640	0,5395	0,2474	0,5665	-

Nota: * Valor significativo para $\alpha = 10\%$.

No entanto, admitindo apenas os modelos onde $p > 0$ e $q > 0$, devido às propriedades em estimação paramétrica pontual, se deve adotar exclusivamente os modelos ARMA(1; 1) e ARMA(4; 3). Neste sentido, a identificação de forma significativa ($P\text{-Valor} < \alpha = 0, 10$), ou seja, com valor da $W_{i,k,j}^{CA}(h; \alpha = 0, 75; b)$ estatisticamente diferentes de zero, obtida neste estudo, difere em comparação ao resultado obtido por Song (2003), o qual identificou um modelo MA(1) pela função de autocorrelação amostral média com *alfa-nível*, o que sugere uma identificação subestimada de um ARMA($p; q$), pelo autor citado, então, segundo os resultados das Tabelas (6.85) e (6.86), o modelo MA(1) não reflete a verdadeira ordem dos parâmetros quanto à identificação da série temporal *fuzzy*, referente a quantidade de matrículas na Universidade do Alabama com um *alfa-nível* = 0,75.

Como foi possível constatar nas Tabelas (6.85) e (6.86), mediante a classe de modelos identificados pela função de autocorrelação estendida *fuzzy*, após algumas simulações de ajustes com o auxílio de um software estatístico a partir da identificação que foi obtida,

apenas o modelo ARIMA(1; 1; 3) apresentou parâmetros estatisticamente diferentes de zero a um nível de significância $\alpha = 10\%$, como se pode observar na Tabela (6.87), a qual mediante os testes de significância dos coeficientes do modelo ajustado, pela estatística de teste *t-Student*, com as respectivas probabilidades de significância, e ainda, com o componente de sazonalidade para este modelo que não se mostrou significativo (tanto para o termo “AR” quanto “MA”), por este motivo não houve inclusão do componente “s”. Os dados da série *fuzzy* utilizada não são independentes, pois, $P\text{-Valor} = 0,018 < \alpha = 0,10$, então, se pode garantir uma relação de dependência entre os dados que compõem a série temporal *fuzzy*, a partir da estatística de teste *Q* “Ljung-Box” para o modelo ARIMA(1; 1; 3), o qual foi obtido mediante a série construída para a quantidade de matrículas com um *alfa-nível* = 0,75.

Tabela 6.87 - Teste de Significância dos Parâmetros do Modelo ARIMA(1; 1; 3).

Modelo	Coefficiente	Erro Padrão	Estatística <i>t</i>	<i>P-Valor</i>
AR(2)	-0,8316	0,4423	-1,88	0,085*
MA(1)	1,4017	0,4157	3,37	0,006*
MA(2)	-1,4806	0,2997	-4,94	0,000*
MA(3)	0,6908	0,3239	2,13	0,054*

Nota: * Valor significativo para $\alpha = 10\%$.

Para o passo 6, no ítem **b**, subitens **(b.1)**, **(b.2)** e **(b.3)** foi possível perceber que a ferramenta de identificação proposta neste trabalho, quando combinada com o *alfa-nível* assumindo os valores 0,25; 0,50 e 0,75; para ativar os conjuntos *fuzzy* e indicar o suporte dos conjuntos, de onde foram selecionados os pontos para compor a série temporal *fuzzy*, que a rotina de identificação apresentou resultados razoáveis, sobretudo, quando o *alfa-nível* = 0,5 ou 0,75, pois com estes valores a identificação foi plena, ou seja, identificou o componente AR e MA no modelo ARIMA($p; d; q$), algo que não ocorreu quando *alfa-nível* = 0,25, onde o melhor modelo identificado foi AR(1).

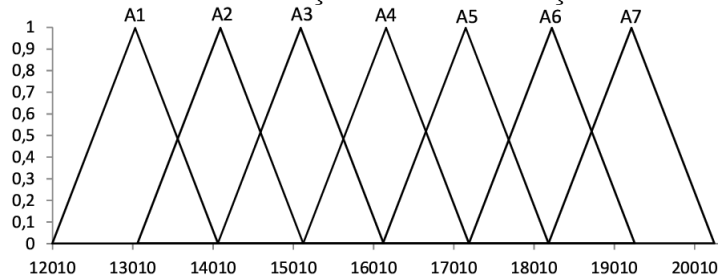
Passo 6: Escolher uma das três funções de autocorrelação estendida *fuzzy*, que estão sendo propostas neste trabalho na Seção 5.6, para identificar a ordem do modelo de série temporal *fuzzy*;

(c) Caso a função escolhida seja $\overline{W}_{i,k,j}^{RA}(h)$, se faz necessário proceder a seleção da função de autocorrelação média aleatória, a partir de cada suporte dos conjuntos *fuzzy* ativados, que foram construídos no passo 5, para em seguida aplicar a função de autocorrelação estendida *fuzzy* sobre a série obtida.

Para calcular a função $\overline{W}_{i,k,j}^{RA}(h)$, primeiramente se deve obter a série de valores da função de autocorrelação amostral média aleatória, mediante a construção dos conjuntos *fuzzy*, com base na definição de número *fuzzy* triangular, tal como, se observa na Figura (6.25). Posteriormente, se obtém a pertinência máxima de cada valor das matrículas ao conjunto *fuzzy* correspondente, o qual será ativado e indicará o suporte de onde será selecionado o ponto para compor, a série temporal dos valores da função de autocorrelação amostral (*fac*) média aleatória. A partir deste estágio, em cada suporte correspondente ao conjunto *fuzzy* ativado, 1.000 (um mil) valores serão simulados, para os quais serão calculados a *fac* e, em seguida, obtida a média destes 1.000 valores da *fac* em cada suporte. Então, no intervalo equivalente ao conjunto *fuzzy* ativado pela pertinência máxima, a média em cada um dos “*n*” intervalos irá compor a série das funções de autocorrelação amostral média aleatória, como mostra a Tabela (6.89), o que possibilitará o cálculo da função de autocorrelação estendida média aleatória *fuzzy* $\overline{W}_{i,k,j}^{RA}(h)$, sobre estes valores em um segundo momento, com o intuito de identificar o modelo de série temporal *fuzzy*, que será resultante da rotina proposta neste trabalho.

É importante destacar que, a rotina para a obtenção da série de valores da *fac* média aleatória em cada suporte, que está sendo adotada neste estudo, segue os procedimentos adotados por Song (2003), pois o objetivo será comparar a eficiência da ferramenta de identificação $\overline{W}_{i,k,j}^{RA}(h)$, que está sendo proposta neste trabalho, a partir da definição de número *fuzzy* triangular, com a ferramenta de identificação que foi proposta por Song (2003), onde este autor utilizou a função de pertinência trivalente no processo de *fuzzificação* dos dados.

Figura 6.25 - Conjuntos *Fuzzy* Pela Definição de Número *Fuzzy* Triangular, Para a Quantidade de Matrículas na Universidade do Alabama no Período de 1971 a 1992, onde o Eixo Y Representa a Pertinência ao Conjunto *Fuzzy* A_j e o Eixo X Indica a Quantidade de Matrículas, Para Auxiliar na Construção da Série dos Valores da Função de Autocorrelação Amostral Média Aleatória.



O processo de *fuzzificação* realizado sobre os valores das *fac*'s médias aleatórias (Tabela

6.89), indicaram relações *fuzzy* lógicas semelhantes às relações dos dados *defuzzificados*, logo, com os mesmos pares de relações entre os conjuntos *fuzzy* ativados anteriormente (Tabela 6.62). Neste contexto, os grupos destas relações *fuzzy* lógicas, também são idênticos aos observados na (Tabela 6.63), devido à transição entre os conjuntos ter sido a mesma, como observado na Tabela (6.88).

Tabela 6.88 - Grupos de Relações *Fuzzy* Lógicas Ativadas Para os Dados da Tabela (6.89).

Grupo Ativado	Relações <i>Fuzzy</i> Lógicas
Grupo 1	$A_1 \rightarrow A_1, A_1 \rightarrow A_2$
Grupo 2	$A_2 \rightarrow A_3$
Grupo 3	$A_3 \rightarrow A_3, A_3 \rightarrow A_4$
Grupo 4	$A_4 \rightarrow A_3, A_4 \rightarrow A_5$
Grupo 5	$A_5 \rightarrow A_4, A_5 \rightarrow A_5, A_5 \rightarrow A_6$
Grupo 6	$A_6 \rightarrow A_7$
Grupo 7	$A_7 \rightarrow A_7$

A partir dos resultados das Tabelas (6.90) e (6.91), caso se adote o critério de $p > 0$ e $q > 0$, a identificação obtida sugere um processo ARMA(1; 1), onde o valor da função $\overline{W}_{i,k,j}^{RA}(h) = -0,5563$, com probabilidade de significância de 0,0129, o que resulta em uma estimação paramétrica pontual estatisticamente diferente de zero, mediante um nível de significância $\alpha = 10\%$, pois, como $P - \text{valor} = 0,0129 < \alpha = 0,10$, o que leva a rejeitar a hipótese de nulidade do resultado obtido na Tabela (6.90), e portanto, se deve aceitar a hipótese de que este ajuste é significativo à 10%, sendo que, esta ordem identificada coincidentemente é a mesma encontrada pela função $W_{i,k,j}^{CA}(h; \alpha; b)$, algo que pode ser explicado mediante os resultados das simulações realizadas no Capítulo 6, deste trabalho, onde se observou que para uma série de baixa duração ($T < 24$), com 7 conjuntos *fuzzy*, o desempenho quanto à identificação da ordem de uma série temporal *fuzzy*, se aproxima bastante, quando a função utilizada é $W_{t,k,j}^{CA}(h; \alpha; b)$ ou $\overline{W}_{i,k,j}^{RA}(h)$.

Diferentemente da identificação de um modelo ARMA(1; 1), que foi obtida com a ferramenta $\overline{W}_{i,k,j}^{RA}(h)$ neste estudo, Song (2003) chegou a um modelo MA(2) em sua rotina de identificação. Este resultado pode ser explicado devido ao fato, deste autor ter utilizado a função de autocorrelação amostral, para identificar o modelo de série temporal, no entanto, como já foi destacado anteriormente, esta função possui capacidade de identificação reduzida, em comparação com a função de autocorrelação estendida, que indica exatamente a ordem do processo ARMA ($p; q$).

Tabela 6.89 - Dados Históricos da Quantidade de Matrículas na Universidade do Alabama entre 1971 e 1992, Processo de *Fuzzificação* Realizado Por Um Número *Fuzzy* Triangular, Relação *Fuzzy* Lógica e a Série dos Valores da Função de Autocorrelação Amostral Média Aleatória.

Ano	Valor de Matrícula	Pertinência Máxima	Intervalo Correspondente	Série Temporal <i>Fuzzy</i>	Relação <i>fuzzy</i> Lógica	<i>fac</i> Média Aleatória
1971	13.055	0,995	[12.010; 14.100]	A_1	-	0,0011
1972	13.563	0,516	[12.010; 14.100]	A_1	$A_1 \rightarrow A_1$	0,0044
1973	13.867	0,776	[13.060; 15.150]	A_2	$A_1 \rightarrow A_2$	0,0021
1974	14.696	0,566	[14.105; 16.195]	A_3	$A_2 \rightarrow A_3$	0,0029
1975	15.460	0,703	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0023
1976	15.311	0,846	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0012
1977	15.603	0,566	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0025
1978	15.861	0,677	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0031
1979	16.807	0,584	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0029
1980	16.919	0,691	[16.200; 18.290]	A_5	$A_5 \rightarrow A_5$	0,0010
1981	16.388	0,819	[15.150; 17.240]	A_4	$A_5 \rightarrow A_4$	0,0003
1982	15.433	0,729	[14.105; 16.195]	A_3	$A_4 \rightarrow A_3$	0,0011
1983	15.497	0,668	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0022
1984	15.145	0,995	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0005
1985	15.163	0,988	[14.105; 16.195]	A_3	$A_3 \rightarrow A_3$	0,0046
1986	15.984	0,794	[15.150; 17.240]	A_4	$A_3 \rightarrow A_4$	0,0016
1987	16.859	0,634	[16.200; 18.290]	A_5	$A_4 \rightarrow A_5$	0,0006
1988	18.150	0,866	[17.245; 19.335]	A_6	$A_5 \rightarrow A_6$	0,0023
1989	18.970	0,648	[18.290; 20.380]	A_7	$A_6 \rightarrow A_7$	0,0059
1990	19.328	0,989	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0014
1991	19.337	0,997	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0008
1992	18.876	0,558	[18.290; 20.380]	A_7	$A_7 \rightarrow A_7$	0,0009

Nota: *fac* (função de autocorrelação amostral).

Tabela 6.90 - Resultados da $\overline{W}_{i,k,j}^{RA}(h)$ para as Matrículas na Universidade do Alabama.

<i>Lag</i>	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,3146	-0,3554	0,2838	-0,0959	-0,1073	0,2288
AR(1)	-0,4447	-0,5563	0,3567	-0,0733	-0,0741	0,2170
AR(2)	-0,0015	-0,3667	0,4071	-0,1278	-0,1821	0,2879
AR(3)	-0,1496	-0,1922	0,1840	-0,0112	-0,2092	0,1933
AR(4)	-0,3678	0,0424	0,0608	-0,1730	-0,2052	0,2670
AR(5)	0,6469	0,3746	-0,1965	-0,1229	-0,1797	-

Os resultados apresentados na Tabela (6.92) indicam que o parâmetro AR(1) não é significativo, ou seja, o valor igual a -0,1043 do coeficiente, que representa uma estimativa para o parâmetro AR(1) é igual a zero, visto que, como a probabilidade de significância desta estimativa é $P\text{-Valor} = 0,684 > \alpha = 0,10$, se deve tomar a decisão de não rejeitar a hipótese nula (inicial), o que implica diretamente em retirar do modelo identificado ARMA(1; 1), este termo específico, logo, este modelo se reduz a um MA(1), pois a es-

timativa deste parâmetro é 0,9371 com $P\text{-Valor} = 0,000 < \alpha = 0,10$, logo, rejeita-se a hipótese nula e o termo é mantido no modelo estimado.

Tabela 6.91 - Probabilidades de Significância de $\overline{W}_{i,k,j}^{RA}(h)$ na Tabela (6.90).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,1494	0,1367	0,2803	0,7292	0,7001	0,4147
AR(1)	0,0467*	0,0129*	0,1983	0,7900	0,7871	0,4547
AR(2)	0,9949	0,1395	0,1657	0,6835	0,5635	0,3938
AR(3)	0,5255	0,4335	0,5181	0,9732	0,5545	0,5628
AR(4)	0,1294	0,8916	0,8474	0,6293	0,5736	0,4519
AR(5)	0,0097*	0,2157	0,5914	0,7193	0,6042	-

Nota: * Valor significativo para $\alpha = 10\%$.

Quanto a estatística de teste Q *Ljung-Box*, se pode comprovar mediante os resultados da Tabela (6.93), que há independência entre os valores da série utilizada pela rotina de identificação $\overline{W}_{i,k,j}^{RA}(h)$ proposta neste trabalho, pois, como $P\text{-Valor} = 0,363 > \alpha = 0,10$, não se deve rejeitar a hipótese inicial, então os dados que compõem a série são independentes entre si, o que inviabiliza a utilização destes dados sob a metodologia de série temporal, em virtude da continuidade e relação de dependência entre os dados em um instante i e $i + 1$, por exemplo, não serem atendidas, comprometendo assim o processo para realização de previsões confiáveis da característica de interesse, mediante o modelo de série temporal *fuzzy* que foi estimado.

Tabela 6.92 - Teste de Significância dos Parâmetros do Modelo ARIMA(1; 1; 1).

Modelo	Coefficiente	Erro Padrão	Estatística t	$P\text{-Valor}$
AR(1)	-0,1043	0,2524	-0,41	0,684
MA(1)	0,9371	0,1830	5,12	0,000*

Nota: * Valor significativo para $\alpha = 10\%$.

Tabela 6.93 - Teste Q *Ljung-Box* Para o Modelo ARIMA(1; 1; 1).

Teste Q “ <i>Ljung-Box</i> ”	Lag	Qui-Quadrado	Grau de Liberdade	$P\text{-Valor}$
	12,000	10,900	10,000	0,363

A determinação da ordem na série temporal *fuzzy*, que representa a quantidade de matrículas na Universidade do Alabama no período de 1971 a 1992, ao utilizar a ferramenta $\overline{W}_{i,k,j}^{RA}(h)$ apresentou resultados satisfatórios, em virtude da significância estatística do valor da função de autocorreção estendida, ao contrário de um dos parâmetros do modelo ARMA(1; 1), que não foi estatisticamente diferente de zero, para $\alpha = 10\%$, fato o qual pode ser justificado, em virtude da rotina de obtenção dos valores em cada suporte dos

conjuntos *fuzzy* ativados, considerar a média aleatória das *fac*, a qual representa um valor com ordem de grandeza muito inferior aos dados de matrículas originais. Neste sentido, a rotina para obtenção da série *fac* média aleatória pode ter influenciado, na significância estatística dos parâmetros identificados pela função $\overline{W}_{i,k,j}^{RA}(h)$.

6.1.4 - Estudo de caso 4: Previsão fuzzy para o índice TAIFEX (Taiwan Futures Exchange)

. Finalmente, nesta seção a rotina de identificação e previsão *fuzzy* que está sendo proposta neste trabalho, será avaliada mediante um último estudo de caso. A base de dados do índice TAIFEX (*Taiwan Futures Exchange*), que se encontra disponível em Wang e Chen (2009), quanto à metodologia de previsão em série temporal, neste estudo de caso servirá novamente para avaliar o algoritmo de previsão proposto no Capítulo 4.

Passo 1: Coletar os dados de interesse que formarão a série temporal [Wang e Chen (2009)].

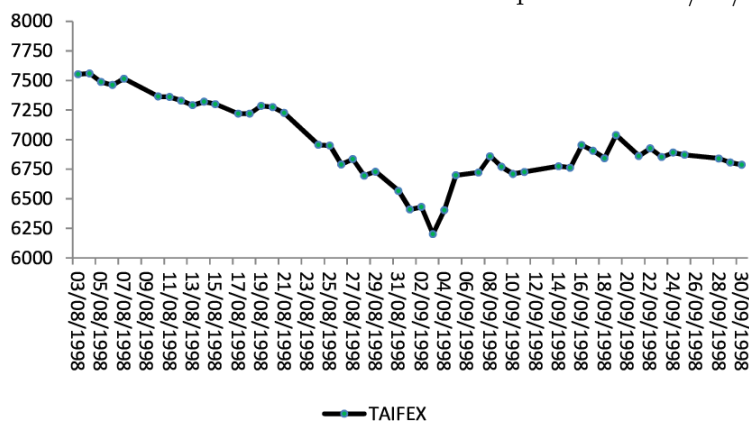
Passo 2: Construir o gráfico da série temporal aos dados do índice TAIFEX (ver, Figura 6.26).

Passo 3: Definir o universo de discurso. Mediante os dados: $Z_{(Min)} = 6200$ e $Z_{(Max)} = 7560$, e ainda, $D_1 = 227$ e $D_2 = 227$. Portanto, o universo de discurso é igual a $U = [5973; 7787]$.

Passo 4: Particionar o universo de discurso em “*m*” conjuntos *fuzzy*.

Pela rotina de identificação na Figura (6.13), como a trajetória $T > 36$ ($T = 47$), o valor de *m* é igual a: 5, 7 ou 9. Segundo Miller (1994), o valor de $m = 7$ é o indicado e será adotado.

Figura 6.26 - Série do índice TAIFEX durante 47 dias no período de 03/08/1998 a 30/09/1998.



Passo 5: Construir os conjuntos *fuzzy* correspondentes a quantidade estabelecida no Passo 4 e *fuzzificar* a série histórica a partir da definição de número *fuzzy* triangular.

Seja A_j a representação do j -ésimo conjunto *fuzzy*, onde $j = 1; 2; \dots; 6; 7$. Logo, caso A_j seja um número *fuzzy* triangular, se chega que, $u_j = [a_j; b_j; c_j]$. Desta forma, os intervalos são:

$$\begin{aligned} u_1 &= [5973; 6200; 6427], u_2 = [6200; 6427; 6653], u_3 = [6427; 6653; 6880], u_4 = [6653; 6880; 7107], \\ u_5 &= [6880; 7107; 7333], u_6 = [7107; 7333; 7560], u_7 = [7333; 7560; 7787]. \end{aligned}$$

Cada um dos sete intervalos resultantes de $m = 7$, em relação aos conjuntos A_j (número *fuzzy* triangular), representam a base de um dos sete triângulos construídos e observados na Figura (6.27). Pela Equação (6.1), os valores da Tabela (6.94), referentes a pertinência máxima em relação aos conjuntos *fuzzy* A_j , se tornaram possíveis de obter. Também são encontradas: as relações *fuzzy* lógicas simples; relações *fuzzy* lógicas agrupadas e a série temporal dos valores *defuzzificados* do TAIFEX, na Tabela (6.94). É importante ressaltar que a *defuzzificação* da série temporal *fuzzy*, se deu mediante o método centróide. A escolha deste método de *defuzzificação* é justificada, por se acreditar que a série resultante deste processo preservará com maior fidelidade, o padrão original dos dados observados e que auxiliarão o ajuste do modelo de série temporal.

Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para identificar a ordem do modelo de série temporal *fuzzy*, sobre os dados históricos de interesse.

(a) Utilizar a função de autocorrelação estendida *fuzzy* $W_{i,k,j}^{DF_z}$ sobre a série *defuzzificada*, a partir da série histórica *defuzzificada* mediante o método centróide.

As Tabelas (6.95 e 6.96), indicam o teste da raiz unitária de *Dickey-Fuller*, para avaliar a estacionariedade da série TAIFEX. É possível verificar que o teste estatístico antes da diferenciação da série histórica (45 dados), não apresenta significância estatística a 10%, pois o p -valor = 0,32634. No entanto, ao diferenciar a série temporal em 1 lag (44 pontos), os dados apresentados na Tabela (6.96), passam a apresentar estacionariedade no período estudado, visto que, p -valor = 0,00014 < 10%. Neste contexto, a função de autocorrelação estendida para série temporal do índice TAIFEX, pode ser plenamente utilizada na identificação paramétrica, por se tratar de uma série temporal estacionária, tal como, a literatura científica desta metodologia necessita.

Tabela 6.94 - Série histórica do índice TAIFEX em 3 de Agosto de 1998 a 30 de Setembro de 1998, com processo de *fuzzificação* utilizando a função de pertinência triangular, relações *fuzzy* lógicas simples, relações *fuzzy* lógicas agrupadas e a série temporal *defuzzificada* do índice TAIFEX.

Data de verificação	TAIFEX Original	Pertinência máxima	Intervalo fuzzy (u_j)	Conjunto fuzzy	Relações fuzzy lógicas	Grupos de relações fuzzy lógicas	Série defuzzificada
03/08/1998	7.552,00	0,965	[7.333; 7.787]	A_7	-	$A_1 \rightarrow A_2$	7.544,30
04/08/1998	7.560,00	1,000	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_2 \rightarrow A_1, A_2, A_3$	7.553,50
05/08/1998	7.487,00	0,678	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_3 \rightarrow A_2, A_3, A_4$	7.568,20
06/08/1998	7.462,00	0,568	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_4 \rightarrow A_3, A_4, A_5$	7.569,30
07/08/1998	7.515,00	0,802	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_5 \rightarrow A_4, A_5, A_6$	7.555,10
10/08/1998	7.365,00	0,859	[7.107; 7.560]	A_6	$A_7 \rightarrow A_6$	$A_6 \rightarrow A_4, A_5, A_6$	7.339,30
11/08/1998	7.360,00	0,881	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$	$A_7 \rightarrow A_6, A_7$	7.327,00
12/08/1998	7.330,00	0,987	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		7.323,90
13/08/1998	7.291,00	0,814	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		7.332,90
14/08/1998	7.320,00	0,942	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		7.325,10
15/08/1998	7.300,00	0,854	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		7.331,00
17/08/1998	7.219,00	0,504	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		7.109,90
18/08/1998	7.220,00	0,501	[6.880; 7.333]	A_5	$A_5 \rightarrow A_5$		7.108,90
19/08/1998	7.285,00	0,788	[7.107; 7.560]	A_6	$A_5 \rightarrow A_6$		7.333,20
20/08/1998	7.274,00	0,739	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		7.340,40
21/08/1998	7.225,00	0,522	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		7.336,70
24/08/1998	6.955,00	0,670	[6.653; 7.107]	A_4	$A_6 \rightarrow A_4$		6.889,10
25/08/1998	6.949,00	0,696	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.883,30
26/08/1998	6.790,00	0,604	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.891,40
27/08/1998	6.835,00	0,802	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.886,20
28/08/1998	6.695,00	0,815	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		6.651,70
29/08/1998	6.728,00	0,670	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		6.654,20
31/08/1998	6.566,00	0,615	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		6.656,00
01/09/1998	6.409,00	0,921	[6.200; 6.653]	A_2	$A_3 \rightarrow A_2$		6.435,30
02/09/1998	6.430,00	0,987	[6.200; 6.653]	A_2	$A_2 \rightarrow A_2$		6.426,50
03/09/1998	6.200,00	1,000	[5.973; 6.427]	A_1	$A_2 \rightarrow A_1$		6.184,30
04/09/1998	6.403,20	0,895	[6.200; 6.653]	A_2	$A_1 \rightarrow A_2$		6.427,10
05/09/1998	6.697,50	0,804	[6.427; 6.880]	A_3	$A_2 \rightarrow A_3$		6.661,00
07/09/1998	6.722,30	0,695	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		6.657,90
08/09/1998	6.859,40	0,909	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		6.878,70
09/09/1998	6.769,60	0,514	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.887,00
10/09/1998	6.709,75	0,750	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		6.653,60
11/09/1998	6.726,50	0,676	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		6.644,70
14/09/1998	6.774,55	0,535	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		6.888,90
15/09/1998	6.762,00	0,520	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		6.648,90
16/09/1998	6.952,75	0,680	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		6.882,30
17/09/1998	6.906,00	0,885	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.881,40
18/09/1998	6.842,00	0,833	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.885,70
19/09/1998	7.039,00	0,700	[6.880; 7.333]	A_5	$A_4 \rightarrow A_5$		7.106,90
21/09/1998	6.861,00	0,916	[6.653; 7.107]	A_4	$A_5 \rightarrow A_4$		6.883,80
22/09/1998	6.926,00	0,797	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.891,50
23/09/1998	6.852,00	0,877	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.881,60
24/09/1998	6.890,00	0,956	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.872,70
25/09/1998	6.871,00	0,960	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.869,30
28/09/1998	6.840,00	0,824	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.887,90
29/09/1998	6.806,00	0,674	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.885,60
30/09/1998	6.787,00	0,590	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		6.879,70

Após a comprovação da estacionariedade da série temporal do índice TAIFEX. Nas Tabelas (6.97 e 6.98), a função $W_{i,k,j}^{DFZ}$ foi utilizada indicando uma identificação do modelo ARMA(5;4), de maneira significativa a 10%, como sendo o modelo de maior ordem, que satisfaz a restrição para $p > 0$ e $q > 0$. Há outros modelos identificados pela função e

que também satisfazem a restrição acima. O processo de otimização do ajuste estrutural do modelo de série temporal, também irá avaliar o ajuste dos modelos ARMA(3;1) e ARMA(4;1), onde o critério de decisão para escolha do modelo ótimo, necessariamente passa pelo modelo com menor erro de previsão.

Figura 6.27 - Representação gráfica dos conjuntos *fuzzy*, onde o eixo X representa os valores do índice TAIEX no período de 03/08 a 30/09/1998 *versus* o eixo Y pertinência ao conjunto A_i .

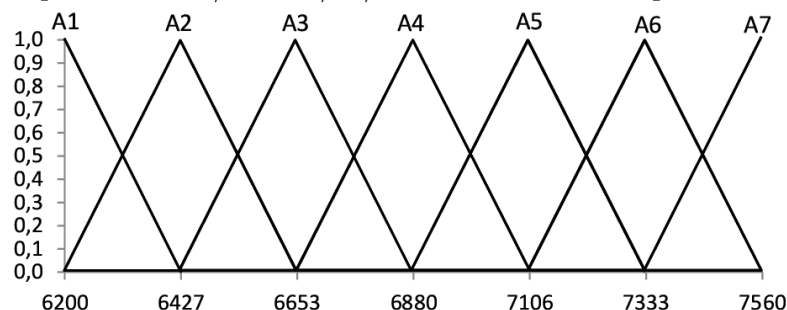


Tabela 6.95 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta Y_t = \beta_1 + \delta Y_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	Tau	P-valor
1	OLS	21042,12	849,25	45	-1,90726	0,32634
2	COV	21042,12	206946,96	45	-1,90726	0,32634
3	COV	21042,12	-29,64	45	-1,90726	0,32634
4	COV	21042,12	10,28	45	-1,90726	0,32634

Tabela 6.96 - Teste da raiz unitária de Dickey-Fuller aumentado no modelo $\Delta X_t = \beta_1 + \delta X_{t-1} + a_t$

Observação	Tipo	MSE	Intercepto	Número de dados	Tau	P-valor
1	OLS	23160,29	-19,922	44	-5,41108	0,00014*
2	COV	23160,29	539,109	44	-5,41108	0,00014*
3	COV	23160,29	0,846	44	-5,41108	0,00014*
4	COV	23160,29	-0,421	44	-5,41108	0,00014*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Tabela 6.97 - Resultados da função $W_{i,k,j}^{DFZ}$ para o índice TAIEX.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,1660	-0,0694	0,0425	-0,1717	0,1448	0,0390
AR(1)	-0,4452	-0,1469	-0,0150	-0,0917	0,1634	0,0918
AR(2)	0,1324	-0,1566	0,0622	-0,0692	0,1310	-0,0068
AR(3)	0,0758	-0,4283	-0,2307	-0,1294	0,1255	0,0005
AR(4)	0,4252	0,3630	0,2187	0,0376	0,0702	0,0533
AR(5)	-0,4277	0,2300	0,1830	-0,0105	0,3177	-0,0121

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6 (a);

Pela rotina de estimação do modelo de previsão *fuzzy* para o índice TAIEX, mediante o menor erro de previsão retornado, entre os modelos otimizados computacionalmente. Os

dados do índice TAIFEX aderiram de forma mais significativa, isto é, com maior acurácia das previsões, ao modelo FSARIMA(2;1;1)(1;0;1)₃₆, como mostra a Tabela (6.99). O teste de significância a 10%, indica exclusivamente o parâmetro de média móvel de ordem 2 (MA(2)), como não significativo, pois $p - valor = 0,143 > 10\%$. Portanto, mediante este resultado se pode concluir que o modelo de previsão ajustado, ao ser utilizado para prever o índice TAIFEX, irá desconsiderar o parâmetro $q = 2$ na estrutura física do modelo, garantindo assim maior capacidade preditiva desta ferramenta de previsão. Quanto ao teste de significância do componente sazonal ($s = 36$), este não é estatisticamente significativo, devido $p - valor = 0,047 > 1\%$, como se pode observar na Tabela (6.100). Então, o termo sazonal $s = 36$, deve ser mantido no modelo de previsão *fuzzy* para o índice TAIFEX, pois constitui um ruído branco para o modelo estimado.

Tabela 6.98 - Probabilidade de significância dos resultados da Tabela 6.97.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,2603	0,6469	0,7798	0,2598	0,3550	0,8065
AR(1)	0,0028*	0,3622	0,9210	0,5907	0,3495	0,5871
AR(2)	0,3796	0,3053	0,7166	0,7295	0,4126	0,9668
AR(3)	0,6190	0,0052*	0,1654	0,5441	0,4402	0,9975
AR(4)	0,0059*	0,0669*	0,3163	0,8653	0,7017	0,8059
AR(5)	0,0062*	0,1499	0,3471	0,9620	0,0948*	0,9525

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Tabela 6.99 - Testes das estimativas dos parâmetros ao modelo FSARIMA(2,1,1)(1,0,1)₃₆ ajustado para o índice TAIFEX.

Modelo	Coeficiente	Erro padrão	Estatística t	P-valor
AR(1)	-0,910	0,104	-8,750	0,000*
AR(2)	-0,711	0,106	-6,680	0,000*
SAR(36)	1,000	0,000	2607,690	0,000*
MA(1)	-0,782	0,032	-24,560	0,000*
SMA(36)	-0,708	0,174	-4,060	0,000*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Tabela 6.100 - Teste *Box-Pierce* modificado (*Ljung-Box*) pela estatística Qui-quadrado ao modelo FSARIMA(2,1,1)(1,0,1)₃₆ ajustado para o índice TAIFEX.

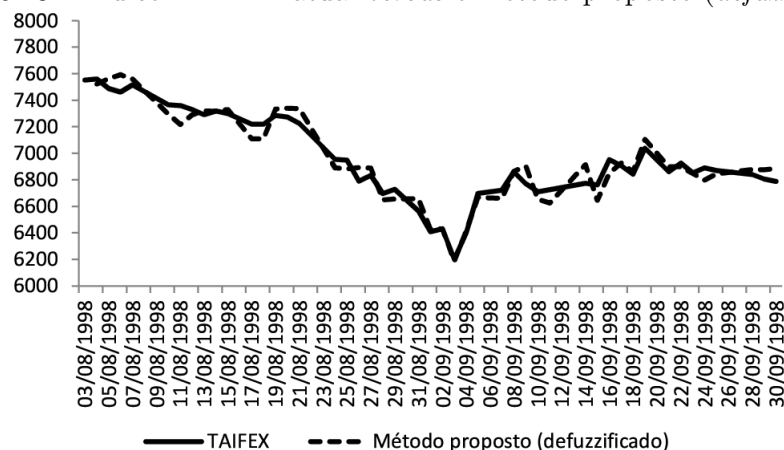
Lag	Qui-quadrado	Grau de liberdade	P-valor
12	5,300	06	0,506
24	7,200	18	0,989
36	44,100	30	0,047
48	-	-	-

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

Mediante a análise dos resultados da Tabela (6.101), a capacidade preditiva do modelo

estimado FSARIMA(2,1,1)(1,0,1)₃₆, pode ser plenamente avaliada. Comparando os valores das previsões obtidas pelo método proposto com os valores defuzzificados, se pode constatar que o modelo de previsão *fuzzy* ajustado ao índice TAIFEX, demonstra uma boa acurácia diante do período estudado. Porém, o modelo ajustado supera apenas dois métodos de previsão, que são propostos em Wang e Chen (2009) e apresentados na Tabela (6.101). Considerando que a métrica de erro utilizada na comparação dos resultados, retornou um resultado do MASE = 0,77, portanto, menor que 1, segundo Hyndman e Koehler (2006), isso é um forte indicativo que os dados da série do índice TAIFEX, se ajustaram muito bem ao modelo de previsão FSARIMA(2,1,1)(1,0,1)₃₆.

Figura 6.28 - Índice TAIFEX atual *versus* o método proposto (*defuzzificado*).



Ratificando o bom ajuste alcançado com o modelo FSARIMA(2,1,1)(1,0,1)₃₆. Na Figura (6.28), a qual representa graficamente os valores reais da série temporal do índice TAIFEX *versus* as previsões, que em vários momentos da trajetória estudada, o padrão de comportamento das duas séries se equivale. Este fato corrobora com as impressões relatadas anteriormente, quanto ao modelo ajustado ter conseguido capturar de forma eficiente, toda variabilidade inerente a base de dados do índice econômico TAIFEX ao longo de grande parte do período estudado. Isto conseqüentemente proporciona um aumento da acurácia do modelo de previsão *fuzzy* em relação a variável estudada neste estudo de caso.

Quanto às seis projeções realizadas pelo modelo proposto (*defuzzificado*). A Tabela (6.102) apresenta seis valores dentro dos intervalos de confiança construídos com um elevado grau de confiabilidade. Devido todas as projeções estarem contidas nos intervalos inferiores e/ou superiores, segundo a metodologia de previsão intervalar, o bom ajuste do modelo é plenamente aceito, o que ratifica novamente a sua excelente capacidade preditiva.

Tabela 6.101 - Comparação dos resultados das previsões pelo método proposto (dados *defuzzificados*) em relação aos cinco métodos propostos em Wang e Chen (2009) para o índice TAIFEX.

Data de Verificação	Índice TAIFEX original	Método de Chen (1996)	Método de Huarng (2001a)	Método de Huarng (2001b)	Método de Lee et al. (2006)	Método de Wang e Chen (2009)	Método proposto: (Treinamento)	Método proposto: (Validação)
03/08/1998	7.552,00	-	-	-	-	-	-	-
04/08/1998	7.560,00	7.450,00	7.450,00	7.450,00	-	-	7.522,61	-
05/08/1998	7.487,00	7.450,00	7.450,00	7.450,00	-	-	7.563,44	-
06/08/1998	7.462,00	7.500,00	7.450,00	7.500,00	7.450,00	7.474,50	7.592,69	-
07/08/1998	7.515,00	7.500,00	7.500,00	7.500,00	7.550,00	7.515,80	7.559,58	-
10/08/1998	7.365,00	7.450,00	7.450,00	7.450,00	7.350,00	7.413,50	7.295,63	-
11/08/1998	7.360,00	7.300,00	7.350,00	7.300,00	7.350,00	7.362,50	7.217,83	-
12/08/1998	7.330,00	7.300,00	7.300,00	7.300,00	7.350,00	7.345,00	7.288,92	-
13/08/1998	7.291,00	7.300,00	7.350,00	7.300,00	7.250,00	7.295,50	7.318,57	-
14/08/1998	7.320,00	7.188,33	7.100,00	7.188,33	7.350,00	7.315,00	7.317,96	-
15/08/1998	7.300,00	7.300,00	7.350,00	7.300,00	7.350,00	7.315,00	7.329,67	-
17/08/1998	7.219,00	7.300,00	7.300,00	7.300,00	7.250,00	7.222,00	7.108,51	-
18/08/1998	7.220,00	7.183,33	7.100,00	7.100,00	7.250,00	7.222,00	7.109,07	-
19/08/1998	7.285,00	7.183,33	7.300,00	7.300,00	7.250,00	7.282,50	7.333,62	-
20/08/1998	7.274,00	7.183,33	7.100,00	7.188,33	7.250,00	7.282,50	7.339,59	-
21/08/1998	7.225,00	7.183,33	7.100,00	7.100,00	7.250,00	7.249,50	7.336,80	-
24/08/1998	6.955,00	7.183,33	7.100,00	7.100,00	6.950,00	6.982,20	6.890,36	-
25/08/1998	6.949,00	6.850,00	6.850,00	6.850,00	6.950,00	6.940,50	6.881,17	-
26/08/1998	6.790,00	6.850,00	6.850,00	6.850,00	6.750,00	6.798,00	6.892,59	-
27/08/1998	6.835,00	6.775,00	6.650,00	6.775,00	6.850,00	6.833,50	6.887,29	-
28/08/1998	6.695,00	6.850,00	6.750,00	6.750,00	6.650,00	6.696,30	6.648,95	-
29/08/1998	6.728,00	6.750,00	6.750,00	6.750,00	6.750,00	6.745,00	6.655,78	-
31/08/1998	6.566,00	6.775,00	6.650,00	6.650,00	6.550,00	6.566,00	6.658,19	-
01/09/1998	6.409,00	6.450,00	6.450,00	6.450,00	6.450,00	6.416,60	6.429,82	-
02/09/1998	6.430,00	6.450,00	6.550,00	6.550,00	6.450,00	6.483,20	6.430,84	-
03/09/1998	6.200,00	6.450,00	6.350,00	6.350,00	6.250,00	6.214,80	6.185,93	-
04/09/1998	6.403,20	6.450,00	6.450,00	6.450,00	6.450,00	6.416,60	6.419,40	-
05/09/1998	6.697,50	6.450,00	6.550,00	6.550,00	6.650,00	6.703,60	6.668,01	-
07/09/1998	6.722,30	6.750,00	6.750,00	6.750,00	6.750,00	6.718,90	6.660,41	-
08/09/1998	6.859,40	6.775,00	6.850,00	6.850,00	6.850,00	6.833,50	6.865,07	-
09/09/1998	6.769,60	6.850,00	6.750,00	6.750,00	6.750,00	6.768,30	6.901,65	-
10/09/1998	6.709,75	6.850,00	6.650,00	6.650,00	6.750,00	6.718,90	6.653,33	-
11/09/1998	6.726,50	6.850,00	6.850,00	6.775,00	6.750,00	6.718,90	6.625,22	-
14/09/1998	6.774,55	6.850,00	6.850,00	6.775,00	6.817,00	6.780,80	6.913,37	-
15/09/1998	6.762,00	6.850,00	6.650,00	6.775,00	6.817,00	6.768,30	6.645,85	-
16/09/1998	6.952,75	6.850,00	6.850,00	6.850,00	6.817,00	6.940,50	6.850,89	-
17/09/1998	6.906,00	6.850,00	6.950,00	6.850,00	6.950,00	6.916,00	6.926,85	-
18/09/1998	6.842,00	6.850,00	6.850,00	6.850,00	6.850,00	6.833,50	6.863,07	-
19/09/1998	7.039,00	6.850,00	6.950,00	6.850,00	7.050,00	7.039,00	7.104,98	-
21/09/1998	6.861,00	6.850,00	6.850,00	6.850,00	6.850,00	6.866,00	6.897,67	-
22/09/1998	6.926,00	6.850,00	6.950,00	6.850,00	6.950,00	6.940,50	6.898,16	-
23/09/1998	6.852,00	6.850,00	6.850,00	6.850,00	6.850,00	6.833,50	-	6.847,70
24/09/1998	6.890,00	6.850,00	6.950,00	6.850,00	6.850,00	6.888,50	-	6.797,30
25/09/1998	6.871,00	6.850,00	6.850,00	6.850,00	6.850,00	6.888,50	-	6.844,46
28/09/1998	6.840,00	6.850,00	6.750,00	6.750,00	6.850,00	6.833,50	-	6.878,70
29/09/1998	6.806,00	6.850,00	6.750,00	6.850,00	6.850,00	6.833,50	-	6.876,69
30/09/1998	6.787,00	6.850,00	6.750,00	6.750,00	6.750,00	6.788,50	-	6.885,65
MASE		0,96	0,88	0,76	0,37	0,14	-	0,77

Tabela 6.102 - Intervalo de confiança *fuzzy* para as 6 projeções do índice TAIFEX pelo método proposto (dados *defuzzificados*).

Data da Projeção	Projeção do índice TAIFEX	Conjunto <i>fuzzy</i>	Intervalo <i>fuzzy</i> (u_j)	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
01/10/1998	6.658,24	A3	[6.427; 6.880]	[6175.47; 7131.10]	[6628.90; 7584.53]
02/10/1998	6.668,12	A3	[6.427; 6.880]	[6175.47; 7131.10]	[6628.90; 7584.53]
03/10/1998	6.882,09	A4	[6.653; 7.107]	[6175.47; 7131.10]	[6628.90; 7584.53]
04/10/1998	6.891,34	A4	[6.653; 7.107]	[6175.47; 7131.10]	[6628.90; 7584.53]
05/10/1998	6.893,60	A4	[6.653; 7.107]	[6175.47; 7131.10]	[6628.90; 7584.53]
06/10/1998	6.437,97	A2	[6.200; 6.653]	[6175.47; 7131.10]	[6628.90; 7584.53]

Passo 6: Escolher a função de autocorrelação estendida *fuzzy*, para identificar o modelo de série temporal *fuzzy*, sobre os dados históricos de interesse.

Passo 6(b): (b1) Para a função $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$, obter a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy* A_i . Em seguida aplicar a função de autocorrelação estendida *fuzzy* na série.

Analisando os resultados da Tabela (6.103), quanto ao processo de *fuzzificação* do índice TAIFEX com α -nível de 0,25. Uma transição completa entre os conjuntos *fuzzy* contruídos, e, que são apresentados na Figura (6.29), pode ser plenamente observado. Este resultado indica que o universo de discurso foi particionado de forma coerente, pois, todos os suportes correspondentes aos conjuntos A_j foram ativados. Quanto as Tabelas (6.105 e 6.106), a função de identificação $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$, retornou um modelo ARMA(3;2), significativo a 10% e satisfazendo a condição $p > 0$ e $q > 0$. Desta forma, o processo de modelagem com busca pelos parâmetros otimizados (menor erro de previsão), irá avaliar todas as combinações possíveis de parâmetros onde: $0 < p \leq 3$ e $0 < q \leq 2$, para o ajuste do modelo de previsão *fuzzy* definitivo.

Figura 6.29 - Conjuntos *fuzzy* construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_j e o eixo X o índice TAIFEX, com α -nível igual a 0,25, para ativar o suporte dos respectivos conjuntos *fuzzy*.

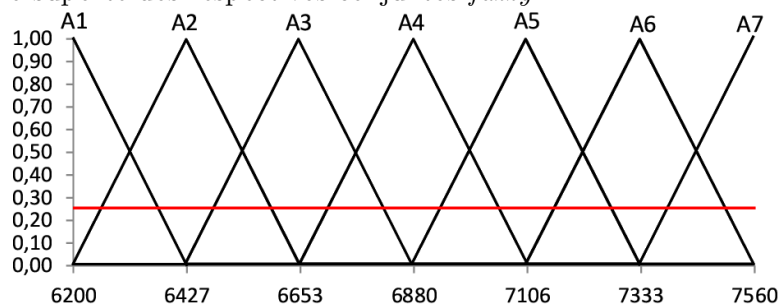


Tabela 6.103 - Série histórica do índice TAIFEX de 03/08/1998 a 30/10/1998, com processo de *fuzzificação* utilizando a função de pertinência triangular, relações *fuzzy* lógicas simples, relações *fuzzy* lógicas agrupadas e a série temporal com α -nível = 0,25 para o índice TAIFEX.

Data de verificação	TAIFEX Original	Intervalo <i>fuzzy</i> (u_i)	Conjunto <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	Série $\alpha = 0,25$
03/08/1998	7.552,00	[7.333; 7.787]	A_7	-	$A_1 \rightarrow A_2$	0,001258
04/08/1998	7.560,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_2 \rightarrow A_1, A_2, A_3$	0,000762
05/08/1998	7.487,00	[7.107; 7.560]	A_6	$A_7 \rightarrow A_6$	$A_3 \rightarrow A_3, A_4$	0,015439
06/08/1998	7.462,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$	$A_4 \rightarrow A_2, A_3, A_4, A_5$	0,006513
07/08/1998	7.515,00	[7.333; 7.787]	A_7	$A_6 \rightarrow A_7$	$A_5 \rightarrow A_3, A_4, A_5, A_6$	0,001952
10/08/1998	7.365,00	[7.107; 7.560]	A_6	$A_7 \rightarrow A_6$	$A_6 \rightarrow A_5, A_6, A_7$	0,001466
11/08/1998	7.360,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$	$A_7 \rightarrow A_6, A_7$	0,003877
12/08/1998	7.330,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,000760
13/08/1998	7.291,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,003990
14/08/1998	7.320,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,000773
15/08/1998	7.300,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,004982
17/08/1998	7.219,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,011536
18/08/1998	7.220,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,005839
19/08/1998	7.285,00	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		0,003244
20/08/1998	7.274,00	[6.880; 7.333]	A_5	$A_5 \rightarrow A_5$		0,000235
21/08/1998	7.225,00	[7.107; 7.560]	A_6	$A_5 \rightarrow A_6$		0,003346
24/08/1998	6.955,00	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		0,003218
25/08/1998	6.949,00	[6.880; 7.333]	A_5	$A_5 \rightarrow A_5$		0,001624
26/08/1998	6.790,00	[6.427; 6.880]	A_3	$A_5 \rightarrow A_3$		0,001386
27/08/1998	6.835,00	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,002817
28/08/1998	6.695,00	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,000989
29/08/1998	6.728,00	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,000595
31/08/1998	6.566,00	[6.200; 6.653]	A_2	$A_4 \rightarrow A_2$		0,003184
01/09/1998	6.409,00	[6.200; 6.653]	A_2	$A_2 \rightarrow A_2$		0,002465
02/09/1998	6.430,00	[6.200; 6.653]	A_2	$A_2 \rightarrow A_2$		0,001581
03/09/1998	6.200,00	[5.973; 6.427]	A_1	$A_2 \rightarrow A_1$		0,000790
04/09/1998	6.403,20	[6.200; 6.653]	A_2	$A_1 \rightarrow A_2$		0,004119
05/09/1998	6.697,50	[6.427; 6.880]	A_3	$A_2 \rightarrow A_3$		0,000573
07/09/1998	6.722,30	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,000027
08/09/1998	6.859,40	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000233
09/09/1998	6.769,60	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,001074
10/09/1998	6.709,75	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,005441
11/09/1998	6.726,50	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,002987
14/09/1998	6.774,55	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,006544
15/09/1998	6.762,00	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,005063
16/09/1998	6.952,75	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,001482
17/09/1998	6.906,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,004166
18/09/1998	6.842,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002448
19/09/1998	7.039,00	[6.880; 7.333]	A_5	$A_4 \rightarrow A_5$		0,002421
21/09/1998	6.861,00	[6.653; 7.107]	A_4	$A_5 \rightarrow A_4$		0,003189
22/09/1998	6.926,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000856
23/09/1998	6.852,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,003548
24/09/1998	6.890,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,005842
25/09/1998	6.871,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,006602
28/09/1998	6.840,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,001173
29/09/1998	6.806,00	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,004273
30/09/1998	6.787,00	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,001462

Passo 6(b): (b2) Para a função $W_{i,k,j}^{CA}(h; \alpha = 0,50; b)$, obter a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy* A_j . Em seguida aplicar a função de autocorrelação estendida *fuzzy* na série.

Aumentando o limite mínimo de ativação dos conjuntos A_j de α -nível = 0,25, para um α -nível = 0,50. De acordo com a Tabela (6.104), o processo de *fuzzificação* do índice

Tabela 6.104 - Série histórica do índice TAIFEX de 03/08/1998 a 30/10/1998, com processo de *fuzzificação* utilizando a função de pertinência triangular, relações *fuzzy* lógicas simples, relações *fuzzy* lógicas agrupadas e a série temporal com α -nível = 0,50 para o índice TAIFEX.

Data de verificação	TAIFEX Original	Intervalo <i>fuzzy</i> (u_j)	Conjunto <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	Série $\alpha = 0, 50$
03/08/1998	7.552,00	[7.333; 7.787]	A_7	-	$A_1 \rightarrow A_2$	0,000405
04/08/1998	7.560,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_2 \rightarrow A_1, A_2, A_3$	0,002308
05/08/1998	7.487,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_3 \rightarrow A_2, A_3, A_4$	0,002009
06/08/1998	7.462,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_4 \rightarrow A_3, A_4, A_5$	0,000828
07/08/1998	7.515,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_5 \rightarrow A_4, A_5, A_6$	0,002872
10/08/1998	7.365,00	[7.107; 7.560]	A_6	$A_7 \rightarrow A_6$	$A_6 \rightarrow A_5, A_6$	0,000350
11/08/1998	7.360,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$	$A_7 \rightarrow A_6, A_7$	0,003904
12/08/1998	7.330,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,006968
13/08/1998	7.291,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,002737
14/08/1998	7.320,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,005564
15/08/1998	7.300,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,002997
17/08/1998	7.219,00	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		0,000738
18/08/1998	7.220,00	[6.880; 7.333]	A_5	$A_5 \rightarrow A_5$		0,002923
19/08/1998	7.285,00	[7.107; 7.560]	A_6	$A_5 \rightarrow A_6$		0,001329
20/08/1998	7.274,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,005814
21/08/1998	7.225,00	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		0,002789
24/08/1998	6.955,00	[6.653; 7.107]	A_4	$A_5 \rightarrow A_4$		0,000842
25/08/1998	6.949,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000915
26/08/1998	6.790,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,001285
27/08/1998	6.835,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002377
28/08/1998	6.695,00	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,001016
29/08/1998	6.728,00	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,001287
31/08/1998	6.566,00	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,002274
01/09/1998	6.409,00	[6.200; 6.653]	A_2	$A_3 \rightarrow A_2$		0,004605
02/09/1998	6.430,00	[6.200; 6.653]	A_2	$A_2 \rightarrow A_2$		0,003292
03/09/1998	6.200,00	[5.973; 6.427]	A_1	$A_2 \rightarrow A_1$		0,005519
04/09/1998	6.403,20	[6.200; 6.653]	A_2	$A_1 \rightarrow A_2$		0,001040
05/09/1998	6.697,50	[6.427; 6.880]	A_3	$A_2 \rightarrow A_3$		0,002235
07/09/1998	6.722,30	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,001620
08/09/1998	6.859,40	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,002202
09/09/1998	6.769,60	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,004263
10/09/1998	6.709,75	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,002221
11/09/1998	6.726,50	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,002523
14/09/1998	6.774,55	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,003708
15/09/1998	6.762,00	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,004174
16/09/1998	6.952,75	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,004135
17/09/1998	6.906,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000582
18/09/1998	6.842,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,004986
19/09/1998	7.039,00	[6.880; 7.333]	A_5	$A_4 \rightarrow A_5$		0,003525
21/09/1998	6.861,00	[6.653; 7.107]	A_4	$A_5 \rightarrow A_4$		0,006322
22/09/1998	6.926,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002556
23/09/1998	6.852,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,003139
24/09/1998	6.890,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,007875
25/09/1998	6.871,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,003524
28/09/1998	6.840,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000228
29/09/1998	6.806,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002907
30/09/1998	6.787,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002726

TAIFEX manteve a transição completa entre os conjuntos *fuzzy* construídos. Isto é, a homogeneidade entre a distribuição dos conjuntos apresentados nas Figuras (6.29 e 6.30), em relação à série temporal do índice TAIFEX particionada, está plenamente preservada. As Tabelas (6.107 e 6.108), indicam a identificação pela função $W_{i,k,j}^{CA}(h; \alpha = 0, 50; b)$, como sendo um modelo ARMA(5;2), para os parâmetros p e q de maior grau, dentre todos os

que são significativos a 10% ($p\text{-valor} < 10\%$). Portanto, a rotina de modelagem otimizada avaliará o desempenho dos processos, tal que, as estimativas dos coeficientes dos modelos de previsão atendam a restrição: $0 < p \leq 5$ e $0 < q \leq 2$.

Tabela 6.105 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,25; b)$ para o índice TAIFEX.

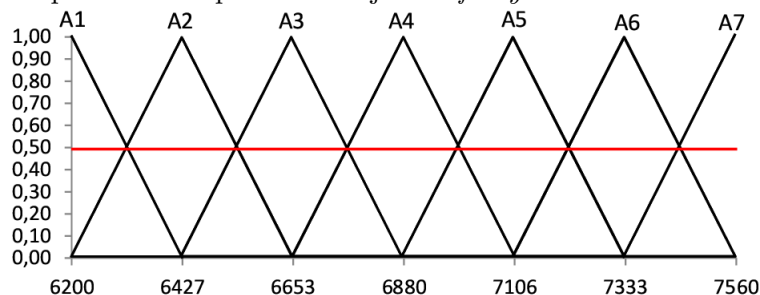
Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,344	-0,116	-0,058	0,057	-0,021	-0,071
AR(1)	-0,485	0,027	-0,169	0,040	-0,002	-0,061
AR(2)	-0,432	-0,202	-0,178	-0,017	0,029	-0,150
AR(3)	-0,381	-0,303	-0,363	-0,050	0,078	-0,073
AR(4)	-0,491	0,048	-0,088	-0,085	0,113	-0,095
AR(5)	-0,418	-0,469	-0,299	-0,210	-0,241	-0,290

Tabela 6.106 - Probabilidade de significância dos resultados da Tabela (6.105).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,020*	0,479	0,724	0,732	0,898	0,669
AR(1)	0,001*	0,885	0,357	0,809	0,994	0,741
AR(2)	0,004*	0,228	0,258	0,930	0,898	0,505
AR(3)	0,012*	0,048*	0,018*	0,826	0,731	0,699
AR(4)	0,002*	0,800	0,717	0,715	0,613	0,578
AR(5)	0,008*	0,003*	0,114	0,272	0,143	0,168

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Figura 6.30 - Conjuntos *fuzzy* construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_j e o eixo X o índice TAIFEX, com α -nível igual a 0,50, para ativar o suporte dos respectivos conjuntos *fuzzy*.



Passo 6(b): (b3) Para a função $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$, obter a série temporal *fuzzy* pela ativação dos conjuntos *fuzzy* A_j . Em seguida aplicar a função de autocorrelação estendida *fuzzy* na série.

A Tabela (6.109) apresenta o processo de *fuzzificação* do índice TAIFEX com α -nível = 0,75. Contudo, assim como ocorreu com os outros dois níveis de ativação (α -nível = 0,25 e 0,50) utilizados anteriormente, neste processo de *fuzzificação*, os suportes dos conjuntos *fuzzy* (ver, Figura (6.31)), a transição entre os conjuntos A_i também foi realizada de forma integral. O resultado do processo de identificação pela função $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$, e, re-

presentado pelas Tabelas (6.110 e 6.111), indica um modelo ARMA(5;3), estatisticamente significativo a 10%, como o modelo de maior ordem para p e q , respectivamente.

Tabela 6.107 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,50; b)$ para o índice TAIFEX.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,457	-0,051	0,060	-0,163	0,247	-0,121
AR(1)	-0,531	-0,447	-0,017	-0,039	0,187	-0,116
AR(2)	-0,401	-0,250	0,063	-0,074	0,067	-0,070
AR(3)	-0,366	-0,422	-0,244	-0,239	0,026	-0,101
AR(4)	0,044	-0,156	-0,379	-0,227	-0,130	0,003
AR(5)	0,282	-0,027	-0,456	0,063	-0,230	-0,012

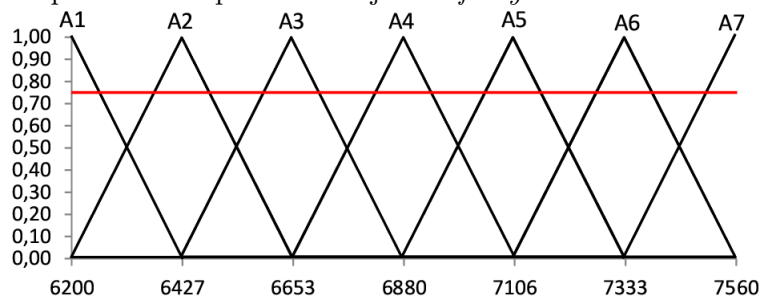
Tabela 6.108 - Probabilidade de significância dos resultados da Tabela (6.107).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,002*	0,774	0,732	0,354	0,168	0,519
AR(1)	0,000*	0,003*	0,917	0,820	0,257	0,579
AR(2)	0,008*	0,116	0,715	0,667	0,724	0,682
AR(3)	0,016*	0,006*	0,128	0,238	0,896	0,596
AR(4)	0,774	0,311	0,027*	0,272	0,565	0,991
AR(5)	0,072*	0,864	0,005*	0,750	0,315	0,962

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Logo, a rotina de otimização paramétrica do modelo de previsão ao índice TAIFEX, irá avaliar o desempenho dos modelos onde: $0 < p \leq 5$ e $0 < q \leq 3$. Devido, a convergência na ordem do parâmetro $p = 5$, quando α -nível = 0,5 e 0,75, e ainda, $q = 2$, mediante α -nível = 0,25 e 0,50. O modelo de previsão *fuzzy* para o índice TAIFEX será ajustado, especificamente considerando os valores de $p = 5$ e $q = 2$.

Figura 6.31 - Conjuntos *fuzzy* construídos pela definição de número *fuzzy* triangular, onde o eixo Y indica a pertinência ao conjunto *fuzzy* A_j e o eixo X o índice TAIFEX, com α -nível igual a 0,75, para ativar o suporte dos respectivos conjuntos *fuzzy*.



Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6 (b);

Consultando a Tabela (6.112), se tem um teste de significância estatística sobre os parâmetros do modelo FSARIMA(4;1;2)(1;0;1)₂₄, que foi ajustado por uma rotina de otimização paramétrica. É possível perceber na Tabela (6.112), uma não significância estatística do parâmetro $p = 3$, isto é, o termo autoregressivo de ordem 3 (AR(3)), apresenta

Tabela 6.109 - Série histórica do índice TAIFEX de 03/08/1998 a 30/10/1998, com processo de *fuzzificação* utilizando a função de pertinência triangular, relações *fuzzy* lógicas simples, relações *fuzzy* lógicas agrupadas e a série temporal com α -nível = 0,75 para o índice TAIFEX.

Data de verificação	TAIFEX Original	Intervalo <i>fuzzy</i> (u_j)	Conjunto <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	Série $\alpha = 0,75$
03/08/1998	7.552,00	[7.333; 7.787]	A_7	-	$A_1 \rightarrow A_2$	0,000675
04/08/1998	7.560,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_2 \rightarrow A_1, A_2, A_3$	0,002683
05/08/1998	7.487,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_3 \rightarrow A_2, A_3, A_4$	0,004289
06/08/1998	7.462,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_4 \rightarrow A_3, A_4, A_5$	0,001906
07/08/1998	7.515,00	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_5 \rightarrow A_4, A_5, A_6$	0,002776
10/08/1998	7.365,00	[7.107; 7.560]	A_6	$A_7 \rightarrow A_6$	$A_6 \rightarrow A_5, A_6$	0,002173
11/08/1998	7.360,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$	$A_7 \rightarrow A_6, A_7$	0,004199
12/08/1998	7.330,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,001169
13/08/1998	7.291,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,002968
14/08/1998	7.320,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,000704
15/08/1998	7.300,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,000050
17/08/1998	7.219,00	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		0,000985
18/08/1998	7.220,00	[6.880; 7.333]	A_5	$A_5 \rightarrow A_5$		0,005993
19/08/1998	7.285,00	[7.107; 7.560]	A_6	$A_5 \rightarrow A_6$		0,003022
20/08/1998	7.274,00	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,003316
21/08/1998	7.225,00	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		0,002015
24/08/1998	6.955,00	[6.653; 7.107]	A_4	$A_5 \rightarrow A_4$		0,001900
25/08/1998	6.949,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002775
26/08/1998	6.790,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000072
27/08/1998	6.835,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002640
28/08/1998	6.695,00	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,002933
29/08/1998	6.728,00	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,000557
31/08/1998	6.566,00	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,001700
01/09/1998	6.409,00	[6.200; 6.653]	A_2	$A_3 \rightarrow A_2$		0,000432
02/09/1998	6.430,00	[6.200; 6.653]	A_2	$A_2 \rightarrow A_2$		0,002639
03/09/1998	6.200,00	[5.973; 6.427]	A_1	$A_2 \rightarrow A_1$		0,001144
04/09/1998	6.403,20	[6.200; 6.653]	A_2	$A_1 \rightarrow A_2$		0,001600
05/09/1998	6.697,50	[6.427; 6.880]	A_3	$A_2 \rightarrow A_3$		0,009926
07/09/1998	6.722,30	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,006207
08/09/1998	6.859,40	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,001359
09/09/1998	6.769,60	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,004185
10/09/1998	6.709,75	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,006433
11/09/1998	6.726,50	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,003955
14/09/1998	6.774,55	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,000636
15/09/1998	6.762,00	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,004649
16/09/1998	6.952,75	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,000028
17/09/1998	6.906,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,001018
18/09/1998	6.842,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000048
19/09/1998	7.039,00	[6.880; 7.333]	A_5	$A_4 \rightarrow A_5$		0,004816
21/09/1998	6.861,00	[6.653; 7.107]	A_4	$A_5 \rightarrow A_4$		0,000220
22/09/1998	6.926,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002221
23/09/1998	6.852,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,003728
24/09/1998	6.890,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,002114
25/09/1998	6.871,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000607
28/09/1998	6.840,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000252
29/09/1998	6.806,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,000874
30/09/1998	6.787,00	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,001277

um p – valor = 0,342 > 10%. Desta forma, o modelo de previsão para o índice TAIFEX ao ser utilizado, não levará em consideração o componente autoregressivo de ordem 3. Os demais coeficientes do modelo de previsão *fuzzy* FSARIMA(4;1;2)(1;0;1)₂₄, são estatísti-

Tabela 6.110 - Resultados da $W_{i,k,j}^{CA}(h; \alpha = 0,75; b)$ para o índice TAIFEX.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,496	0,038	-0,081	0,213	-0,194	-0,088
AR(1)	-0,475	-0,277	0,009	0,158	-0,185	-0,205
AR(2)	-0,515	-0,220	0,080	0,069	-0,178	0,057
AR(3)	0,148	-0,130	-0,270	-0,229	-0,148	-0,109
AR(4)	0,490	-0,025	-0,368	-0,134	0,115	-0,154
AR(5)	-0,167	0,219	-0,243	-0,352	-0,058	-0,121

Tabela 6.111 - Probabilidade de significância dos resultados da Tabela (6.110).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,001*	0,834	0,652	0,239	0,298	0,646
AR(1)	0,001*	0,073*	0,955	0,386	0,393	0,272
AR(2)	0,001*	0,168	0,640	0,666	0,472	0,837
AR(3)	0,334	0,417	0,086*	0,164	0,472	0,568
AR(4)	0,002*	0,873	0,088*	0,529	0,548	0,420
AR(5)	0,286	0,202	0,339	0,044*	0,754	0,532

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

camente significantes a 10%. Como o modelo estimado possui o termo de sazonalidade integrado em sua estrutura física, um teste de *Qui-quadrado* se faz necessário. É possível constatar pela Tabela (6.113), que o termo sazonal $s = 24$ retornou uma probabilidade de significância igual a 0,342, isto é, um *p-valor* maior que 10%. Portanto, o termo de sazonalidade integrado ao modelo ajustado, deve ser mantido no mesmo, pois constitui um ruído branco, como comprovado também por testes computacionais que otimizaram, todos os parâmetros deste modelo de previsão *fuzzy*.

Tabela 6.112 - Testes das estimativas dos parâmetros ao modelo FSARIMA(4,1,2)(1,0,1)₃₆ ajustado para o índice TAIFEX.

Modelo	Coeficiente	Erro padrão	Estatística <i>t</i>	P-valor
AR(1)	0,240	0,093	2,580	0,014*
AR(2)	-0,312	0,090	-3,480	0,001*
AR(4)	-0,202	0,086	-2,350	0,024*
AR(5)	0,805	0,088	9,180	0,000*
SAR(36)	1,000	0,013	76,240	0,000*
MA(1)	1,099	0,089	12,390	0,000*
MA(2)	-1,038	0,037	-28,130	0,000*
SMA(36)	-0,707	0,187	-3,780	0,001*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Tabela 6.113 - Teste *Box-Pierce* modificado (*Ljung-Box*) pela estatística Qui-quadrado ao modelo FSARIMA(4,1,2)(1,0,1)₂₄ ajustado para o índice TAIFEX.

Lag	Qui-quadrado	Grau de liberdade	P-valor
12	14,000	03	0,003*
24	16,600	15	0,342
36	62,800	27	0,000*
48	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Passo 8: Comparar as previsões do passo 7, com dados reais da série e medir o MASE.

A avaliação do desempenho para o modelo de previsão *fuzzy* FSARIMA(4;1;2)(1;0;1)₂₄, será realizada mediante uma comparação da medida de acurácia MASE, retornada pelo modelo ajustado, em relação a outros métodos de previsão consolidados na literatura. Verificando a Tabela (6.114), se pode perceber que o algoritmo de previsão *fuzzy* proposto neste trabalho, auxiliou no ajuste de um modelo de previsão com baixíssimo erro (muito próximo de zero). Este fato pode ser ratificado, devido o valor da medida de desempenho MASE = 0,17, em comparação a outros métodos de previsão apresentados na Tabela (6.114), destacar apenas o método proposto por Wang e Chen (2009), com valor do MASE (0,64) superior ao obtido pelo método proposto com α -nível. No entanto, é importante destacar que o modelo proposto por Wang e Chen (2009), apesar de retornar um valor do MASE = 0,64, o mesmo não foi capaz de prever os instantes de tempo: 03/08/1998 a 05/08/1998. Neste contexto, o modelo de previsão FSARIMA(4;1;2)(1;0;1)₂₄, retornou um valor do MASE menor, em relação aos métodos avaliados e não gerou previsão apenas para o primeiro instante de tempo observado (03/08/1998).

Figura 6.32 - Índice TAIFEX atual *versus* o método proposto (α -nível).

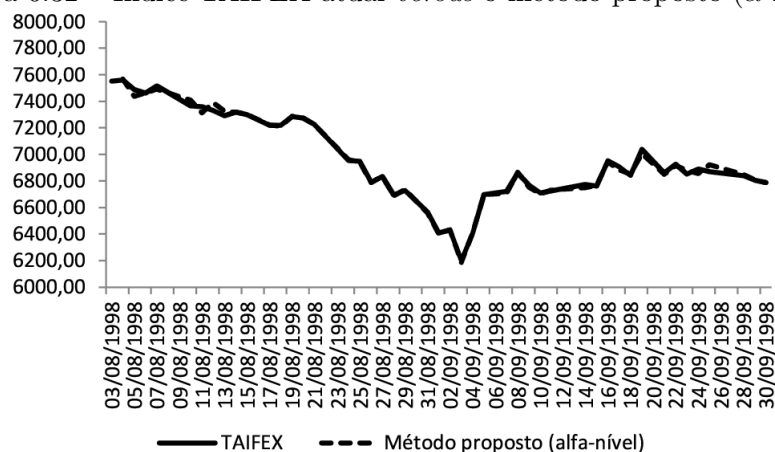


Tabela 6.114 - Comparação dos resultados das previsões pelo método proposto (α -nível) em relação aos cinco métodos propostos em Wang e Chen (2009) para o índice TAIFEX.

Data de Verificação	Índice TAIFEX original	Método de Chen (1996)	Método de Huarng (2001a)	Método de Huarng (2001b)	Método de Lee et al. (2006)	Método de Wang e Chen (2009)	Método proposto: (Treinamento)	Método proposto: (Validação)
03/08/1998	7.552,00	-	-	-	-	-	-	-
04/08/1998	7.560,00	7.450,00	7.450,00	7.450,00	-	-	7.568,32	-
05/08/1998	7.487,00	7.450,00	7.450,00	7.450,00	-	-	7.438,41	-
06/08/1998	7.462,00	7.500,00	7.450,00	7.500,00	7.450,00	7.474,50	7.462,15	-
07/08/1998	7.515,00	7.500,00	7.500,00	7.500,00	7.550,00	7.515,80	7.491,47	-
10/08/1998	7.365,00	7.450,00	7.450,00	7.450,00	7.350,00	7.413,50	7.407,70	-
11/08/1998	7.360,00	7.300,00	7.350,00	7.300,00	7.350,00	7.362,50	7.313,82	-
12/08/1998	7.330,00	7.300,00	7.300,00	7.300,00	7.350,00	7.345,00	7.389,27	-
13/08/1998	7.291,00	7.300,00	7.350,00	7.300,00	7.250,00	7.295,50	7.324,73	-
14/08/1998	7.320,00	7.188,33	7.100,00	7.188,33	7.350,00	7.315,00	7.317,52	-
15/08/1998	7.300,00	7.300,00	7.350,00	7.300,00	7.350,00	7.315,00	7.300,30	-
17/08/1998	7.219,00	7.300,00	7.300,00	7.300,00	7.250,00	7.222,00	7.221,10	-
18/08/1998	7.220,00	7.183,33	7.100,00	7.100,00	7.250,00	7.222,00	7.212,02	-
19/08/1998	7.285,00	7.183,33	7.300,00	7.300,00	7.250,00	7.282,50	7.286,44	-
20/08/1998	7.274,00	7.183,33	7.100,00	7.188,33	7.250,00	7.282,50	7.272,52	-
21/08/1998	7.225,00	7.183,33	7.100,00	7.100,00	7.250,00	7.249,50	7.222,59	-
24/08/1998	6.955,00	7.183,33	7.100,00	7.100,00	6.950,00	6.982,20	6.960,72	-
25/08/1998	6.949,00	6.850,00	6.850,00	6.850,00	6.950,00	6.940,50	6.940,20	-
26/08/1998	6.790,00	6.850,00	6.850,00	6.850,00	6.750,00	6.798,00	6.787,59	-
27/08/1998	6.835,00	6.775,00	6.650,00	6.775,00	6.850,00	6.833,50	6.835,27	-
28/08/1998	6.695,00	6.850,00	6.750,00	6.750,00	6.650,00	6.696,30	6.686,67	-
29/08/1998	6.728,00	6.750,00	6.750,00	6.750,00	6.750,00	6.745,00	6.735,42	-
31/08/1998	6.566,00	6.775,00	6.650,00	6.650,00	6.550,00	6.566,00	6.560,12	-
01/09/1998	6.409,00	6.450,00	6.450,00	6.450,00	6.450,00	6.416,60	6.401,13	-
02/09/1998	6.430,00	6.450,00	6.550,00	6.550,00	6.450,00	6.483,20	6.433,83	-
03/09/1998	6.200,00	6.450,00	6.350,00	6.350,00	6.250,00	6.214,80	6.185,61	-
04/09/1998	6.403,20	6.450,00	6.450,00	6.450,00	6.450,00	6.416,60	6.406,07	-
05/09/1998	6.697,50	6.450,00	6.550,00	6.550,00	6.650,00	6.703,60	6.696,91	-
07/09/1998	6.722,30	6.750,00	6.750,00	6.750,00	6.750,00	6.718,90	6.706,85	-
08/09/1998	6.859,40	6.775,00	6.850,00	6.850,00	6.850,00	6.833,50	6.866,97	-
09/09/1998	6.769,60	6.850,00	6.750,00	6.750,00	6.750,00	6.768,30	6.754,14	-
10/09/1998	6.709,75	6.850,00	6.650,00	6.650,00	6.750,00	6.718,90	6.701,17	-
11/09/1998	6.726,50	6.850,00	6.850,00	6.775,00	6.750,00	6.718,90	6.731,49	-
14/09/1998	6.774,55	6.850,00	6.850,00	6.775,00	6.817,00	6.780,80	6.749,33	-
15/09/1998	6.762,00	6.850,00	6.650,00	6.775,00	6.817,00	6.768,30	6.767,25	-
16/09/1998	6.952,75	6.850,00	6.850,00	6.850,00	6.817,00	6.940,50	6.944,06	-
17/09/1998	6.906,00	6.850,00	6.950,00	6.850,00	6.950,00	6.916,00	6.881,59	-
18/09/1998	6.842,00	6.850,00	6.850,00	6.850,00	6.850,00	6.833,50	6.851,28	-
19/09/1998	7.039,00	6.850,00	6.950,00	6.850,00	7.050,00	7.039,00	7.006,11	-
21/09/1998	6.861,00	6.850,00	6.850,00	6.850,00	6.850,00	6.866,00	6.850,70	-
22/09/1998	6.926,00	6.850,00	6.950,00	6.850,00	6.950,00	6.940,50	6.916,55	-
23/09/1998	6.852,00	6.850,00	6.850,00	6.850,00	6.850,00	6.833,50	-	6.877,88
24/09/1998	6.890,00	6.850,00	6.950,00	6.850,00	6.850,00	6.888,50	-	6.855,76
25/09/1998	6.871,00	6.850,00	6.850,00	6.850,00	6.850,00	6.888,50	-	6.922,38
28/09/1998	6.840,00	6.850,00	6.750,00	6.750,00	6.850,00	6.833,50	-	6.850,98
29/09/1998	6.806,00	6.850,00	6.750,00	6.850,00	6.850,00	6.833,50	-	6.809,12
30/09/1998	6.787,00	6.850,00	6.750,00	6.750,00	6.750,00	6.788,50	-	6.787,08
MASE		0,96	0,88	0,76	0,65	0,64	-	0,17

Analisando graficamente o comportamento das previsões produzidas, em relação aos valores utilizados para o ajuste do modelo FSARIMA(4;1;2)(1;0;1)₂₄. A Figura (6.32) evidencia uma fortíssima associação entre os valores atuais e as previsões, pois em grande parte da série temporal do índice TAIFEX e da série temporal das previsões, as trajetórias des-

tas duas séries se sobrepõem. Isto é, claramente há uma proximidade muito acentuada em relação aos pontos plotados na Figura (6.32), para os instantes de tempo entre 03/08/1998 a 30/09/1998. Quantificando a relação entre os valores das duas séries apresentadas na Figura (6.32), mediante uma correlação de *Pearson*. Esta estatística de teste apresenta valor igual a 0,998, com $p\text{-valor} = 0,000$, portanto, estatisticamente significativa a 10%. Estes resultados ratificam as fortes impressões de um bom ajuste da base de dados do índice TAIFEX, em relação ao modelo de previsão *fuzzy* ajustado.

Com o intuito de avaliar a capacidade preditiva do modelo FSARIMA(4;1;2)(1;0;1)₂₄. Na Tabela (6.115), seis projeções foram realizadas para o índice TAIFEX, referentes aos dias: 01/10/1998 a 06/01/1998. É possível perceber na Tabela (6.115), que todas as projeções estão contidas em pelo menos, um dos intervalos construídos com 99% de confiança.

Tabela 6.115 - Intervalo de confiança *fuzzy* para as 6 projeções do índice TAIFEX pelo método proposto (α -nível).

Data da Projeção	Projeção do índice TAIFEX	Conjunto <i>fuzzy</i>	Intervalo <i>fuzzy</i> (u_j)	I.C($\bar{\mu}_h^L$; 99%)	I.C($\bar{\mu}_h^U$; 99%)
01/10/1998	6.756,49	A3	[6.427; 6.880]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
02/10/1998	6.777,49	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
03/10/1998	6.842,03	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
04/10/1998	6.786,59	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
05/10/1998	6.719,09	A3	[6.427; 6.880]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
06/10/1998	6.490,25	A2	[6.200; 6.653]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]

Face aos resultados obtidos na Tabela (6.115), onde, devido o altíssimo grau de confiabilidade implementado na construção dos intervalos *fuzzy*, é plenamente possível aceitar que o modelo de previsão, foi muito bem ajustado.

Passo 6(c): Para utilizar a função de identificação $\bar{W}_{i,k,j}^{RA}(h)$, se deve selecionar o valor da função de autocorrelação amostral (fac) média aleatória, no suporte do conjunto *fuzzy* ativado pela pertinência máxima em relação a A_j , mediante os intervalos u_j construídos no passo 5.

Após o processo de *fuzzificação* apresentado na Tabela (6.116), se pode perceber que houve uma transição completa entre os conjuntos A_j , isto é, ocorreu a ativação de todos os conjuntos *fuzzy* que foram construídos sobre o universo de discurso U . Este fato, sugere uma partição de U realizada de forma adequada, em relação à série temporal do índice TAIFEX. A representação gráfica dos conjuntos A_j sobre o universo de discurso pode ser observada na Figura (6.33). Os resultados da função $W_{i,k,j}^{RA}(h; b)$, quando aplicada sobre a

série temporal *fuzzy* do TAIFEX, são apresentados nas Tabelas (6.117 e 6.118), onde para cada resultado há uma probabilidade de significância associada.

Tabela 6.116 Série histórica do índice TAIFEX em 03/08/1998 a 30/10/1998, processo de *fuzzificação* mediante a pertinência máxima do número *fuzzy* triangular, lógica *fuzzy* simples e agrupada, série da função de autocorrelação amostral média aleatória em cada intervalo u_j .

Data de verificação	TAIFEX Original	Pertinência máxima	Intervalo <i>fuzzy</i> (u_j)	Conjunto <i>fuzzy</i>	Relações <i>fuzzy</i> lógicas	Grupos de relações <i>fuzzy</i> lógicas	fac média aleatória
03/08/1998	7.552,00	0,965	[7.333; 7.787]	A_7	-	$A_1 \rightarrow A_2$	0,00225
04/08/1998	7.560,00	1,000	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_2 \rightarrow A_1, A_2, A_3$	0,00209
05/08/1998	7.487,00	0,678	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_3 \rightarrow A_2, A_3, A_4$	0,00356
06/08/1998	7.462,00	0,568	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_4 \rightarrow A_3, A_4, A_5$	0,00358
07/08/1998	7.515,00	0,802	[7.333; 7.787]	A_7	$A_7 \rightarrow A_7$	$A_5 \rightarrow A_4, A_5, A_6$	0,00047
10/08/1998	7.365,00	0,859	[7.107; 7.560]	A_6	$A_7 \rightarrow A_6$	$A_6 \rightarrow A_4, A_5, A_6$	0,00272
11/08/1998	7.360,00	0,881	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$	$A_7 \rightarrow A_6, A_7$	0,00146
12/08/1998	7.330,00	0,987	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,00369
13/08/1998	7.291,00	0,814	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,00018
14/08/1998	7.320,00	0,942	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,00137
15/08/1998	7.300,00	0,854	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,00191
17/08/1998	7.219,00	0,504	[6.880; 7.333]	A_5	$A_6 \rightarrow A_5$		0,00124
18/08/1998	7.220,00	0,501	[6.880; 7.333]	A_5	$A_5 \rightarrow A_5$		0,00159
19/08/1998	7.285,00	0,788	[7.107; 7.560]	A_6	$A_5 \rightarrow A_6$		0,00177
20/08/1998	7.274,00	0,739	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,00295
21/08/1998	7.225,00	0,522	[7.107; 7.560]	A_6	$A_6 \rightarrow A_6$		0,00305
24/08/1998	6.955,00	0,670	[6.653; 7.107]	A_4	$A_6 \rightarrow A_4$		0,00097
25/08/1998	6.949,00	0,696	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00561
26/08/1998	6.790,00	0,604	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00246
27/08/1998	6.835,00	0,802	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00085
28/08/1998	6.695,00	0,815	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,00355
29/08/1998	6.728,00	0,670	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,00505
31/08/1998	6.566,00	0,615	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,00246
01/09/1998	6.409,00	0,921	[6.200; 6.653]	A_2	$A_3 \rightarrow A_2$		0,00100
02/09/1998	6.430,00	0,987	[6.200; 6.653]	A_2	$A_2 \rightarrow A_2$		0,00095
03/09/1998	6.200,00	1,000	[5.973; 6.427]	A_1	$A_2 \rightarrow A_1$		0,00021
04/09/1998	6.403,20	0,895	[6.200; 6.653]	A_2	$A_1 \rightarrow A_2$		0,00095
05/09/1998	6.697,50	0,804	[6.427; 6.880]	A_3	$A_2 \rightarrow A_3$		0,00214
07/09/1998	6.722,30	0,695	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,00004
08/09/1998	6.859,40	0,909	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,00636
09/09/1998	6.769,60	0,514	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00387
10/09/1998	6.709,75	0,750	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,00181
11/09/1998	6.726,50	0,676	[6.427; 6.880]	A_3	$A_3 \rightarrow A_3$		0,00252
14/09/1998	6.774,55	0,535	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,00553
15/09/1998	6.762,00	0,520	[6.427; 6.880]	A_3	$A_4 \rightarrow A_3$		0,00238
16/09/1998	6.952,75	0,680	[6.653; 7.107]	A_4	$A_3 \rightarrow A_4$		0,00079
17/09/1998	6.906,00	0,885	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00215
18/09/1998	6.842,00	0,833	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00077
19/09/1998	7.039,00	0,700	[6.880; 7.333]	A_5	$A_4 \rightarrow A_5$		0,00128
21/09/1998	6.861,00	0,916	[6.653; 7.107]	A_4	$A_5 \rightarrow A_4$		0,00284
22/09/1998	6.926,00	0,797	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00492
23/09/1998	6.852,00	0,877	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00262
24/09/1998	6.890,00	0,956	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00340
25/09/1998	6.871,00	0,960	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00653
28/09/1998	6.840,00	0,824	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00100
29/09/1998	6.806,00	0,674	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00065
30/09/1998	6.787,00	0,590	[6.653; 7.107]	A_4	$A_4 \rightarrow A_4$		0,00307

Na Tabela (6.118), é possível perceber que o modelo identificado de forma significativa (p -valor $< 10\%$), é um modelo autoregressivo de média móvel com parâmetros $p = 5$ e $q = 5$ (0,0388). Portanto, para todo modelo que satisfaça a restrição $0 < p \leq 5$ e

$0 < q \leq 5$, a rotina de modelagem com otimização dos parâmetros (modelo com menor erro de previsão), invariavelmente irá avaliar o desempenho apresentado pelo modelo de previsão, até que seja obtido o modelo com parâmetros considerados otimizados, para prever o índice TAIFEX.

Figura 6.33 - Representação gráfica dos conjuntos *fuzzy*, onde o eixo X representa os valores do índice TAIFEX no período de 03/08 a 30/09/1998 *versus* o eixo Y pertinência ao conjunto A_j .

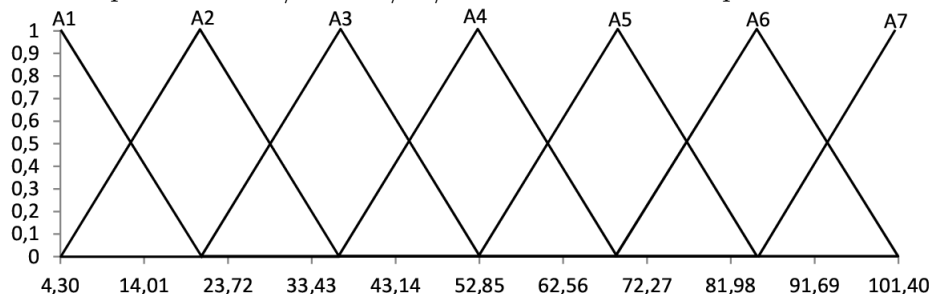


Tabela 6.117 - Resultados da função $W_{i,k,j}^{RA}(h; b)$ para o índice TAIFEX.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	-0,438	-0,188	0,213	0,062	-0,214	-0,045
AR(1)	-0,546	-0,474	0,182	0,066	-0,186	-0,080
AR(2)	-0,340	0,298	-0,289	-0,028	-0,042	-0,302
AR(3)	0,452	-0,004	-0,232	0,065	0,010	-0,277
AR(4)	0,468	-0,020	-0,200	0,061	0,013	-0,241
AR(5)	-0,236	0,336	-0,292	-0,124	-0,151	-0,359

Tabela 6.118 - Probabilidade de significância dos resultados da Tabela 6.117.

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,003*	0,279	0,232	0,735	0,245	0,811
AR(1)	0,000*	0,002*	0,331	0,711	0,327	0,669
AR(2)	0,024*	0,131	0,162	0,904	0,857	0,094*
AR(3)	0,003*	0,983	0,356	0,811	0,968	0,101
AR(4)	0,002*	0,914	0,432	0,821	0,959	0,212
AR(5)	0,131	0,093*	0,154	0,525	0,555	0,039*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Passo 7: Gerar previsões pelo modelo de série temporal *fuzzy* identificado no passo 6(c); Na Tabela (6.119), a estatística de teste *t-Student*, para os parâmetros do modelo ajustado, indica uma significância estatística a 10% (parâmetros estatisticamente diferentes de zero). Este fato é observado, em relação a todos os coeficientes otimizados para o modelo que foi ajustado, mediante testes computacionais, baseados na identificação obtida pela função $W_{t,k,j}^{RA}(h; b)$.

Tabela 6.119 - Testes estatísticos para as estimativas dos parâmetros do modelo FSARIMA(2,1,5)(1,0,1)₃₆ ajustado para o índice TAIFEX.

Modelo	Coefficiente	Erro padrão	Estatística t	P-valor
AR(1)	0,3126	0,0788	3,97	0,000*
AR(2)	-0,9505	0,0736	-12,92	0,000*
SAR(36)	1,0010	0,0034	293,22	0,000*
MA(1)	1,2505	0,0458	27,29	0,000*
MA(2)	-1,7979	0,1121	-16,04	0,000*
MA(3)	0,9343	0,0822	11,36	0,000*
MA(4)	-0,6651	0,1082	-6,15	0,000*
MA(5)	-0,2562	0,1161	-2,21	0,034*
SMA(36)	-0,7017	0,2076	-3,38	0,002*

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 10\%$.

Como se trata de um modelo com o componente de sazonalidade integrado, a Tabela (6.120) indica um teste *Qui-quadrado*, para o termo de sazonalidade identificado. O resultado do teste para o componente sazonal $s = 36$, não foi significativo, devido o $p\text{-valor} = 0,030 > 1\%$. Desta forma, se obteve um modelo de previsão híbrido FSARIMA(2;1;5)(1;0;1)₃₆, para prever o índice TAIFEX, isto é, se tem um modelo de previsão SARIMA *fuzzy* de ordem relativamente elevada.

Tabela 6.120 - Teste *Box-Pierce* modificado (*Ljung-Box*) pela estatística *Qui-quadrado* ao modelo FSARIMA(2,1,5)(1,0,1)₃₆ ajustado para o índice TAIFEX.

Lag	<i>Qui-quadrado</i>	Grau de liberdade	P-valor
12	5,500	3	0,136
24	6,400	15	0,972
36	42,400	27	0,030*
48	-	-	-

Nota: * Valor significativo (estatisticamente diferente de zero) para $\alpha = 1\%$.

A Tabela (6.121) fornece os resultados das previsões para o índice TAIFEX, mediante vários métodos de previsão consolidados na literatura de série temporal. Também são apresentados os resultados de previsão pelo modelo FSARIMA(2;1;5)(1;0;1)₃₆, que foi obtido com o auxílio do algoritmo de identificação proposto neste trabalho. Pela medida de desempenho MASE, que é utilizada na comparação dos resultados, o modelo sugerido pela identificação da função $W_{i,k,j}^{RA}(h; b)$, produziu os menores erros, em relação aos cinco métodos de previsão comparados.

Analisando a Tabela (6.122), o resultado do MASE = 0,12, para o modelo de previsão *fuzzy* proposto no trabalho, mostra uma boa acurácia do modelo FSARIMA(2;1;5)(1;0;1)₃₆,

Tabela 6.121 - Comparação dos resultados das previsões pelo método proposto ($W_{i,k,j}^{RA}(h; b)$) em relação aos cinco métodos propostos em Wang e Chen (2009) para o índice TAIFEX.

Data de Verificação	Índice TAIFEX original	Método de Chen (1996)	Método de Huarng (2001a)	Método de Huarng (2001b)	Método de Lee et al. (2006)	Método de Wang e Chen (2009)	Método proposto: (Treinamento)	Método proposto: (Validação)
03/08/1998	7.552,00	-	-	-	-	-	-	-
04/08/1998	7.560,00	7.450,00	7.450,00	7.450,00	-	-	7.597,21	-
05/08/1998	7.487,00	7.450,00	7.450,00	7.450,00	-	-	7.385,27	-
06/08/1998	7.462,00	7.500,00	7.450,00	7.500,00	7.450,00	7.474,50	7.439,42	-
07/08/1998	7.515,00	7.500,00	7.500,00	7.500,00	7.550,00	7.515,80	7.523,69	-
10/08/1998	7.365,00	7.450,00	7.450,00	7.450,00	7.350,00	7.413,50	7.363,37	-
11/08/1998	7.360,00	7.300,00	7.350,00	7.300,00	7.350,00	7.362,50	7.359,53	-
12/08/1998	7.330,00	7.300,00	7.300,00	7.300,00	7.350,00	7.345,00	7.332,66	-
13/08/1998	7.291,00	7.300,00	7.350,00	7.300,00	7.250,00	7.295,50	7.294,84	-
14/08/1998	7.320,00	7.188,33	7.100,00	7.188,33	7.350,00	7.315,00	7.318,99	-
15/08/1998	7.300,00	7.300,00	7.350,00	7.350,00	7.300,00	7.315,00	7.293,95	-
17/08/1998	7.219,00	7.300,00	7.300,00	7.300,00	7.250,00	7.222,00	7.214,22	-
18/08/1998	7.220,00	7.183,33	7.100,00	7.100,00	7.250,00	7.222,00	7.221,54	-
19/08/1998	7.285,00	7.183,33	7.300,00	7.300,00	7.250,00	7.282,50	7.290,81	-
20/08/1998	7.274,00	7.183,33	7.100,00	7.188,33	7.250,00	7.282,50	7.277,74	-
21/08/1998	7.225,00	7.183,33	7.100,00	7.100,00	7.250,00	7.249,50	7.223,66	-
24/08/1998	6.955,00	7.183,33	7.100,00	7.100,00	6.950,00	6.982,20	6.951,78	-
25/08/1998	6.949,00	6.850,00	6.850,00	6.850,00	6.950,00	6.940,50	6.948,53	-
26/08/1998	6.790,00	6.850,00	6.850,00	6.850,00	6.750,00	6.798,00	6.789,95	-
27/08/1998	6.835,00	6.775,00	6.650,00	6.650,00	6.775,00	6.850,00	6.833,50	6.831,11
28/08/1998	6.695,00	6.850,00	6.750,00	6.750,00	6.650,00	6.696,30	6.690,54	-
29/08/1998	6.728,00	6.750,00	6.750,00	6.750,00	6.750,00	6.745,00	6.732,12	-
31/08/1998	6.566,00	6.775,00	6.650,00	6.650,00	6.550,00	6.566,00	6.576,98	-
01/09/1998	6.409,00	6.450,00	6.450,00	6.450,00	6.450,00	6.416,60	6.412,72	-
02/09/1998	6.430,00	6.450,00	6.550,00	6.550,00	6.450,00	6.483,20	6.420,42	-
03/09/1998	6.200,00	6.450,00	6.350,00	6.350,00	6.250,00	6.214,80	6.189,63	-
04/09/1998	6.403,20	6.450,00	6.450,00	6.450,00	6.450,00	6.416,60	6.405,40	-
05/09/1998	6.697,50	6.450,00	6.550,00	6.550,00	6.650,00	6.703,60	6.705,19	-
07/09/1998	6.722,30	6.750,00	6.750,00	6.750,00	6.750,00	6.718,90	6.721,58	-
08/09/1998	6.859,40	6.775,00	6.850,00	6.850,00	6.850,00	6.833,50	6.855,42	-
09/09/1998	6.769,60	6.850,00	6.750,00	6.750,00	6.750,00	6.768,30	6.776,12	-
10/09/1998	6.709,75	6.850,00	6.650,00	6.650,00	6.750,00	6.718,90	6.720,91	-
11/09/1998	6.726,50	6.850,00	6.850,00	6.775,00	6.750,00	6.718,90	6.721,99	-
14/09/1998	6.774,55	6.850,00	6.850,00	6.775,00	6.817,00	6.780,80	6.755,14	-
15/09/1998	6.762,00	6.850,00	6.650,00	6.775,00	6.817,00	6.768,30	6.753,07	-
16/09/1998	6.952,75	6.850,00	6.850,00	6.850,00	6.817,00	6.940,50	6.951,55	-
17/09/1998	6.906,00	6.850,00	6.950,00	6.850,00	6.950,00	6.916,00	6.877,87	-
18/09/1998	6.842,00	6.850,00	6.850,00	6.850,00	6.850,00	6.833,50	6.865,81	-
19/09/1998	7.039,00	6.850,00	6.950,00	6.850,00	7.050,00	7.039,00	6.953,39	-
21/09/1998	6.861,00	6.850,00	6.850,00	6.850,00	6.850,00	6.866,00	6.843,69	-
22/09/1998	6.926,00	6.850,00	6.950,00	6.850,00	6.950,00	6.940,50	6.940,03	-
23/09/1998	6.852,00	6.850,00	6.850,00	6.850,00	6.850,00	6.833,50	-	6.848,68
24/09/1998	6.890,00	6.850,00	6.950,00	6.850,00	6.850,00	6.888,50	-	6.878,65
25/09/1998	6.871,00	6.850,00	6.850,00	6.850,00	6.850,00	6.888,50	-	6.873,66
28/09/1998	6.840,00	6.850,00	6.750,00	6.750,00	6.850,00	6.833,50	-	6.809,45
29/09/1998	6.806,00	6.850,00	6.750,00	6.850,00	6.850,00	6.833,50	-	6.805,68
30/09/1998	6.787,00	6.850,00	6.750,00	6.750,00	6.750,00	6.788,50	-	6.824,84

sobretudo, em comparação com os cinco métodos de previsão propostos para a série do índice TAIFEX, considerada neste estudo. Segundo Hyndman e Koehler (2006), um valor do MASE < 1 , indica que os dados da série temporal se ajustaram bem ao modelo de previsão, e, portanto, há uma grande capacidade de previsão deste modelo, em relação à base de dados utilizada em seu ajuste. Portanto, o modelo FSARIMA(2;1;5)(1;0;1)₃₆,

apresentou o menor MASE, além de conseguir prever o índice TAIFEX a partir dos instantes $i \geq 2$, diferentemente do método de Wang e Chen (2009), que só inicia as previsões a partir de $i \geq 4$, como mostra a Tabela (6.122).

Considerando apenas os últimos 16 valores do TAIFEX na Tabela (6.121), vários métodos de previsão são propostos em Egrioglu (2014) e Bas et al. (2014). Os resultados dos métodos propostos são apresentados na Tabela (6.123) e comparados ao método proposto neste estudo de caso. Como a base de dados TAIFEX reduziu de 47 a 16 pontos, um novo modelo de previsão foi identificado, tal que, o método FSARIMA(3;1;2)(1;0;1)₁₀ retornou o menor MASE (0,03). A Tabela (6.123) indica que o modelo proposto possui capacidade preditiva elevada (MASE ≈ 0), sobretudo, em comparação com métodos de previsão considerados nesta tabela. A Figura (6.34) ratifica plenamente a condição de confiabilidade do modelo de previsão, após avaliação pelo valor do MASE para $T = 16$ e 47. Portanto, a excelente acurácia do método proposto foi confirmada.

Tabela 6.122 - Comparação do MASE obtido por vários métodos para o TAIFEX no período de 3 de agosto de 1998 a 30 de setembro de 1998.

Fonte	Model	MASE
Chen (1996)	Método de Chen	0,96
Huarng (2001a)	Método com duas variáveis heurísticas	0,88
Huarng (2001b)	Método com três variáveis heurísticas	0,76
Lee et al. (2006)	Método de Lee et al	0,37
Wang and Chen (2009)	Método de Wang e Chen	0,14
Método proposto	FSARIMA(2,1,5)(1,0,1) ₃₆	0,12

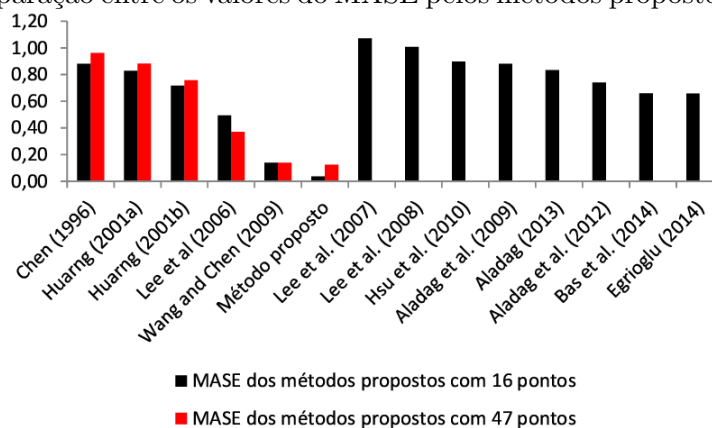
Tabela 6.123 - Comparação do MASE obtido por vários métodos para o TAIFEX no período de 10 de setembro de 1998 a 30 de setembro de 1998.

Fonte	Model	MASE
Lee et al. (2007)	Método de Lee et al.	1,07
Lee et al. (2008)	Método de Lee et al.	1,01
Hsu et al. (2010)	Método de Hsu et al.	0,90
Aladag et al. (2009)	Método de Aladag et al.	0,88
Aladag (2013)	Método de Aladag	0,83
Aladag et al. (2012)	Método de Aladag et al.	0,74
Bas et al. (2014)	Método de Bas et al.	0,66
Egrioglu (2014)	Método de Egrioglu	0,65
Método proposto	FSARIMA(3,1,2)(1,0,1) ₁₀	0,03

Tabela 6.124 - Intervalo de confiança *fuzzy* para as 6 projeções do índice TAIFEX pelo método proposto ($W_{t,k,j}^{RA}(h; b)$).

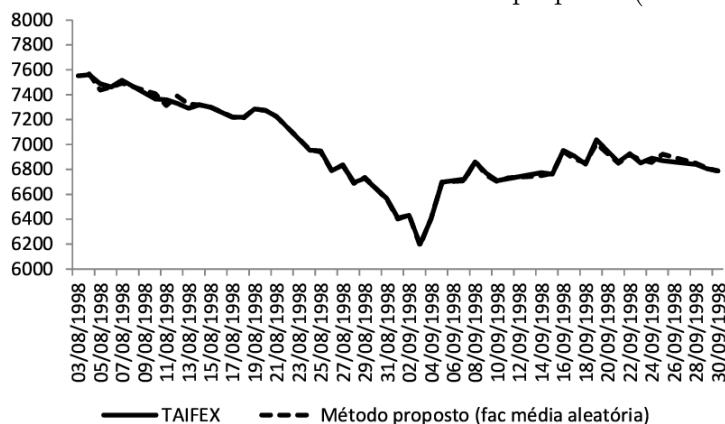
Data da Projeção	Projeção do índice TAIFEX	Conjunto <i>fuzzy</i>	Intervalo <i>fuzzy</i> (u_i)	I.C($\tilde{\mu}_h^L$; 99%)	I.C($\tilde{\mu}_h^U$; 99%)
01/10/1998	6.787,31	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
02/10/1998	6.763,52	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
03/10/1998	6.785,99	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
04/10/1998	6.760,54	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
05/10/1998	6.739,01	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]
06/10/1998	6.491,37	A4	[6.653; 7.107]	[6.175,47; 7.131,10]	[6.628,90; 7.584,53]

Figura 6.34 - Comparação entre os valores do MASE pelos métodos propostos ao índice TAIFEX.



Adicionalmente à comparação dos erros de previsão realizada anteriormente, para ratificar o resultado da identificação do modelo FSARIMA(2;1;5)(1;0;1)₃₆, foram realizadas seis projeções para o índice TAIFEX. Sobre as projeções foi aplicado a técnica de intervalo de confiança *fuzzy* (ver, Wu (2009)). A construção dos intervalos na Tabela (6.124), foi concebida com 99% de confiança, tal que, as seis projeções realizadas pelo modelo FSARIMA(2;1;5)(1;0;1)₃₆, observam-se dentro de pelo menos um dos intervalos que foram construídos. Portanto, face aos resultados das projeções obtidas, o modelo pode ser considerado necessariamente muito bem identificado.

O excelente ajuste do modelo FSARIMA(2;1;5)(1;0;1)₃₆, também é plenamente comprovado analisando a Figura (6.35), que trata da comparação gráfica dos valores atuais *versus* os valores das previsões pelo método proposto. Nesta Figura (6.35), a excelência quanto ao ajuste do modelo de previsão, fica claramente evidenciado, pois, há uma sobreposição, na maioria da extensão de trajetória da série temporal. Portanto, entre os instantes 03/08/1998 a 30/09/1998, uma quantidade muito pequena, em relação ao to-

Figura 6.35 - Índice TAIFEX atual *versus* o método proposto (fac média aleatória).

tal de pontos dos valores atuais ($T = 47$), não apresentam valores coincidentes com os resultados das previsões obtidas para o índice TAIFEX.

6.2 - Conclusão do Capítulo

Neste capítulo buscou-se aplicar todas as ferramentas propostas neste trabalho, a partir da utilização das mesmas em dados reais disponíveis na literatura, para que fosse possível desta forma medir a acurácia dos métodos de identificação que foram desenvolvidos, e ainda ser possível ratificar a eficácia e aplicabilidade destas metodologias de identificação paramétrica, em dados das mais diversas áreas do conhecimento científico, como por exemplo, para dados de matrículas em uma universidade, o preço e a carga de energia elétrica, o índice econômico TAIFEX.

Inicialmente, ao aplicar as metodologias desenvolvidas neste trabalho sobre dados do preço de energia elétrica em New South Wales na Austrália, se constatou uma eficiência razoável das ferramentas propostas quanto à identificação do modelo de série temporal *fuzzy* resultante, pois ao estabelecer o modelo SARIMA *fuzzy*, como a representação estocástica do comportamento apresentado pelos dados em questão, durante o período que os dados foram analisados o modelo identificado se mostrou capaz de gerar previsões confiáveis, em relação à característica estudada.

Além disso, o modelo identificado retornou um erro de previsão muito pequeno, sobretudo, quando comparado com os erros obtidos mediante outras metodologias de previsão clássicas e *fuzzy*, que se encontram disponíveis na literatura, na forma de *Benchmark* para

série temporal. Portanto, os resultados obtidos ratificaram de forma direta e objetiva a eficiência das ferramentas propostas neste estudo, quanto à previsão do preço de energia elétrica em New South Wales.

Um segundo estudo de caso foi realizado sobre a carga de energia elétrica, onde novamente as ferramentas de identificação propostas neste estudo, demonstraram uma grande eficiência no auxílio da diminuição do erro de previsão, do modelo de série temporal *fuzzy* estabelecido aos dados. Em comparação a outros modelos de previsão via série temporal, para a série da carga de energia elétrica, uma das ferramentas propostas neste estudo conseguiu superar o erro de previsão, em relação a todos os modelos propostos por uma pesquisa científica publicada na literatura, o que corroborou com os bons resultados já alcançados até então.

Quanto à diminuição do erro de previsão do modelo de série temporal *fuzzy*, a partir da incorporação na rotina de identificação paramétrica do modelo, pelas ferramentas propostas neste trabalho com três variações específicas, se tornou perceptível o considerável aumento da acurácia dos modelos de previsão *fuzzy*, os quais foram utilizados para prever e projetar valores da carga de energia elétrica, que foram utilizados no ajuste do modelo que melhor capturasse, toda dependência inerente aos dados modelados estocasticamente, ao longo de todo o período compreendido pela série temporal da carga de energia elétrica.

Foi possível ainda, com o auxílio das metodologias de identificação para modelos de série temporal *fuzzy*, gerar projeções para um horizonte de seis anos (matrículas) e seis instantes de tempo (preço e carga de energia elétrica, índice TAIEX), em relação ao período que foi utilizado para ajustar os modelos de previsão *fuzzy* da série de matrículas e das demais variáveis. Isto possibilitou medir novamente a acurácia das ferramentas de identificação paramétricas utilizadas neste estudo, visto que, a estas projeções se atribuiu o conceito metodológico de intervalo de confiança *fuzzy*, que representa uma ferramenta recente, todavia, extremamente eficiente quanto à estimação paramétrica intervalar, no ambiente clássico, isto é, para dados *Crisp*, como também no universo *fuzzy*, que foi adotado para este estudo.

Quanto ao terceiro e ao quarto estudo de caso realizado neste trabalho, sobre o número de matrículas na Universidade do Alabama, e para o índice TAIEX, respectivamente, os resultados obtidos com a modelagem estatística temporal foram considerados satis-

fatórios. Os erros de previsão adotados como critério de comparação dos modelos, superaram o desempenho de vários métodos de previsão *fuzzy*, e também, de métodos clássicos quando comparados. Além do que, os erros retornados pelos modelos identificados neste estudo apresentaram um comportamento de ruído branco, tal como, a metodologia de série temporal clássica, e a metodologia de série temporal *fuzzy* necessita, para produzir uma confiabilidade significativa às previsões realizadas.

No próximo capítulo serão apresentadas as conclusões finais deste trabalho, evidenciando as principais contribuições alcançadas pelos métodos propostos neste estudo, desde uma breve discussão sobre o estado da arte relacionado à metodologia que foi estudada, como também, a proposição de novas ferramentas para auxiliar a identificação dos modelos de série temporal *fuzzy*. Os resultados de simulações realizadas e a validação das técnicas de identificação implementadas mediante dois estudos de caso, também serão comentados, com intuito de consolidar a plena funcionalidade das ferramentas de identificação paramétrica propostas. Perspectivas de novas propostas metodológicas de identificação, sobretudo, para modelos de série temporal *fuzzy*, também serão consideradas durante a concepção do próximo capítulo, tentando delinear um novo horizonte na pesquisa científica, relacionado especificamente ao tema abordado neste trabalho.

Capítulo 7

Conclusões Finais da Tese e Recomendações Para Trabalhos Futuros

7.1 - Conclusões Finais Sobre os Resultados Alcançados Com Este Trabalho

A proposta de identificação desenvolvida neste trabalho, se mostrou eficiente e, portanto, passível de ser indicada para auxiliar na estimação dos parâmetros de processos autoregressivos sazonais, para séries estruturalmente de curta duração, como também em séries de média e longa duração. E ainda, este estudo propôs três ferramentas de identificação inéditas na literatura científica, quanto à determinação e estimação paramétrica pontual, dentro da metodologia de série temporal clássica combinada com a teoria de conjuntos *fuzzy*. Neste contexto, o intuito foi desenvolver mecanismos capazes de corroborar, sobretudo, de uma forma eficiente e precisa, para determinar a estrutura física que melhor represente os modelos de previsão *fuzzy* aos dados.

Mediante os resultados obtidos neste trabalho, se pode constatar que a hipótese básica desta tese não foi refutada, ou seja, a função de autocorrelação estendida *fuzzy* apresentou uma considerável capacidade, quanto a identificação dos modelos de série temporal *fuzzy*, assim como, esta função contribuiu para a escolha de um modelo que gerou previsões com erro inferior, em relação a outras funções de autocorrelação e métodos de previsão *fuzzy* disponíveis na literatura.

A suposição da função de pertinência triangular, quando combinada com a função de autocorrelação estendida, ser capaz de apresentar eficiência significativa para modelar, monitorar e gerar previsões de uma característica observável, a partir de um modelo de série temporal *fuzzy*, sendo este oriundo de um processo autoregressivo *fuzzy* de ordem “ p ”, neste caso, para $p = 2$, também foi confirmada neste trabalho. Houve um desempenho considerado satisfatório das três ferramentas propostas neste estudo, com relação à iden-

tificação dos processos autoregressivos sazonais, que foram considerados nas simulações e em aplicações às bases de dados reais.

O objetivo geral deste trabalho foi alcançado de forma satisfatória, pois a função de autocorrelação estendida, combinada com a função de pertinência triangular tornou possível, identificar de forma eficiente e precisa o modelo de série temporal *fuzzy*, sobretudo, a partir de processos autoregressivos sazonais. De forma específica, mediante as simulações que foram realizadas neste trabalho, o número de conjuntos *fuzzy*, que de forma geral também pode ser considerado como o número de termos linguísticos, para auxiliar a rotina de identificação da série temporal *fuzzy*, foram dimensionados de acordo com a trajetória simulada, isto é, o comprimento da série temporal que foi simulada neste estudo. Neste contexto, também foi incorporado ao processo de *fuzzificação* da série simulada a função de pertinência triangular, o que possibilitou traçar um perfil para uma série de comprimento T , quanto à função de identificação ótima (maior desempenho) e o número de conjuntos *fuzzy* necessários, a serem construídos para identificar com maior precisão o modelo de série temporal *fuzzy*.

Com o desenvolvimento deste estudo foi possível perceber, que a função de autocorrelação estendida, ao ser associada à metodologia contemplada pela teoria de conjuntos *fuzzy*, se mostrou uma ferramenta capaz de auxiliar a rotina de identificação paramétrica, em processos autoregressivos sazonais e modelos de série temporal *fuzzy* com trajetórias distintas. As funções de autocorrelação estendida $W_{t,k,j}^{DFZ}$; $W_{t,k,j}^{CA}(h; \alpha; b)$ e $\bar{W}_{t,k,j}^{RA}(h)$, que foram desenvolvidas neste trabalho, mostraram a partir das simulações de processos autoregressivos de ordem 2 realizados, um desempenho satisfatório quanto à identificação dos modelos simulados. Em termos gerais, para séries temporais *fuzzy* de comprimento 24, 36 e 96, com 4 combinações de parâmetros nos modelos e 5 diferentes quantidades de conjuntos *fuzzy*, o desempenho foi considerado satisfatório.

Posteriormente as simulações que foram realizadas, um estudo experimental sobre a metodologia desenvolvida foi implementada neste trabalho. Os procedimentos foram realizados sobre bases de dados disponíveis em artigos científicos, publicados em pesquisas que trataram do problema de identificação e que já foram utilizadas como séries de validação, quanto a rotinas de identificação e previsão em série temporal clássica e *fuzzy*. Os

resultados obtidos com as ferramentas propostas neste estudo se mostraram satisfatórias, pois apresentaram um erro de previsão considerado baixo, sobretudo, em comparação com os resultados obtidos por outros métodos, que foram apresentados neste trabalho. As projeções geradas para instantes posteriores ao período utilizado para estimar o modelo SARIMA *fuzzy*, também alcançaram excelentes resultados, devido a maioria dos valores situarem-se dentro do intervalo de confiança com 95% de probabilidade.

Uma das contribuições relevantes deste trabalho se deu, pelo fato da metodologia de identificação que foi proposta, ter sido concebida em sua etapa de *fuzzificação* dos dados, pela condição de ativação dos conjuntos *fuzzy* a partir da definição de número *fuzzy* triangular. Este fato tornou possível que dois conjuntos adjacentes, fossem associados a um mesmo valor da série temporal, no entanto, o conjunto *fuzzy* ativado necessariamente foi o que apresentou pertinência máxima, em relação ao valor pontual da série temporal.

Torna-se importante destacar também, que ao utilizar a função de pertinência triangular para ativar os conjuntos *fuzzy*, o grau de aderência que cada valor obteve em relação aos conjuntos construídos na *fuzzificação*, variou obrigatoriamente dentro do intervalo [0; 1]. Esta característica em especial, se diferencia da maioria das propostas de previsão *fuzzy*, que estão disponíveis na literatura, pois, estas utilizam em sua maioria a função de pertinência trivalente, isto é, uma função de ativação dos conjuntos *fuzzy*, que obrigatoriamente só assume os valores 0; 0,5 ou 1.

Neste contexto, como todos os modelos de série temporal *fuzzy* identificados neste trabalho, tiveram seus resultados comparados a outros modelos de previsão, que utilizaram em suas estruturas de *fuzzificação* a função de pertinência trivalente, acredita-se que a utilização da definição de número *fuzzy* triangular nesta proposta, tenha corroborado de forma significativa para aumentar a acurácia destes modelos de previsão, aliada a outras características importantes do algoritmo desenvolvido neste estudo, como o número de conjuntos *fuzzy*, função de autocorrelação, etc., face aos resultados relevantes que foram retornados pelos modelos de previsão *fuzzy* identificados.

7.2 - Recomendações Para Trabalhos Futuros

1. Utilizar outras funções de pertinência provenientes da definição de número *fuzzy*, diferentes da função de pertinência triangular que foi utilizada neste estudo. Desta forma, testar se uma função de pertinência específica influencia de forma significativa, na eficiência e desempenho das ferramentas de identificação para modelos de série temporal *fuzzy*, a partir da metodologia proposta neste trabalho.
2. Utilizar no processo de identificação dos modelos de série temporal, com auxílio da função de autocorrelação estendida *fuzzy* alfa-nível, valores aleatórios gerados no suporte de cada conjunto *fuzzy* ativado, a partir de um valor do alfa-nível definido a priori por um especialista. No entanto, que estes valores não necessariamente sejam resultados da função de autocorrelação amostral, calculada para os números aleatórios gerados em cada intervalo, como foi realizado neste estudo, e sim, obter, por exemplo, a média, a moda, a mediana, o mínimo ou máximo da amostra dos valores aleatórios em cada intervalo *fuzzy* construído.
3. Ampliar o estudo que foi realizado neste trabalho, para o universo multivariado de séries temporais, onde o conceito de função de transferência, conseqüentemente será incorporado à rotina de identificação. Neste contexto, a partir de uma série temporal de saída única, haverá uma associação a uma ou mais séries de entrada, como por exemplo, a série temporal do consumo diário de energia elétrica, associada a variáveis ambientais: máxima temperatura externa e umidade relativa. Para tal, a identificação desta série temporal de saída, será concebida com auxílio das ferramentas propostas neste trabalho.
4. Estudar modelos lineares e não lineares de série temporal, com saída e entrada simples, como também, saída e entrada múltipla da série temporal. Neste contexto, possibilitar a identificação do modelo de série temporal, considerando a função de autocorrelação estendida proposta neste estudo, como também, proceder o diagnóstico dos modelos identificados e realizar previsões para os modelos concebidos mediante a função de transferência.
5. Avaliar o desempenho dos modelos de série temporal destinados a trajetórias de curta

duração, tais como, o modelo de alisamento exponencial simples e duplo, o método de *Winters* aditivo e multiplicativo, aplicados sobre dados históricos oriundos de um processo *fuzzy* triangular, e, comparar os resultados das previsões obtidas, com outros métodos de previsão *fuzzy* consolidados na literatura.

6. Desenvolver um estudo sobre gráficos de controle estatístico *fuzzy*, estabelecendo a teoria conceitual para obtenção dos limites de controle *fuzzy* e seus demais termos integrantes, mediante um grau de confiabilidade pré-estabelecido, como também, pela função de pertinência adotada no processo de *fuzzificação* e pelo número de conjuntos *fuzzy* adotados, etc.

Apêndice A

Série Temporal Clássica

A.1 - Introdução

Recentemente, uma infinidade de propostas metodológicas passou a ser apresentada na literatura, com vistas ao melhor conhecimento dos fenômenos de interesse nas mais diversas áreas do conhecimento científico, sendo estes suscetíveis à variabilidade com o decorrer do tempo. Isto implica no desenvolvimento de ferramentas com capacidade de capturar, com grande eficiência e confiabilidade as características particulares destes fenômenos, sobretudo com a intenção de analisar, interpretar, monitorar e prever, tais fenômenos em um cenário de longo, médio ou curto prazo, sem desconsiderar as características inerentes a cada processo. Dentre os modelos de previsão para curto prazo, se pode destacar: os modelos lineares generalizados, como o modelo de regressão simples e múltiplo; modelos de série temporal; dentre outros modelos associados à inteligência artificial, como redes neurais artificiais, por exemplo. Neste sentido, este capítulo destina-se a apresentação e discussão dos modelos de previsão, mediante série temporal segundo a metodologia clássica “*Crisp*”. Assim, a partir da Seção 4.2, expõem-se, os conceitos básicos sobre a metodologia desenvolvida por Box-Jenkins (1970), para previsão de uma série temporal.

A.2 - Visão Geral Sobre Série Temporal

Segundo Moretin e Tolo (2004), uma série temporal é uma seqüência de observações ordenadas no tempo de uma variável de interesse X_i , onde $i = 0; \pm 1; \pm 2; \dots$, que apresenta dependência serial, isto é, dependência no instante i . Geralmente, se admite que observações sejam coletadas em espaços de tempo equivalentes, ou seja, em períodos idênticos como: diariamente, mensalmente, semestralmente, anualmente, etc. Para Cryer e Chan (2008), o conhecimento da estrutura completa de probabilidade de cada processo, é determinado pelo conjunto da distribuição de toda a coleção finita dos X'_{is} . A análise de série temporal possui basicamente três objetivos gerais em sua estrutura, como se observa em Lima (2008), e que destacam-se a seguir.

1. **Descrição:** Identificar a natureza do fenômeno representado procurando um padrão de comportamento, como o padrão de tendência, a existência de variação sazonal, *outliers* (valores discrepantes), alterações estruturais, etc;
2. **Explicação:** Usar a variação em uma série para explicar a variação em outra série;
3. **Predição:** Prever a evolução da variável da série temporal a partir de um modelo matemático que descreva o comportamento das observações.

Quando se trabalha com a metodologia de série temporal, se parte inicialmente da suposição de que a série seja estacionária (livre de tendência), ou seja, uma série é considerada estacionária, desde que, se desenvolva aleatoriamente no tempo ao redor de uma média constante, onde esta média pode ser interpretada como o nível da série, como se verifica na Figura (A.2), no entanto, o que se observa na maioria das séries são características de não estacionariedade, como por exemplo em séries Econômicas e de Energia Elétrica, dentre outras, causada pela presença de tendência e/ou sazonalidade, sendo o caso mais simples aquele em que a série flutua ao redor de uma reta, com inclinação positiva ou negativa que remete a uma tendência linear (ver, Figura A.1). A sazonalidade, quando presente mostra-se de forma bastante marcante, como por exemplo, o aumento do consumo de energia elétrica na época do verão. “O estudo dos processos estacionários pode ser feito no domínio da frequência ou no domínio do tempo. O estudo no domínio da frequência dá ênfase aos conceitos de periodograma e de densidade espectral, o domínio no tempo, às funções de autocovariância e autocorrelação” (Beltrão, 1991).

Figura A.1 - Exemplo de Uma Série Temporal Sobre o Consumo de Energia Elétrica no Brasil, entre Janeiro de 2004 e Fevereiro de 2011.

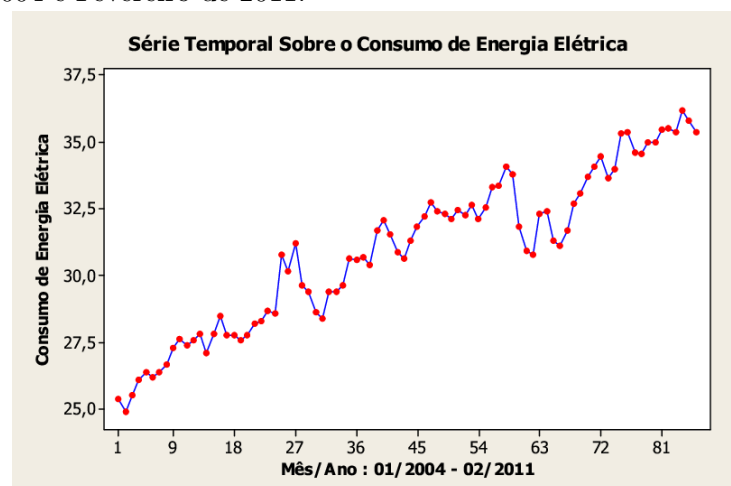
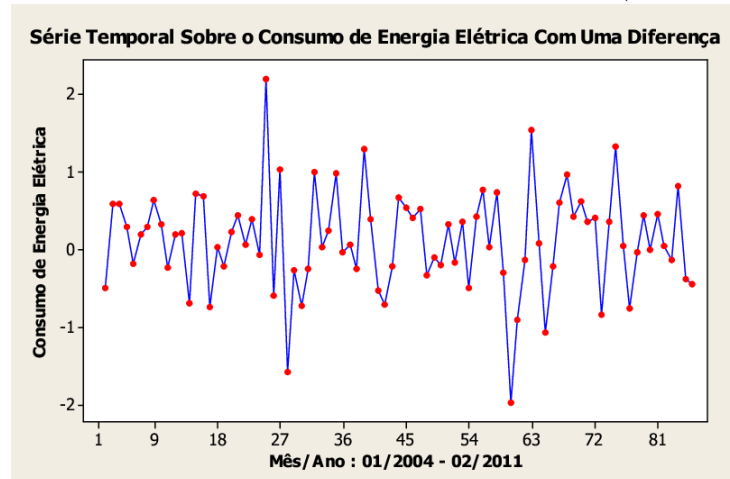


Figura A.2 - Exemplo de Uma Série Temporal Sobre o Consumo de Energia Elétrica no Brasil, entre Janeiro de 2004 e Fevereiro de 2011, Na Forma Estacionária (Livre de Tendência).



A.3 - Função de Autocorrelação Amostral e Função de Autocorrelação Parcial Amostral

Segundo Box et al. (2013), para os modelos Auto Regressivos (AR) informações adicionais sobre a ordem dos mesmos podem ser obtidas, mediante um exame na Função Autocorrelação Parcial Amostral (FACPA). A FACPA pode ser usada para identificar a ordem de um processo $AR(p)$. A Função de Autocorrelação Amostral (FACA) de um processo de médias móveis (MA) de ordem 1 se comporta, tal como, a FACPA do processo $AR(1)$, com um corte “abrupto” na primeira defasagem. A FACPA do processo $MA(1)$ decai Exponencialmente, exatamente como a FACA do processo $AR(1)$. A dualidade entre os processos MA e AR se reproduz nas “assinaturas” reveladas nas FACA e FACPA, sugerindo assim a provável ordem do modelo de série temporal.

Dado ϕ_{kj} o j -ésimo coeficiente de um processo $AR(k)$, sob a forma ϕ_{kk} , se pode comprovar a relação expressa na Equação (A.1), onde, $j = 1; 2; \dots; k$.

$$\rho_j = \phi_{k1} \times \rho_{j-1} + \phi_{k2} \times \rho_{j-2} + \dots + \phi_{kk} \times \rho_{j-k}, \quad (A.1)$$

Resolvendo as equações de *Yule-Walker*, através de um produto matricial entre a matriz de correlações e o vetor de parâmetros da FACPA, se obtém o vetor de Auto Correlações e consequentemente origina a FACPA, tal como, descrito pela Equação (A.2),

$$\phi_{kk} = \frac{\rho_k^*}{\rho_k}, \quad (A.2)$$

onde, ρ_k representa a matriz de correlações, ρ_k^* caracteriza a matriz de correlações ρ_k , com a última coluna substituída pelo vetor de Auto Correlações.

A FACPA ϕ_{kk} se equivale à correlação parcial entre Z_i e Z_{i-k} , eliminados os efeitos intermediários (influência) de $Z_{i-1}; \dots; Z_{i-k+1}$, onde $k = 0; 1; 2; \dots$, e ainda, $i = 1; 2; \dots; n$ ($n \leq T$) e T é o comprimento de trajetória da série temporal.

Definição A.1. Seja uma série temporal livre de tendência, portanto caracterizando um processo estacionário, caso este seja centrado de segunda ordem e dado por $X = (X_i)$. Se pode garantir segundo Lago Neto (2011), que a função de autocovariância e a função de autocorrelação amostral são expressas respectivamente, pelas Equações (A.3) e (A.4).

$$\gamma_X(k) = E(X_i; X_{i-k}) \tag{A.3}$$

e

$$\rho_X(k) = \frac{\gamma_X(k)}{\gamma_X(0)}, \tag{A.4}$$

tal que, $k = 0; 1; 2; \dots$, de maneira que, caso se utilize o algoritmo desenvolvido por *Durbin-Levinson*, tal como, se pode observar em Wei (2005), a função de autocorrelação parcial na defasagem $k \geq 1$ é definida por $\phi_{k;k}$, de onde, os valores de $\phi_{k;j}$ são obtidos de forma recursiva, como apresentado pelas Equações (A.5) a (A.7).

$$\phi_{1;1} = \rho_X(1), \tag{A.5}$$

$$\phi_{k;k} = \frac{\rho_X(k) - \sum_{j=1}^{k-1} \phi_{k-1;j} \rho_X(k-j)}{1 - \sum_{j=1}^{k-1} \phi_{k-1;j} \rho_X(j)} \tag{A.6}$$

e

$$\phi_{k;j} = \phi_{k-1;j} - \phi_{k;k} \phi_{k-1;k-j}, \tag{A.7}$$

tal que, $j = 1; 2; \dots; k-1$.

As rotinas dadas pelas Equações (A.6) e (A.7) são refeitas para defasagens em $k = 2; 3; 4; \dots$ e desta forma, se pode obter $\phi_{k;k}$, mediante a Equação (A.6). Admitindo como verdadeira a hipótese inicial (nula), que a série temporal em questão representa um processo que caracterize uma sequência de ruído branco, neste sentido, a variância de ϕ_{kk} pode ser aproximada por $Var(\phi_{kk}) \approx \frac{1}{T}$, onde T é o comprimento de trajetória da série

temporal, como afirma Lago Neto (2011). Em consequência, se origina um intervalo de confiança (limites críticos) dado pela expressão $\pm \frac{2}{\sqrt{T}}$ com aproximadamente 95% de confiança, para avaliar a hipótese de ruído branco no processo em análise, segundo Wei (2005).

A.3.1 - Formas de Identificação dos Processos Autoregressivos de Médias Móveis Mediante a Função de Autocorrelação Amostral e Parcial

- (a) FACA apresenta um corte abrupto (Íngreme) depois de poucas defasagens, o que caracteriza um processo de médias móveis. Isso se confirma se consequentemente a FACPA decai exponencialmente, como se pode verificar de forma clara na Figura (A.5).
- (b) FACA decai Exponencialmente, o que representa um indício do processo autoregressivo. Nesse caso, a FACPA ajuda a determinar a ordem do processo (ver, Figura A.6).

A.3.2 - Análise Estatística da Função de Autocorrelação Amostral e Parcial.

Há de ressaltar que, só se dispõe de estimativas amostrais da FACA e FACPA, que são variáveis aleatórias. Deve-se então efetuar um teste de significância para grandes amostras, assim pode-se mostrar que, sob a hipótese nula de um ruído branco, se chega as Equações (A.8 e A.9).

$$\text{FACA} : r_k \approx N \left(0; \frac{1}{\sqrt{T}} \right) \tag{A.8}$$

e

$$\text{FACPA} : \hat{\phi}_{kk} \approx N \left(0; \frac{1}{\sqrt{T}} \right). \tag{A.9}$$

O intervalo com aproximadamente 95% de confiança para uma FACA ou FACPA, é dado por: $\pm \frac{2}{\sqrt{T}}$, se r_k ou $\hat{\phi}_{kk}$ estiver fora desse intervalo, isso representa uma indicação de que o processo é um MA ou AR, respectivamente, tal que, N indica uma função densidade de probabilidade normal, com média zero e variância $\frac{1}{\sqrt{T}}$, e ainda, T denota o tamanho da série temporal observada. Portanto, as principais ferramentas utilizadas na identificação e especificação de um modelo linear são a função de autocorrelação amostral e função de autocorrelação parcial amostral.

Se uma série temporal apresenta um par “FACA e FACPA”, com comportamento similar à de um processo estocástico teórico, então esse processo se torna candidato natural para modelar a série. As mesmas duas funções FACA e FACPA, também são utilizadas sobre a série de resíduos do modelo estimado, para verificar se restaram informações. Caso hajam sobras de informações sobre os resíduos, o modelo não está bem especificado, comprometendo o desempenho do modelo.

Figura A.3 - Exemplo da FACA Para o Consumo de Energia Elétrica Com 2 lags Significativos.

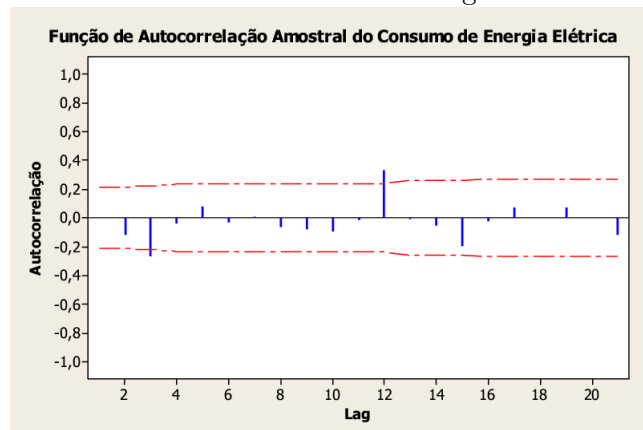


Figura A.4 - Exemplo da FACPA Para o Consumo de Energia Elétrica Com 2 lags Significativos.

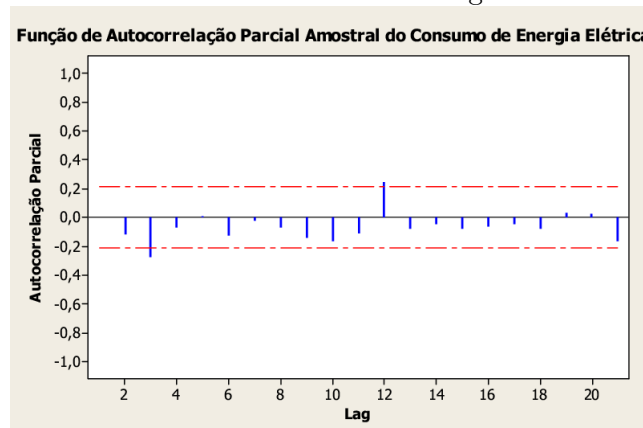


Figura A.5 - Modelo de Médias Móveis Identificado Pela Função de Autocorrelação Amostral.

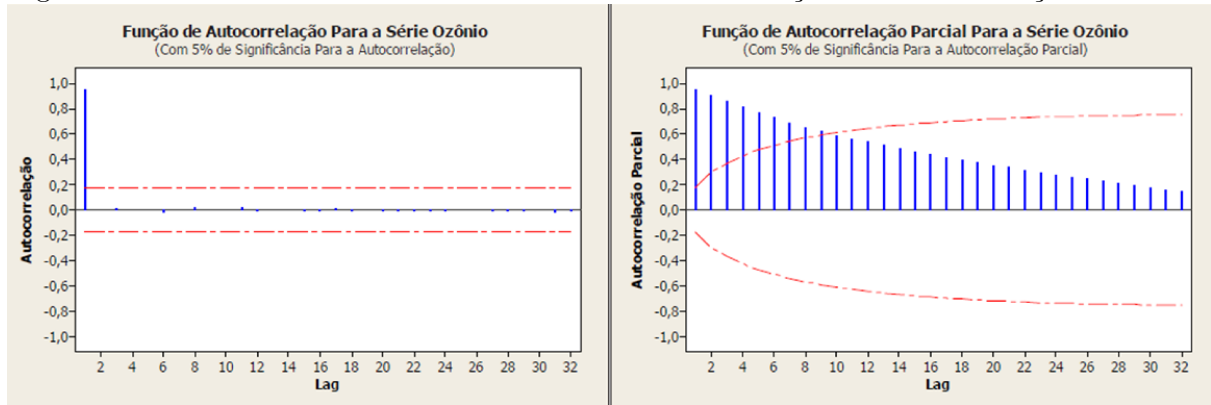
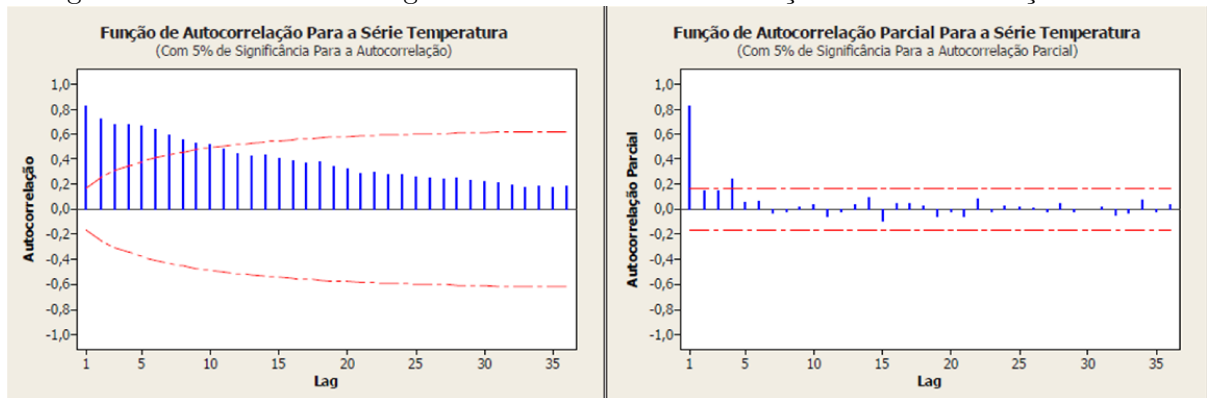


Figura A.6 - Modelo Autoregressivo Identificado Pela Função de Autocorrelação Parcial.



De acordo com as características dos dados a serem modelados, tal como, as possíveis previsões através da metodologia de série temporal, a investigação de quatro hipóteses faz-se necessário.

- (a) Hipótese de permanência (estacionariedade);
- (b) Hipótese de trajetória padrão (tendência);
- (c) Hipótese sazonal com permanência (sazonalidade);
- (d) Hipótese sazonal com trajetória (tendência e sazonalidade).

Como uma série temporal é um processo controlado a partir de leis probabilística, então, uma série temporal representa um processo estocástico, logo, assumindo uma variável T observada ao longo de um intervalo de tempo, e ainda, seja T um conjunto arbitrário, se pode concluir que um processo estocástico representa uma família, a qual pode ser descrita pela Equação (A.10).

$$Z = Z(t), t \in T, \tag{A.10}$$

tal que, para todo $t \in T$, $Z(t)$ é uma variável aleatória. Assim, um processo estocástico é uma família de variáveis aleatórias definidas em um mesmo espaço de probabilidades. O conjunto T é, em geral, uma variável aleatória discreta $Z = 0; \pm 1; \pm 2; \pm 3; \dots$, ou pode ser, uma variável aleatória contínua $Z \in \mathbb{R}$ (conjunto dos números reais).

Neste sentido, uma série temporal é composta basicamente por quatro características:

1. Tendência: Indica o comportamento da série a “longo prazo”, isto é, se ela “cresce”, “decrece” ou “se mantém” estabilizada, e qual a velocidade que ocorreram estas mudanças;
2. Sazonalidade: Consiste em flutuações periódicas específicas da variável, ou seja, indica um comportamento distinto na trajetória da série temporal durante um período específico;
3. Ciclo: Indica “formas” (padrões) na série que se repetem em períodos superiores a um ano. A identificação de ciclos se torna possível apenas quando há disponibilidade de séries longas, sendo frequente ignorar esta componente para séries temporais de curto prazo;
4. Erro: Reflete flutuações a curto prazo, com aspecto impreciso e imprevisível.

A partir das quatro características descritas acima, se tem o modelo de série temporal aditivo dado pela Equação (A.11).

$$Z(t) = T_i + S_i + C_i + a_i, \tag{A.11}$$

onde, T_i representa o componente de tendência do modelo, S_i o componente de sazonalidade, C_i o componente de ciclicidade e a_i o ruído branco do modelo o qual se supõe que apresenta média zero e variância constante $\sigma_{a_i}^2$, onde $i = 1; 2; \dots; n$ ($n \leq T$) e T é o comprimento de trajetória da série temporal. Caso as amplitudes sazonais apresentem variação em função da tendência, o modelo aditivo expresso pela Equação (A.11), perde eficiência e o modelo multiplicativo dado pela Equação (A.12) torna-se o mais adequado.

$$Z(t) = T_i \times S_i \times C_i \times a_i. \tag{A.12}$$

O modelo de *Box-Jenkins*, o método de decomposição e o modelo de suavização exponencial representam uma classe de métodos e modelos de previsão que se fundamentam na idéia que as observações passadas da série contém informações sobre o seu padrão de comportamento futuro.

A.4 - Modelo Sazonal Autoregressivo Integrado de Médias Móveis

Quando uma série temporal apresenta o componente sazonal em sua estrutura, tal que, os procedimentos para estimar os parâmetros do modelo, e ainda, a eliminação da componente de sazonalidade determinística desta série pode não ter sido alcançada, pelo fato de restar ainda uma correlação significativa, segundo Morettin e Toloí (2004), isso pode ocorrer em virtude de duas situações específicas:

Situação 1: Uma baixa ordem dos *lags*, o que representa um indicativo de que ainda há correlação entre os resíduos, algo que pode ser contornado mediante um modelo autoregressivo integrado de médias móveis (ARIMA), dentre outros procedimentos metodológicos que também podem resolver, este tipo de problema em um processo de modelagem estatística.

Situação 2: Existem *lags* sazonais, ou seja, múltiplos de períodos s . Neste sentido, não se pode desconsiderar a necessidade de considerar uma sazonalidade estocástica, o que invariavelmente resulta em ajustar a série temporal original dos dados, segundo um modelo ARIMA com o componente de sazonalidade incorporado ao mesmo, suscitando por conseguinte, um modelo denominado de sazonal autoregressivo integrado de médias móveis (SARIMA).

A.4.1 - Sazonalidade Determinística

Dada uma série temporal com padrão estrutural sazonal determinística e indicada pela variável Z_i não estacionária, com “ s ” períodos observados anualmente, ou seja, $s = 12$ para uma série coletada mensalmente. Então, Z_i , onde, $i = 1; 2; \dots; n$ ($n \leq T$) e T é o comprimento de trajetória da série temporal, para Morettin e Toloí (2004) pode ser representada pela Equação (A.13),

$$Z_i = \mu_i + N_i, \quad (\text{A.13})$$

tal que, μ_i indica uma função determinística periódica, a qual satisfaz a relação $\mu_i - \mu_{i-s} = 0$, ou de forma equivalente, se pode garantir que a Equação (A.14) também é verdadeira,

$$(1 - B^s) \times \mu_i = 0, \quad (\text{A.14})$$

e ainda, N_i representa um processo estacionário, onde se torna possível utilizar um modelo autoregressivo de médias móveis (ARMA) com parâmetros p e q , para indicar este termo. Portanto, se pode garantir que o componente estacionário N_i satisfaz plenamente a Equação (A.15),

$$\phi(B)N_i = \theta(B)a_i, \quad (\text{A.15})$$

tal que, o componente de erro ou “ruído branco” incorporado a Equação (A.15) é expresso pelo termo a_i , e ainda, uma representação algébrica para o termo μ_i presente nas Equações (A.13, A.14) é dada segundo a Equação (A.16), onde μ , α_j , β_j , para $j = 1; \dots; 6$, representam constantes desconhecidas, de acordo com Morettin e Toloi (2004).

$$\mu_i = \mu + \sum_{j=1}^6 \left[\alpha_j \times \cos\left(\frac{2\pi j i}{s}\right) + \beta_j \times \sin\left(\frac{2\pi j i}{s}\right) \right]. \quad (\text{A.16})$$

Neste sentido, ao aplicar a diferença sazonal sobre a Equação (A.13), a qual indica um modelo determinístico sazonal, se chega na expressão indicada pela Equação (A.17), com suas implicações decorrentes.

$$(1 - B^s) \times [Z_i] = [\mu_i + N_i] \times (1 - B^s) \Rightarrow (1 - B^s) \times [Z_i] = [\mu_i] \times (1 - B^s) + [N_i] \times (1 - B^s), \quad (\text{A.17})$$

no entanto, mediante a Equação (A.14) a Equação (A.17), se chega que,

$$(1 - B^s) \times [Z_t] = 0 + [N_i] \times (1 - B^s) \Rightarrow (1 - B^s) \times [Z_i] = [N_i] \times (1 - B^s). \quad (\text{A.18})$$

Procedendo a substituição da Equação (A.15) sobre a Equação (A.18), se a expressão,

$$\phi(B) \times (1 - B^s) \times [Z_i] = \theta(B) \times (1 - B^s)a_i, \quad (\text{A.19})$$

ou de forma equivalente, a Equação (A.19) pode ser reescrita como na Equação (A.20),

$$\phi(B) \times K_i = \theta(B) \times (1 - B^s)a_i, \quad (\text{A.20})$$

tal que, o termo $K_i = (1 - B^s) \times [Z_i]$.

A.4.2 - Sazonalidade Estocástica

Seja uma série temporal sazonal representada por uma variável Z_i não estacionária, onde $i = 1; 2; \dots; n$ ($n \leq T$) e T é o comprimento de trajetória da série temporal, com “ s ” períodos observados anualmente, ou seja, $s = 12$ para uma série coletada mensalmente.

Segundo Lago Neto et al. (2011), a série Z_i pode ser estabilizada (estacionária), a partir de diferenciações sazonais na mesma, dando origem a uma série X_i , removendo desta forma a sazonalidade contida na série original, tal como, implementado por Hamilton (1994), no modelo Autoregressivo Integrado de Média Móvel (ARIMA), como se observa a seguir.

$$Z_i - Z_{i-s} = (1 - B^s)Z_i = X_i. \quad (\text{A.21})$$

Em função do padrão sazonal de Z_i na Equação (A.21), se faz necessário na maioria dos casos estabelecer um modelo estocástico para representar esta série temporal, o que suscita o desenvolvimento da Equação (A.22), logo, há uma modelagem de Z_i mediante o seu padrão sazonal.

$$\Phi(B^s)(1 - B^s)^D Z_i = \Theta(B^s)X_i, \quad (\text{A.22})$$

onde a Equação (A.23) indica o polinômio autoregressivo sazonal (ordem P), na Equação (A.22),

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{Ps}, \quad (\text{A.23})$$

e a Equação (A.24) indica o polinômio de média móvel sazonal de ordem Q , na Equação (A.22),

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}. \quad (\text{A.24})$$

Segundo Lago Neto (2011), na Equação (A.22) o padrão sazonal é aleatório entre os ciclos “ s ”, se a sazonalidade presente na Equação (A.21) for filtrada, uma estrutura estocástica regular da série Z_i será obtida, mediante o modelo ARIMA ($p; d; q$), como observado na Equação (A.25),

$$\phi(B)(1 - B)^d Z_i = \theta(B)a_i, \quad (\text{A.25})$$

e a Equação (A.26) é um polinômio autoregressivo não sazonal de ordem p , na Equação (A.25),

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \quad (\text{A.26})$$

Contudo, a Equação (A.27), indica um polinômio de média móvel não sazonal de ordem q , que conseqüentemente, ao combinar-se as Equações (A.22 e A.25), se tem origem uma classe de modelos multiplicativos sazonais ARIMA($p; d; q$)($P; D; Q$) $_s$, também denominada de SARIMA e representada pela Equação (A.28),

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (\text{A.27})$$

$$\text{E ainda,} \quad \phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D Z_i = \theta(B)\Theta(B^s)a_i, \quad (\text{A.28})$$

onde a_i é um erro aleatório e $a_i \sim N(0, \sigma^2)$, isto é, o erro segue uma distribuição de probabilidade normal com média zero e variância σ^2 . Segundo Dong et al. (2011), os termos $p; d; q; P; D; Q$ são números inteiros e s representa o componente de sazonalidade do modelo. É possível garantir ainda, que as Equações (A.26, A.23, A.27 e A.24) indicam polinômios em B de grau $p; q; P; Q$. O termo B caracteriza um vetor de deslocamento retrógrado, d representa o número de diferenças realizadas na série temporal, D representa o número de diferenças sazonais na série temporal, Z_i denota o valor da série temporal no instante i , onde $i = 1; 2; \dots; n$ ($n \leq T$), e ainda, a_i representa o resíduo estimado (ruído branco) no instante i . De acordo com Lago Neto (2011), na classe de modelos acima é possível realizar tanto uma diferenciação regular quanto sazonal.

A utilização de uma série temporal sazonal é justificada quando um comportamento periódico (cíclico) é observado no tempo de monitoração, em virtude da ocorrência de um padrão anormal. Esta “anormalidade” precisa ser capturada pelo modelo estimado para que a capacidade preditiva do modelo de série temporal resultante não seja comprometida e não influencie na acurácia do modelo de previsão.

A.5 - Adequação do Modelo de Série Temporal

Os modelos de série temporal obrigatoriamente precisam atender a alguns pressupostos, tais como: a condição de estacionariedade na série temporal utilizada (sem tendência e/ou sem mudança de níveis); a condição de invertibilidade da série temporal (ser invertível), como também, os erros de previsão ser um ruído branco. Modelos oriundos de processos que não atendam a estas condições, necessariamente tendem a produzir previsões com reduzida acurácia, logo, a adequação do modelo de previsão faz parte de uma etapa do processo de identificação, extremamente relevante quanto à confiabilidade das previsões produzidas pelo método proposto aos dados.

A.5.1 - Estacionariedade

A condição de estacionariedade pode ser comprovada, pelo teste estatístico da raiz unitária sazonal do componente autoregressivo no modelo estimado, isto é, o teste de *Dickey-Fuller* e o teste HEGY (ver, Hylleberg 1995). Caso a hipótese nula de raiz unitária do

teste estatístico não seja refutada, a capacidade preditiva do modelo de série temporal sazonal estará comprometida, o que conseqüentemente resultará num modelo estimado com previsões imprecisas, e portanto, com baixa confiabilidade, devido a série temporal não ser estacionária.

Segundo Hylleberg (1995), os testes sugeridos por Canova e Hansen (1993), fundamentam-se na condição de que em uma regressão com dummies sazonais e outros componentes determinísticos, além de fatores não estocásticos como regressores de raiz unitária podem ser identificados nos resíduos de forma geral. Para Hylleberg (1995) particularmente uma raiz unitária em uma determinada frequência pode afetar a covariância dos resíduos de forma específica. Ainda, basicamente o desempenho do teste Canova e Hansen (teste CH), o qual considera os resíduos de uma regressão que extrai o componente sazonal determinístico e outros componentes determinísticos em um processo, onde a matriz de covariância é definida sob a hipótese de raiz unitária sazonal em uma determinada frequência que pode ser afetada. Neste contexto, segundo Hylleberg (1995), o teste da raiz unitária pode ser aplicado às hipóteses nas Equações (A.29)-(A.31), sob a forma de teste unilaterais à esquerda e a Equação (A.32) como um teste bilateral. Na Equação (A.29), a hipótese nula é de uma raiz unitária de longo prazo igual a zero, contra uma hipótese alternativa para esta raiz ser menor que zero, com utilização do valor da distribuição de *Dickey-Fuller* tabelado para t sobre π_1 . Assim, num cenário de longo prazo o teste da Equação (A.29) é o indicado.

$$\begin{cases} H_0 : \pi_1 = 0, & \text{a série temporal não é estacionária;} \\ H_1 : \pi_1 < 0, & \text{a série temporal é estacionária.} \end{cases} \quad (\text{A.29})$$

No caso do teste da raiz unitária semestral, se indica utilizar a Equação (A.30), o que resulta em uma raiz unitária de frequência $\frac{1}{2}$, pelo valor de t na distribuição de *Dickey-Fuller* sobre π_2 .

$$\begin{cases} H_0 : \pi_2 = 0, & \text{a série temporal não é estacionária;} \\ H_1 : \pi_2 < 0, & \text{a série temporal é estacionária.} \end{cases} \quad (\text{A.30})$$

Quanto a Equação (A.31), esta trata de um teste da raiz unitária anual, ou seja, uma raiz unitária de frequência $\frac{1}{4}$ e o conjugado complexo equivalente a $\frac{3}{4}$ de um círculo

agregando maior complexidade ao processo. Caso $\pi_4 \equiv 0$, a raiz unitária irá existir com uma frequência de $\frac{1}{4}$ se $\pi_3 = 0$ e um teste de hipóteses será baseado em um valor de i sobre π_3 seguindo uma distribuição como em Dickey et al. (1984) em um processo $x_i = \rho x_{i-2} + \varepsilon_i$ (Hylleberg, 1995).

$$\begin{cases} H_0 : \pi_3 = 0, & \text{a série temporal não é estacionária;} \\ H_1 : \pi_3 < 0, & \text{a série temporal é estacionária.} \end{cases} \quad (\text{A.31})$$

No caso de $\pi_4 \neq 0$ um teste onde a hipótese nula $H_0 : \pi_3 = 0, \pi_4 = 0$, baseado sobre o valor da distribuição F correspondente para π_3 e π_4 é sugerido, tal como se observa na Equação (A.32).

$$\begin{cases} H_0 : \pi_3 \cap \pi_4 = 0, & \text{a série temporal não é estacionária;} \\ H_1 : \pi_3 \cup \pi_4 \neq 0, & \text{a série temporal é estacionária.} \end{cases} \quad (\text{A.32})$$

A.5.2 - Invertibilidade

Seja um processo de média móvel Y_i ($i = 1; 2; \dots; n$, $n \leq T$ e T é o comprimento de trajetória da série temporal) de ordem $q = 1$, isto é, MA(1), segundo Cryer and Chan (2008), a Equação (A.33), pode plenamente representar a estrutura física deste processo.

$$Y_i = e_i - \theta e_{i-1}. \quad (\text{A.33})$$

Toda via, a função de autocorrelação obtida para o parâmetro θ e parâmetro $\frac{1}{\theta}$ na Equação (A.33), remete a exatamente um único resultado, tal como demonstra Cryer and Chan (2008), que também afirma ocorrer o mesmo para MA(2). Desta forma, antes da determinação dos parâmetros do modelo de série temporal pela função de autocorrelação, o processo de média móvel deve ser muito bem avaliado, para que a identificação do modelo não seja comprometida.

A condição invertibilidade será comprovada, reescrevendo a Equação (A.33) em função de e_i , resultando na Equação (A.34), a qual indicará os erros do modelo no instante i .

$$e_i = Y_i + \theta e_{i-1}, \quad (\text{A.34})$$

igual a:

porém, substituindo na Equação (A.34) i por $i - 1$, se chega ao erro e_i necessariamente

$$e_i = Y_i + \theta (Y_{i-1} + \theta e_{i-2}) = Y_i + \theta Y_{i-1} + \theta^2 e_{i-2}. \quad (\text{A.35})$$

Assumindo que $|\theta| < 1$, se pode continuar a obter expressões de forma infinita, isto é,

$$e_i = Y_i + \theta Y_{i-1} + \theta^2 Y_{i-2} + \dots,$$

ou seja, a Equação (A.36) será verdadeira.

$$Y_i = (-\theta Y_{i-1} - \theta^2 Y_{i-2} - \theta^3 Y_{i-3} - \dots) + e_i. \quad (\text{A.36})$$

Portanto, segundo Cryer and Chan (2008), caso $|\theta| < 1$, um modelo MA(1) pode ser invertido num modelo autoregressivo de ordem infinita, em seus próprios valores defasados. Logo, diz-se que MA(1) é um modelo invertível, se somente se, $|\theta| < 1$. De maneira equivalente, caso a condição de invertibilidade seja atendida ($|\theta| < 1$), se pode garantir que existe um único processo MA(1), para uma específica função de autocorrelação. Generalizando, para o modelo MA(q) se define um polinômio característico de MA, onde todas as raízes da equação satisfaçam $|\mathbf{B}| > 1$; no modelo ARMA($p; q$), as condições de invertibilidade e estacionariedade devem ser atendidas.

A.5.3 - Ruído Branco

Durante o processo de modelagem em série temporal, segundo Cryer e Chan (2008), um dos pontos fundamentais para que a excelência deste procedimento seja alcançada, diz respeito a este processo se comportar como um ruído branco. Neste sentido, Cryer e Chan (2008) garantem que caso a série temporal analisada seja livre de tendência (estacionária), a condição de ruído branco é plenamente atendida no processo de modelagem. O ruído branco é admitido como uma sequência de variáveis aleatórias independentes e identicamente distribuídas ($a_i, i = 1; 2; \dots; n, n \leq T$). Portanto, uma quantidade razoável de processos eficientes pode ser construída, mediante a obtenção de um processo de ruído branco a_i . Para comprovar a condição de estacionariedade de um processo, se faz necessário a comprovação das condições dadas na Equação (A.37 - A.39).

$$\begin{aligned} & P(a_{i_1} \leq x_1, a_{i_2} \leq x_2, \dots, a_{i_n} \leq x_n) \\ = & P(a_{i_1} \leq x_1) P(a_{i_2} \leq x_2) \dots P(a_{i_n} \leq x_n). \end{aligned} \quad (\text{Condição de Independência}) \quad (\text{A.37})$$

$$\begin{aligned} & P(a_{i_1} \leq x_1, a_{i_2} \leq x_2, \dots, a_{i_n} \leq x_n) \\ = & P(a_{i_1-k} \leq x_1, a_{i_2-k} \leq x_2, \dots, a_{i_n-k} \leq x_n). \end{aligned} \quad (\text{Condição de Distribuição Idêntica}) \quad (\text{A.38})$$

$$\begin{aligned}
& P(a_{i_1} \leq x_1, a_{i_2} \leq x_2, \dots, a_{i_n} \leq x_n) \\
&= P(a_{i_1-k} \leq x_1) P(a_{i_2-k} \leq x_2) \\
&\dots P(a_{i_n-k} \leq x_n). \quad (\text{Condição de Distribuição Idêntica Pela Independência}) \quad (\text{A.39})
\end{aligned}$$

A.6 - Função de Autocorrelação Inversa Amostral

A função de autocorrelação inversa amostral (FACIA) foi desenvolvida por *Cleveland* no início dos anos setenta (ver, *Cleveland 1972*), como uma forma normalizada da transformada de *Fourier* e como forma recíproca da função de densidade espectral. Portanto, a FACIA considera um processo de parâmetro $\{X\}$ discreto, cuja função de densidade espectral $h(w)$, existe para todo w e é tal que $h(w) > 0$, para todo w , e $\{h(w)\}^{-1}$ admite uma expansão em série de *Fourier*. A auto covariância inversa de lag r , $R_i(r)$, é então definida como o r -ésimo coeficiente de *Fourier* $\frac{1}{4\pi^2} \{h(w)\}^{-1}$, ou seja,

$$\frac{1}{4\pi^2} \times \{h(w)\}^{-1} = \frac{1}{2\pi} \times \sum_{r=-\infty}^{\infty} R_i(r) \times e^{-iwr}, \quad (\text{A.40})$$

de maneira que pela inversão da fórmula padrão,

$$R_i(r) = \frac{1}{4\pi^2} \times \int_{-\pi}^{\pi} \{h(w)\}^{-1} \times e^{iwr} dw, \quad (\text{A.41})$$

onde, $r = 0; \pm 1; \pm 2; \dots$

A FACIA ($\rho_i(r)$), se define pela Equação (A.42) a seguir,

$$\rho_i(r) = \frac{R_i(r)}{R_i(0)}, \quad (\text{A.42})$$

onde, $r = 0; \pm 1; \pm 2; \dots$

A motivação para a introdução das funções $R_i(r)$, $\rho_i(r)$, é que em certo sentido, elas possuem as propriedades “*Dual*” de $R_i(r)$ e $\rho_i(r)$.

Em particular, se X_i ($i = 1; 2; \dots; n$, $n \leq T$) obedece a um modelo $AR(k)$ puro, então $R_i(r)$ se comporta exatamente, tal como, a FACA de um modelo $MA(k)$ puro, isto é, $R_i(r)$ deve ser zero para todos os lags maiores do que k . Porém, se X_i obedece a um modelo $MA(s)$ puro então $R_i(r)$ se comporta, tal como, a FACA de um modelo $AR(s)$ puro, e conseqüentemente decai gradativamente para zero segundo *Cleveland (1972)*.

Segundo *Lago Neto (2011)*, a estimação da função de autocorrelação inversa pode ser

realizada, mediante um método autoregressivo, de maneira que um modelo autoregressivo de ordem “ r ” seja ajustado, como recomenda Chatfield (1979), ou seja, pela forma descrita na Equação (A.43),

$$X_i = \sum_{j=1}^r \phi_{r;j} X_{i-j} + a_i, \quad (\text{A.43})$$

necessitando que, $r \rightarrow \infty$, para se calcular a função de autocorrelação inversa de forma eficiente.

Neste sentido, a Equação (A.44) a seguir, pode descrever uma forma alternativa de calculo da função de autocorrelação inversa, tal como, encontrado em Lago Neto (2011), diferentemente da forma observada na Equação (A.42), definida anteriormente.

$$\rho_{iX}(h) = \frac{-\phi_{r;h} + \sum_{j=1}^{r-h} \phi_{r;j} \phi_{r;j+h}}{1 + \sum_{j=1}^r \phi_{r;j}^2}, \quad (\text{A.44})$$

onde, $1 \leq h \leq r$, e ainda, a obtenção dos valores de $\phi_{r;k}$, se torna possível mediante o algoritmo de *Durbin-Levinson* na forma recursiva, de acordo com o utilizado nas Equações (A.5) a (A.7).

Há possibilidade segundo Lago Neto (2011), de obter um modelo autoregressivo de ordem “ r ” ajustado, tal como, recomenda Chatfield (1979) (Ver, Equação (A.43)), de forma equivalente mediante programação em regressão múltipla, no entanto, o algoritmo de *Durbin-Levinson* é comprovadamente mais eficiente no ajuste do modelo autoregressivo de ordem “ r ”.

Em termos gerais, a função de autocorrelação inversa amostral apresenta o mesmo desempenho no processo de modelagem ARIMA, em comparação com a função de autocorrelação parcial amostral. O diferencial da FACIA em relação à FACPA se dá pela capacidade desta primeira função indicar subconjuntos e modelos autoregressivos sazonais de forma mais eficiente, o que lhe confere um *status* de superioridade no processo de identificação do modelo de série temporal.

A função de autocorrelação inversa amostral auxilia na identificação de um modelo de série temporal, em especial se esta série estiver sobre diferenciação fornecendo as seguintes implicações:

Implicação 1: Dados oriundos de um modelo não estacionário ou quase não estacionário, a FACIA apresenta um comportamento de um modelo de médias móveis não invertível.

Implicação 2: Caso a série histórica apresente dados que sigam um modelo de médias móveis não invertível, então a FACIA apresenta comportamento de não estacionariedade, e conseqüentemente, sua representação gráfica irá demonstrar uma trajetória decaindo lentamente.

Claramente a função de autocorrelação inversa amostral de um modelo autoregressivo de médias móveis, é definida como a função de autocorrelação amostral do modelo ARMA $(q; p)$. Portanto, significa afirmar que a FACIA implica na obtenção de um modelo com os operadores, de média móvel e autoregressivo invertidos, em relação à composição padrão do modelo ARMA $(p; q)$.

A.6.1 - Características Para o “Bom Ajuste” de Um Modelo ARIMA

As características desejáveis de um modelo ARIMA incluem as seguintes condições:

- (1) Um modelo com o menor número de coeficientes possível;
- (2) O modelo deve ser estacionário, ou seja, livre de tendência;
- (3) O modelo deve ser inversível;
- (4) O modelo deve apresentar estimativas de seus coeficientes com alta qualidade:
 - (4.1) O módulo do instante t , para cada estimativa do coeficiente, é maior ou igual a 2;
 - (4.2) Não existência de “alta” correlação entre os coeficientes e as estimativas.
- (5) O modelo deve apresentar resíduos não correlacionados.

A FACIA de um modelo ARMA (p, q) é definida como a FACA do modelo ARMA (q, p) . É obtida pela troca dos operadores de Médias Móveis e Auto Regressivo. Por exemplo, a Auto Correlação Inversa do processo autoregressivo de ordem 1, dado pela Equação (A.45)

$$AR(1) : (1 - 0,8B) Y_i = a_i, \quad (A.45)$$

onde, $Y_i = a_i + 0,8 Y_{i-1}$, se defini como a Auto Correlação da Equação (A.46)

$$Y_i = a_i - 0,8 a_{i-1}. \quad (A.46)$$

Tabela A.1 - Resumo sobre as características observadas nas funções de identificação para modelos de série temporal: FAC, FACP e FACL.

Função de identificação	Método Autoregressivo (p)	Método Média Móvel (q)	Método ARMA($p; q$)
Função de Autocorrelação Amostral	Decai para zero exponencialmente	Corte na defasagem q	Decai para zero exponencialmente
Função de Autocorrelação Parcial	Corte na defasagem p	Decai para zero exponencialmente	Decai para zero exponencialmente
Função de Autocorrelação Inversa	Corte na defasagem p	Decai para zero exponencialmente	Decai para zero exponencialmente

A.7 - Função de Autocorrelação Estendida

Quando o objetivo de um estudo trata da identificação paramétrica em modelos de série temporal, naturalmente a Função de Autocorrelação Amostral (FAC) e Função de Autocorrelação Parcial (FACP) surgem como ferramentas para atender tal necessidade, sobretudo, se o modelo a ser identificado (estimado) for representado por um processo puro Auto Regressivo (AR) de ordem “ p ” ou um processo puro de Médias Móveis (MA) de ordem “ q ”. Porém, caso o modelo seja representado por um processo misto Auto Regressivo de Médias Móveis (ARMA), tanto a FAC quanto a FACP teóricas possuem infinitos valores diferentes de zero, o que dificulta substancialmente a identificação de forma mais precisa do modelo de série temporal, oriundo deste tipo de processo misto ARMA a partir da FAC e FACP, tal como afirma Cryer e Chan (2008).

Com o intuito de suprir a necessidade de identificação dos processos mistos ARMA, várias ferramentas metodológicas têm sido propostas com o decorrer dos anos, dentre as quais se podem citar o método de Canto desenvolvido por Becuinj et. al. (1980), o método da Função de Autocorrelação Estendida (FACE) defendido por Tsay e Tiao (1984), e ainda, o método da Correlação Canônica Minimizado (SCAN) proposto por Tsay e Tiao (1985), dentre outras propostas. Dentro deste contexto, a Função de Autocorrelação Estendida se apresenta como uma excelente proposta metodológica, devido as excelentes propriedades de amostragem que garantem uma menor margem de erro, sobretudo para tamanhos amostrais tendendo ao infinito, tal como, foi comprovado mediante simulações

comparativas realizadas por Chan (1999), além de esta função realizar a identificação de forma mais precisa mediante um teste de significância sobre “ p ” e “ q ”.

Segundo Cryer e Chan (2008), a Função de Autocorrelação Estendida é um método que parte da suposição de que, se é conhecido o componente Auto Regressivo de um modelo misto ARMA, ao filtrar a auto regressão deste modelo a partir dos resultados observados da série temporal de um processo puro de médias móveis com ordem “ q ”, o qual beneficia a propriedade de corte na sua Função de Autocorrelação Amostral, então, os coeficientes do processo puro auto regressivo com ordem “ p ” podem ser estimados mediante uma sequência finita de regressões. Na Equação (A.47) é apresentado um modelo teórico exato que descreve um processo misto ARMA(1;1),

$$Z_i = \phi Z_{i-1} + e_i - \theta e_{i-1}. \quad (\text{A.47})$$

Segundo ainda Cryer e Chan (2008), uma regressão linear simples (Equação (A.47)) de Z_i sobre Z_{i-1} ($i = 1; 2; \dots; n, n \leq T$) resulta em um estimador inconsistente de ϕ , mesmo para uma infinidade de dados. Além disso, o coeficiente teórico desta regressão é dado pela Equação (A.48) e não por ϕ .

$$\rho_1 = \frac{(\phi - \theta) \times (1 - \phi \theta)}{(1 - 2 \phi \theta + \theta^2)}. \quad (\text{A.48})$$

Porém, os resíduos da regressão representada pela Equação (A.47), necessariamente contém informações sobre os erros (e_i) do processo de modelagem. Na sequência, uma segunda regressão múltipla é realizada, para Z_i sobre Z_{i-1} , e sobre o lag 1 dos resíduos da primeira regressão. O coeficiente de Z_{i-1} na segunda regressão é denotado por $\tilde{\phi}$, o qual representa um estimador consistente de ϕ . Cryer e Chan (2008) definem ainda, um processo estocástico como na Equação (A.49),

$$W_i = Z_i - \tilde{\phi} Z_{i-1}, \quad (\text{A.49})$$

que representa um modelo MA(1) aproximadamente. Para um modelo ARMA(1;2), uma terceira regressão de Z_i, Z_{i-1} , sobre o lag 1 (um) dos resíduos da segunda regressão e sobre o lag 2 (dois) dos resíduos da primeira regressão resulta no coeficiente de Z_{i-1} , que é um estimador consistente de ϕ .

Definição A.2. (Consistência): Uma sequência $\{\tilde{\phi}_n, n = 1, 2, \dots\}$ de estimadores de um

parâmetro ϕ é dita “*consistente*”, desde que, para todo, $\varepsilon > 0$, a Equação (A.50) seja verdadeira.

$$P\left(|\tilde{\phi} - \phi| > \varepsilon\right) \rightarrow 0, \quad (\text{A.50})$$

para $n \rightarrow \infty$. Ou então, $\{\tilde{\phi}_n, n = 1, 2, \dots\}$ é uma sequência consistente de estimadores de ϕ se

$$(i) \lim_{n \rightarrow \infty} E\left(\tilde{\phi}_n\right) = \phi;$$

$$(ii) \lim_{n \rightarrow \infty} Var\left(\tilde{\phi}_n\right) = 0.$$

Neste sentido, de maneira similar, os coeficientes autoregressivos do processo ARMA(p ; q), podem ser estimados de maneira consistente mediante uma sequência de “ q ” modelos de regressão. Em virtude da ordem dos processos autoregressivos e de médias móveis serem desconhecidos a priori, segundo Cryer e Chan (2008), se faz necessário a utilização de procedimentos iterativos, para a estimação dos parâmetros desejados, tal como, dado pela Equação (A.51).

$$W_{i,k,j} = Z_i - \tilde{\phi}_1 Z_{i-1} - \dots - \tilde{\phi}_k Z_{i-k}, \quad (\text{A.51})$$

onde, os resíduos autoregressivos são definidos com os coeficientes autoregressivos estimados iterativamente, assumindo que a ordem do processo AR é k e a ordem do processo MA é j . E ainda, Tsay e Tiao (1984) estabeleceram condições de convergência em probabilidade, aos parâmetros autoregressivos dos estimadores de mínimos quadrados obtidos de forma iterativa, tal como, definido no Teorema 4.7.1, o qual origina consequentemente a Equação (A.52).

Teorema A.1. Supondo que Z_t representa um processo estacionário ARMA(p ; q), neste sentido se pode garantir que,

$$\hat{\phi}_{l(k)}^{(j)} = \phi_l + O_p(n^{-\frac{1}{2}}), \quad (\text{A.52})$$

onde, $l = 1; \dots; k$. Se $k = p$ e $j > q$ ou $k \geq p$ e $j = q$, onde sabe-se que, $\phi_l = 0$ para $l > p$.

Prova A.1. Ver Tiao e Tsay (1983), pp. 866.

Para obter mais informações sobre as propriedades de consistência e convergência dos estimadores de mínimos quadrados ordinários, relacionados aos parâmetros autoregressivos em modelos mistos ARMA, mediante Corolários, Lemas e Teoremas, ver maiores detalhes em Tiao e Tsay (1983).

As autocorrelações amostrais de $W_{i,k,j}$ ($i = 1; 2; \dots; n$, $n \leq T$, $k \geq 0$ e $j \geq 0$) são referentes as autocorrelações amostrais estendidas. Para $k = p$ e $j \geq q$, $W_{i,k,j}$ é um modelo MA(q) aproximadamente, de maneira que, as autocorrelações teóricas de lag “ $q + 1$ ” ou superiores a isso são iguais à zero. Para “ $k > p$ ”, surge um problema de modelagem, para tal quantidade de defazagens na série, que acaba resultando no aumento da ordem do modelo de MA, no processo W para no mínimo “ $k - p$ ” e “ $j - q$ ” parâmetros.

Tsay e Tiao (1984), sugerem um resumo das informações obtidas a partir da função de autocorrelação estendida amostral, mediante uma tabela de dupla entrada com “ k ” linhas e “ j ” colunas, tal que, o k -ésimo elemento e o j -ésimo elemento da tabela são indicados pelo simbolo “X”, se no lag “ $j + 1$ ” da correlação amostral de $W_{i,k,j}$ for significativamente diferente de zero, isto é, o valor absoluto da correlação amostral é maior que a amplitude de $\frac{1,96}{\sqrt{n-k-j}}$, visto que, a correlação amostral assintoticamente segue uma distribuição de probabilidade normal, tal como, se observa na Equação (A.53).

$$\text{FACE} \approx N\left(0; \frac{1}{n-k-j}\right), \quad (\text{A.53})$$

para $n \rightarrow \infty$, se os W 's são aproximadamente um processo de MA(j) e 0 (zero) caso contrário.

Na Tabela (A.2), se pode verificar que o processo MA($p; q$) terá um modelo teórico representado por uma matriz triangular onde os elementos acima da diagonal, ou seja, na parte superior esquerda são todos iguais à zero, indicando assim um vértice para a ordem do modelo ARMA. Analisando a Tabela (A.2), se torna possível identificar a forma padrão de um modelo ARMA (1;1). Na parte superior esquerda da tabela se verifica o vértice do triângulo de zeros, que é marcado com o simbolo 0*, o qual está localizado em $p = 1$ (linha 1) e $q = 1$ (coluna 1), que por conseguinte representa a indicação na forma exata de um processo ou modelo ARMA(1;1).

Tabela A.2 - Exemplo do Processo de Identificação Pela Função de Autocorrelação Estendida Teórica Para Um Modelo ARMA(1;1), tal como, Descrito em Cryer e Chan (2008).

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1	x	0*	0	0	0	0	0	0	0	0	0	0	0	0
2	x	x	0	0	0	0	0	0	0	0	0	0	0	0
3	x	x	x	0	0	0	0	0	0	0	0	0	0	0
4	x	x	x	x	0	0	0	0	0	0	0	0	0	0
5	x	x	x	x	x	0	0	0	0	0	0	0	0	0
6	x	x	x	x	x	x	0	0	0	0	0	0	0	0
7	x	x	x	x	x	x	x	0	0	0	0	0	0	0

A.7.1 - Características da Função de Autocorrelação Estendida

1. Representa uma função eficiente quanto à determinação da ordem de um processo *Misto* que é caracterizado por um modelo ARMA(p ; q);
2. Os seus resultados são apresentados mediante uma tabela de dupla entrada com “ k ” linhas por “ j ” colunas;
3. Na Tabela (A.2) as linhas indicam a ordem “ p ” do componente autoregressivo do modelo;
4. Na Tabela (A.2) as colunas indicam a ordem “ q ” da parte de médias móveis do modelo;
5. Na Tabela (A.2) a primeira linha contém a função de autocorrelação amostral, a segunda linha contém a primeira FACE amostral, etc.

Com o auxílio de uma generalização das equações de Yule-Walker (ver, Cryer e Chan, 2008, pg. 72), se chega à proposta conceitual inicial da função de autocorrelação estendida, a qual é descrita a partir de duas etapas. Na primeira etapa, a obtenção de estimadores consistentes para os coeficientes do modelo autoregressivo é exigida. Encontrados os estimadores, se torna possível transformar uma série temporal originalmente representada por um modelo ARMA, em uma série temporal indicada segundo um modelo de médias móveis puro. A segunda etapa deste procedimento consiste, em utilizar a função de auto-

correlação amostral sobre o modelo de médias móveis resultante da transformação, para assim poder identificar a ordem do termo q .

Ao calcular a função de autocorrelação estendida implicitamente um teste estatístico está sendo realizado, com o objetivo de comprovar a um nível de significância (α) de probabilidade, se a hipótese nula (H_0) deve ser rejeitada. O resultado de cada $W_{i,k,j}$, para $k = 0, \dots, 5$ e $j = 0, \dots, 5$, deve ser estatisticamente diferente de zero, ou seja, se está testando as hipóteses apresentadas na Equação (A.54). Neste contexto, a regra de decisão ou tomada de decisão do teste será: caso $P\text{-Valor} > \alpha$, se deve tomar a decisão de não rejeitar a hipótese H_0 , caso contrário, se deve rejeitar H_0 à um nível de significância α , pré-estabelecido pelo especialista. Os valores mais usuais na literatura científica para α são 1%; 5% e 10%, onde esta probabilidade pode assumir diversos valores de interesse ao universo que está sendo estudado, desde que, o valor de $\alpha \in [0; 1]$.

$$\begin{cases} H_0 : \text{ARMA}(p; q) = 0, \text{ os parâmetros } p \text{ e } q \text{ são iguais a zero;} \\ H_1 : \text{ARMA}(p; q) \neq 0, \text{ os parâmetros } p \text{ e } q \text{ são diferentes de zero,} \end{cases} \quad (\text{A.54})$$

para, $p = 0; \dots; 5$ e $q = 0; \dots; 5$, os parâmetros AR e MA, respectivamente.

Exemplo A.1. Seja a série temporal do consumo total de energia elétrica (MWh), durante seis anos observados mensalmente em um sistema elétrico Brasileiro. Deseja-se identificar a ordem do modelo de série temporal mediante a função de autocorrelação estendida (Equação, A.51).

Tabela A.3 - Resultados da $W_{i,k,j}$ Para o Consumo Total de Energia Elétrica (MW/h).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,631	0,310	0,033	-0,058	-0,084	-0,116
AR(1)	0,088	0,204	0,039	-0,057	-0,044	-0,047
AR(2)	-0,305	0,248	-0,112	0,014	0,045	-0,104
AR(3)	-0,038	-0,177	-0,098	0,011	0,073	0,011
AR(4)	-0,177	-0,038	-0,323	0,050	0,053	0,057
AR(5)	-0,334	0,029	-0,400	0,044	0,118	0,015

Analisando os resultados da função $W_{t,k,j}$ nas Tabelas (A.3 e A.4), a identificação realizada retornou sete modelo significativos, ou seja, estatisticamente diferentes de zero a um nível de significância de 10%. Porém, considerando apenas os modelos onde $p > 0$ e $q > 0$, a identificação processada sugeriu os modelos: ARMA(2;1), ARMA(4;2) e ARMA(5;2) de forma significativa.

Tabela A.4 - Probabilidades de Significância de $W_{i,k,j}$ Para os Resultados da Tabela (A.3).

Lag	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	0,0001*	0,0708*	0,8560	0,7501	0,6412	0,5230
AR(1)	0,4965	0,1746	0,8378	0,6833	0,7395	0,7199
AR(2)	0,0192*	0,0633*	0,4486	0,9294	0,7896	0,4320
AR(3)	0,7706	0,1777	0,5335	0,9455	0,6316	0,9383
AR(4)	0,1809	0,7910	0,0175*	0,7594	0,7053	0,7290
AR(5)	0,0126*	0,8445	0,0068*	0,7914	0,4573	0,9224

Nota: * Valor significativo para $\alpha = 10\%$.

A.8 - Erros de Previsão dos Modelos de Série Temporal Ajustados

Após a determinação de um modelo de série temporal, a sua capacidade preditiva deve ser aferida, a partir de observações não inclusas na amostra que deu origem ao modelo em questão. Em série temporal, se deve satisfazer a suposição de que, os resíduos ou erros de previsão do modelo de série temporal ajustado, obrigatoriamente devem apresentar aleatoriedade em torno de uma média zero, assim como estes erros precisam possuir variância constante e não serem correlacionados ao longo do tempo. Neste sentido, a avaliação do desempenho preditivo de um modelo de série temporal deve atender aos passos descritos a seguir.

Passo 1: Estimar um modelo de série temporal até um horizonte de tempo “ i ” ($i = 1; 2; \dots; n$), onde $n \leq T$ (tamanho da série), desconsiderando as últimas n observações da série histórica;

Passo 2: Calcular previsões para as últimas n observações da série histórica;

Passo 3: Comparar os valores das previsões obtidos no passo 2, com os valores reais da série.

Para medir a exatidão das previsões, a maioria dos métodos empregam o resíduo em seus cálculos, ou seja, a diferença entre o valor observado na série no instante de tempo “ i ” em relação ao valor da previsão neste mesmo instante ($e_i = Z_i - \hat{Z}_i$). Dentre os métodos mais conhecidos para se medir a performance preditiva se pode citar: o erro quadrático médio (MSD), o erro absoluto médio (MAD), o erro percentual absoluto médio (MAPE) e a raiz do erro médio quadrático (RMSE), os quais são descritos pelas Equações (A.55)

- (A.58), respectivamente, a seguir.

$$\text{MSD} = \left(\frac{1}{T}\right) \times \sum_{i=1}^n (Z_i - \hat{Z}_i)^2 = \left(\frac{1}{T}\right) \times \sum_{i=1}^n (e_i)^2. \quad (\text{A.55})$$

$$\text{MAD} = \left(\frac{1}{n}\right) \times \sum_{i=1}^n |(Z_i - \hat{Z}_i)| = \left(\frac{1}{n}\right) \times \sum_{i=1}^n |e_i|. \quad (\text{A.56})$$

$$\text{MAPE} = \left(\frac{100}{T}\right) \times \sum_{i=1}^n \left| \frac{(Z_i - \hat{Z}_i)}{Z_i} \right| = \left(\frac{100}{T}\right) \times \sum_{i=1}^n \left| \frac{(e_i)}{Z_i} \right|. \quad (\text{A.57})$$

$$\text{RMSE} = \sqrt{\left(\frac{1}{T}\right) \times \sum_{i=1}^n (Z_i - \hat{Z}_i)^2} = \sqrt{\left(\frac{1}{T}\right) \times \sum_{i=1}^n (e_i)^2}, \quad (\text{A.58})$$

onde n é a quantidade observada da série histórica, T o tamanho de trajetória da série temporal, $i = 1; 2; \dots; n$, ($n \leq T$) é o instante observado da série temporal, Z_i é o valor observado no instante i e \hat{Z}_i indica a previsão do modelo de série temporal ajustado à série no instante i .

Quanto menor for o erro de previsão melhor estará ajustado o modelo a série de dados, o MSD é geralmente a medida de erro mais utilizada, pois penaliza desproporcionalmente os grandes erros (positivos ou negativos), porém, o MAD devido calcular os resíduos de forma absoluta, isto é, na mesma ordem de grandeza dos valores originais da série, é o mais indicado em alguns problemas de série temporal, por exemplo, quando combinada aos gráficos de controle da qualidade.

Destaca-se que no meio acadêmico não existe consenso a respeito do melhor método a utilizar. Se o analista desejar obter um modelo de variância mínima recomenda-se o MSD. Se for possível desprezar alguns erros elevados, o MAD pode ser a melhor opção. “O MAPE é indicado quando se deseja comparar a precisão de duas séries temporais distintas” (Brockwell e Davis, (2002)).

Recentemente muitos pesquisadores tem apresentado propostas de algoritmos de previsão para série temporal em um ambiente *fuzzy*, adotando como medida de performance preditiva, a raiz do erro médio quadrático, visto que, devido esta medida de acurácia representar uma função do erro quadrático médio, os erros com valores absolutos muito elevados são penalizados de forma desproporcional, além de expressar o seu resultado na

mesma ordem de grandeza dos dados observados na série histórica original, o que também contribui para uma melhor interpretação dos resultados obtidos, e que deverá ser comparado ao de outros modelos de previsão ajustados.

Neste trabalho, assim como em Carvalho Jr. e Costa Jr. (2017), foi utilizada a medida de acurácia conhecida como escala de erro médio absoluto (MASE), tal como definido na Equação (A.59) e visto em Hyndman e Koehler (2006), no intuito de contornar problemas incorporados a algumas medidas de acurácia das previsões, por se acreditar que algumas dessas medidas como as que foram citadas acima, e geralmente são adotadas sobre modelos de previsão *fuzzy*, como por exemplo, medidas de erro relativas e medidas baseadas em erros relativos, onde a primeira medida segue uma distribuição de probabilidade, com média indefinida e variância infinita, já a segunda medida de erro só é possível o cálculo a partir de várias previsões sobre a mesma série de dados, o que impossibilita esta medida de ser aplicada a previsões fora da amostra observada, de acordo com Hyndman e Koehler (2006).

$$\text{MASE} = \text{Média}(|q_i|), \quad (\text{A.59})$$

onde, $|q_i|$ representa o valor absoluto de q_i , o qual é expresso pela Equação (A.60).

$$q_i = \frac{e_i}{\frac{1}{n-1} \sum_{i=2}^n |Z_i - Z_{i-1}|}, \quad (\text{A.60})$$

tal que, $e_i = Z_i - \hat{Z}_i$ indica o erro observado no instante i , tal que, $i = 1, \dots, n$; ($n \leq T$), onde n representa a quantidade observada da série histórica, T o tamanho de trajetória da série temporal, Z_i e Z_{i-1} , os valores da série temporal no instante i e $i-1$, respectivamente.

Quanto aos erros retornados pelas previsões do modelo de série temporal, estes devem satisfazer os pressupostos de independência, distribuídos identicamente segundo uma função densidade de probabilidade normal, com média zero e variância constante, ou seja, há necessidade de uma inspeção e avaliação dos erros de previsão obtidos pelo modelo ajustado, tal como, em Duru (2010) e Hyndman e Koehler (2006), para que desta forma as previsões realizadas em um horizonte esperado, retornem resultados confiáveis e sem a influência direta de aspectos externos no cenário da variável estudada, o que em série temporal é conhecido como um

erro representando um ruído branco, no modelo de previsão estimado por algum método de identificação paramétrica.

A.9 - Critérios de Informação Para os Modelos de Série Temporal

Quando se determina um modelo de série temporal, com o intuito de representar o comportamento de um determinado fenômeno de interesse (variável), em um intervalo de tempo específico, a variabilidade inerente a este fenômeno obrigatoriamente deve ser capturada pelo modelo ajustado aos dados, ou seja, o modelo deve representar de maneira fiel o real comportamento dos dados utilizados no processo. Neste sentido, como se trata de um modelo capaz de gerar previsões sobre uma característica de interesse, se faz necessário que os erros (resíduos) originados pelo modelo ajustado sejam os menores possíveis, e ainda, que estes não interfiram de forma relevante nas previsões que possam vir a ser realizadas pelo modelo ajustado. Logo, estudar o comportamento dos resíduos de um modelo ajustado, significa avaliar a sua capacidade preditiva, que em muitos casos se mostra com eficiência equivalente, em vários modelos distintos. Portanto, o critério da informação tem por finalidade, mensurar a qualidade do ajuste do modelo em questão, como ainda, penalizar a inclusão de parâmetros extras no modelo estabelecido.

Caso um modelo possua um número elevado de parâmetros, há possibilidade do ajuste (estimação) deste, ser superior ao de um modelo com um número inferior de parâmetros, em relação a uma mesma característica de interesse, no entanto, isso não significa que necessariamente quanto ao critério da informação a sua avaliação será superior. Vale ressaltar que, em processos de modelagem existe o critério da parcimônia, ou seja, um modelo com o menor número de parâmetros possíveis para descrever o comportamento de uma característica específica, sem desprezar a capacidade preditiva e a adequação em termos do comportamento dos resíduos, do modelo ajustado. Assim, deve-se indicar como melhor modelo dentre os ajustados a uma série de dados, aquele que apresente o menor critério da informação. A seguir, destacam-se três critérios da informação mais utilizados em processos de modelagem em série temporal, dentre os disponíveis na literatura científica.

1. Critério da Informação de Akaike's: (*AIC*)

Segundo este critério, o modelo de série ajustado que deve ser escolhido para representar

uma série de dados, obrigatoriamente deve minimizar a expressão dada pela Equação (A.61).

$$AIC = -2 \log[Máxima Verossimilhança] + 2k, \quad (A.61)$$

onde, $k = p + q + 1$, se o modelo contém um intercepto ou constante e $k = p + q$, caso contrário. E ainda, “ p ” e “ q ” são chamados parâmetros autoregressivo e de médias móveis, respectivamente.

Segundo Cryer e Chan (2008), caso o número de parâmetros do modelo de série temporal seja muito elevado, o critério de *Akaike's* se torna um estimador viciado. No entanto, Hurvich e Tsai (1989) mostram que este problema de viés do estimador pode ser contornado, mediante a adição de outro termo de penalização estocástica, no critério *AIC*, o que resulta em um critério de *Akaike's corrigido*, denotado por *AIC_C*, como mostra a Equação (A.62).

$$AIC_C = AIC + \frac{2(k+1)(k+2)}{n-k-2}, \quad (A.62)$$

onde, n é o tamanho amostral efetivo da série temporal observada; há uma exclusão da variância do erro; $k = p + q + 1$, se o modelo contém um intercepto ou constante e $k = p + q$, caso contrário. E ainda, “ p ” e “ q ” são ditos parâmetros autoregressivo e de médias móveis, respectivamente.

Caso os dados referentes à série temporal sigam uma distribuição de probabilidade normal, e ainda, para estimativas de máxima verossimilhança sendo utilizadas nos parâmetros, do modelo de série temporal ajustado aos dados, o critério *AIC* passa a ser indicado pela Equação (A.63).

$$AIC = [n \times \ln(RSS)] + 2k, \quad (A.63)$$

tal que, n é o tamanho da série temporal observada; *RSS* representa a soma de quadrados dos resíduos e k é tal como definido nas Equações (A.61) e (A.62), que foram descritas anteriormente.

2. Critério da Informação Bayesiana: (*BIC*)

Outra forma de avaliar o ajuste de modelos autoregressivos de médias móveis, é mediante o critério da informação bayesiana *schwarz*, que é expresso pela Equação (A.64).

$$BIC = -2 \log[Máxima Verossimilhança] + k \log(n), \quad (A.64)$$

onde, n e k são definidos de forma idêntica ao que foi apresentado na Equação (A.62).

Para Cryer e Chan (2008), caso a série temporal se comporte segundo um processo autoregressivo de médias móveis $(p; q)$ puro, então a minimização empregada pelo critério da informação bayesiana será consistente, ou seja, a mediada que o tamanho da amostra aumente o valor deste critério se torna cada vez mais preciso. Porém, caso o processo em questão não se trate de um ARMA $(p; q)$ finito, então a minimização do critério AIC em relação a uma grande classe de modelos ARMA $(p; q)$, irá convergir para um modelo ideal, que melhor represente o processo em análise, entre todos os modelos de série temporal ajustados pelos dados históricos de interesse.

3. Critério da Informação de Schwarz: (SIC)

Na literatura há uma outra opção de avaliação do ajuste de um modelo de série temporal, que é denominado de critério da informação de *schwarz* e dado pela Equação (A.65).

$$SIC = [n \times \ln(RSS)] + [k \times \ln(n)], \quad (A.65)$$

tal que, n e k são definidos como na Equação (A.62), e ainda, RSS é definido na Equação (A.63).

Observação A.1. É importante ressaltar que o efeito de comparação entre os critérios apresentados acima, necessita que o “ n ” seja obrigatoriamente o mesmo, em todas as séries de dados para as quais se ajuste um modelo de série temporal e que consequentemente resultem em um modelo autoregressivo de médias móveis, o que necessariamente garante o efeito de comparação.

A.10 - Conclusão do Apêndice A

Neste capítulo, os conceitos básicos sobre série temporal clássica foram discutidos, desde uma formalização generalizada, passando pelo conceito de tendência, estacionariedade, sazonalidade e diferenciação, até características particulares de quatro funções que auxiliam na identificação dos processos autoregressivos. Especificamente para função de autocorrelação amostral, função de autocorrelação parcial, função de autocorrelação inversa e função de autocorrelação estendida, se destacou a forma *dual* de identificação inerente às duas funções iniciais mencionadas anteriormente, tal como, o objetivo, definições básicas e particularidades de cada uma destas funções. Testes para a significância

estatística destas funções e as características de um bom ajuste, para um modelo série temporal que represente sobretudo um processo ARIMA foram avaliados. Destacaram-se também os erros de previsão e os critérios de informação, em um modelo de série temporal ajustado, visto que, estes erros se apresentam relacionados a um modelo identificado, com o objetivo de previsão de dados, logo, os erros de previsão e critérios de informação são uma forma de estabelecer, uma medida de precisão e qualidade do ajuste, respectivamente, em um modelo de série temporal, como foi possível constatar neste capítulo. Portanto, as medidas de acurácia dos modelos de previsão mediante uma série temporal, que foram apresentadas neste capítulo são de suma importância, para sustentar, fundamentar e consolidar a metodologia proposta, que passa pelos métodos de identificação que serão abordados nos próximos capítulos.

No próximo capítulo, o modelo de série temporal no ambiente *fuzzy* será alvo de estudo, com o auxílio para o seu desenvolvimento das definições básicas, conceitos fundamentais, algoritmos de previsão e os métodos de identificação dos modelos de série temporal *fuzzy*, que estão disponíveis e consolidados na literatura científica. Com isto, acredita-se ser possível estabelecer e propor novas metodologias a serem utilizadas, como ferramentas de suporte na identificação em série temporal *fuzzy*. Neste sentido, algumas propostas inéditas de ferramentas para identificação paramétrica serão apresentadas, considerando os conceitos alcançados pela metodologia dos conjuntos *fuzzy*, combinados com a função de autocorrelação estendida. Desta forma, o objetivo será identificar de forma eficiente modelos autoregressivo *fuzzy*, com ou sem sazonalidade estocástica incorporada a série temporal em questão, o que invariavelmente suscitará uma forma híbrida de identificação sobre série temporal *fuzzy*, corroborando assim com os métodos encontrados atualmente na literatura especializada, algo que hipoteticamente já vem sendo considerado neste trabalho.

Referências Bibliográficas

1. Abdollahzade, M., Miranian, A., Hassani, H. and Iranmanesh, H. A new hybrid enhanced local linear neuro-fuzzy model based on the optimized singular spectrum analysis and its application for nonlinear and chaotic time series forecasting, *Information Sciences*, vol. 295, pp. 107-125, 2015.
2. Abdulshahed, A. M., Longstaff, A. P. and Fletcher, S. The application of ANFIS prediction models for thermal error, *Applied Soft Computing*, vol. 27, pp. 158-168, 2015.
3. Aguirre, L. A. *Introdução à identificação de sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais/ Luis Antonio Aguirre*. 3. Ed. rev. e ampl. - Belo Horizonte: Editora UFMG, 2007.
4. Alwan, L. C. e Roberts, H. V. Times series modeling for statistical process control, *Journal of Business Economics Statistics*, vol. 16, pp. 83-89, 1988.
5. Aström. *Introduction to Stochastic Control Theory*, Academic Press, N.Y, 1970.
6. Askari, S. and Montazerin, N. A high-order multi-variable fuzzy time series forecasting algorithm based on fuzzy clustering, *Expert Systems with Applications*, vol. 42, pp. 2121-2135, 2015.
7. Bai, E., Wong, W. K., Chu, W. C., Xia, M. and Pan, F. A heuristic time-invariant model for fuzzy time series forecasting, *Expert Systems with Applications*, vol. 38, pp. 2701-2707, 2011.
8. Baraldi, P., Maio, F. D., Genini, D. and Zio, E. Reconstruction of missing data in multidimensional time series by fuzzy similarity, *Applied Soft Computing*, vol. 26, pp. 1-9, 2015.
9. Barnard, G. A. Control Chart and Stochastic Process, *Journal of the Royal Statistical Society, Series B*, vol.21, pp. 239-271, 1959.
10. Becuinj, M., Gourieroux, S., and Monfort, A. Identification of a Mixed Autoregressive-Moving Average Process: The Corner Method. In *Time Series*, Edited by O. D. Anderson, pp. 423-436, Amsterdam: North-Holland, 1980.
11. Bilobrovec, M., Marçal, R. F. M., & Kovaleski, J. L. *Implementação de Um Sistema de Controle Inteligente Utilizando a Lógica Fuzzy*. XI Simpósio de Engenharia de Produção, Bauru, São Paulo, Brasil, 2004.
12. Bitar, S. D. B., *Modelo de Programação Matemática Fuzzy para Auxílio ao Planejamento da Expansão dos Sistemas Elétricos Isolados na Amazônia*. Tese de Doutorado. Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica. 2009.
13. Boukezzoula, R., Galichet, S., Foulloy, L. MIN and MAX Operators for Fuzzy Intervals and Their Potential Use in Aggregation Operators, *IEEE Transactions Fuzzy Systems*, vol. 15, pp. 1135-44, 2007.
14. Bowden, C. D. e Graybill, A. F. Confidence Bands of Uniform and Proportional Width for Linear Models, *American Statistical Association Journal*, March, pp. 182-198, 1966.
15. Bowden, C. D. e Graybill, A. F. Linear Segment Confidence Bands for Simple Linear Models, *American Statistical Association Journal*, June, pp. 403-408, 1967.
16. Box, G. P, Jenkins, G. M. and MacGregor. Some Recent Advances in Forecasting and Control, *Applied Statistics*, vol.23, pp. 158-179, 1974.
17. Box, G. E., Jenkins, G. M., Reinsel, G. C. *Time series analysis: Forecasting and Control*, Wiley.com, 2013.
18. Bradshaw, C. W. A Fuzzy Set Theory Interpretation of Economic Control Limits, *European Journal of Operation Research*, vol. 13, pp. 403-408, 1983.
19. Brockwell, P. J. and Davis, R. A. *Introduction to Time Series and Forecasting*, Springer-Verlag, Second Edition, New York, 2002.
20. Bulut, E. Modeling Seasonality Using the Fuzzy Integrated Logical Forecasting (FILF) Approach, *Expert Systems with Applications*, vol. 41, pp. 1806-1812, 2014.
21. Cai Q., Zhang, D., Zheng, W., Leung, S. C. H. A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression, *Knowledge-Based Systems*, vol. 74, pp. 61-68, 2015.
22. Cai, Q., Zhang, D., Wu, B. and Leung, S. C. H. A Novel Stock Forecasting Model Based On Fuzzy Time Series and Genetic Algorithm, *Procedia Computer Science*, vol. 18, pp. 1155-1162, 2013.
23. Campos, R. J. *Previsão de Séries Temporais Com Aplicações a Séries de Consumo de Energia Elétrica*, Dissertação de Mestrado: UFMG, Belo Horizonte, 2008.
24. Cardoso, G. C. *Modelo de Previsão Baseado em Agrupamento e Base de Regras Nebulosas*, Dissertação de Mestrado: UNICAMP, Campinas, 2003.
25. Carvalho Jr. J. G., Costa Jr. C. T. Identification method for fuzzy forecasting models of time series, *Applied Soft Computing*, vol. 50, pp. 166-182, 2017.

26. Carvalho Jr., J. G., Costa Jr., C. T., Lago Neto, J. C. Fuzzy Time Series With “ n ” Differences For Identification Model Order by Autocorrelation Function, Learning and Nonlinear Models vol. 2 no. 10, pp. 109-118, 2012.
27. Chan, W. S. A Comparison of Some Pattern Identification Methods for Order Determination of Mixed ARMA Models. Statistics and Probability Letters, vol. 42, pp. 69-79. 1999.
28. Chang, S. I. and Aw, C. A. A Neural Fuzzy control Chart for Detecting and Classifying Process Mean Shifts, International Journal of Production Research, vol. 34, pp. 2265-2278, 1996.
29. Chen, S. M. Forecasting Enrollments Based on Fuzzy Time Series. Fuzzy Sets and Systems, vol. 81, pp. 311-319, 1996.
30. Chen, S. M. Forecasting Enrollments Based on High-Order Fuzzy Time Series. Cybernetics and Systems, vol. 33, pp. 1-16, 2002.
31. Chen, L.-H. A Demerit Control Chart With Linguistic Weights, Journal of Intelligent Manufacturing, vol. 16, pp. 349-359, 2003.
32. Chen, S. M. and Chung, N. Yi. Forecasting Enrollments of Students by Using Fuzzy Time Series and Genetic Algorithms, Information and Management Sciences, vol. 3, pp. 1-17, 2006.
33. Chen, S. M. and Hsu, C. C. A New Method to Forecast Enrollments Using Fuzzy Time Series, International Journal of Applied Science and Engineering, vol. 2, pp. 234-244, 2004.
34. Chen, S. M., and Tanuwijaya, K. Fuzzy Forecasting Based on High-Order Fuzzy, Logical Relationships and Automatic Clustering Techniques. Expert Systems with Applications, vol. 38, pp. 15425-15437, 2011.
35. Chen, M. Y. and Chen, B. T. A hybrid fuzzy time series model based on granular computing for stock price forecasting, Information Sciences, vol. 294, pp. 227-241, 2015.
36. Chen, M. Y. and Chen, B. T. Online fuzzy time series analysis based on entropy discretization and a Fast Fourier Transform, Applied Soft Computing, vol. 14, pp. 156-166, 2014.
37. Cheng, C. H., Wang, J. W., & Cheng, G. W. Multi-Attribute Fuzzy Time Series, Method Based on Fuzzy Clustering. Expert Systems with Applications, vol. 34, pp. 1235-1242, 2008a.
38. Cheng, C. H., Chen T. L., Teoh, H. J., Chiang, C. H. Fuzzy Time-Series Based on Adaptive Expectation Model for TAIEX Forecasting. Expert Systems with Applications vol. 34, pp. 1126-1132, 2008b.
39. Cleveland, W. S. The Inverse Autocorrelations of a Time Series and Their Applications, Technometrics vol. 14, pp. 277-298, 1972.
40. Conejo, A. J., Contreras, J., Espínola, R. and Plazas, M. A. Forecasting Electricity Prices for a Day-Ahead Pool-Based Electric Energy Market, International Journal of Forecasting, doi. 10.1016.2004.12.005, 2004.
41. Cryer, J. D., Chan, K. -S. Time Series Analysis - With Applications in R, Springer, Second Edition, 2008.
42. Cuaresma, J. C., Hlouskova, J., Kossmeier S., and Obersteiner, M. Forecasting Electricity Spot-Prices Using Linear Univariate Time Series Models, Applied Energy, vol. 77, pp. 87-106, 2004.
43. Darbellay, G. A. and Slama, M. Forecasting the Short Term Demand for Electricity. Do Neural Networks Stand a Better Chance, International Journal of Forecasting, vol. 16, pp. 71-83, 2000.
44. Dong, Y., Wang, J., Jiang, H. and Wu, J. Short-Term Electricity Price Forecast Based On The Improved Hybrid Model, Energy Conversion and Management, vol. 52, pp. 2987-2995, 2011.
45. Duru, O. A Fuzzy Integrated Logical Forecasting Model For Dry Bulk Shipping Index Forecasting: An Improved Fuzzy Time Series Approach, Expert Systems with Applications, vol. 37, pp. 5372-5380, 2010.
46. Duru, O., Bulut, E., Yoshida, S. A Fuzzy Extended DELPHI Method for Adjustment of Statistical Time Series Prediction an Empirical Study on Dry Bulk Freight Market Case, Expert Systems with Applications, vol. 39, pp. 840-848, 2012.
47. Efendi R., Ismail Z., Deris M. M., A new linguistic out-sample approach of fuzzy time series for daily forecasting of Malaysian electricity load demand, Applied Soft Computing, vol. 28, pp. 422-430, 2015.
48. Egrioglu, E., Aladag, C. H., Yolcu, U., Basaran, M. A., and Uslu, V. R. A New Hybrid Approach Based on SARIMA and Partial High Order Bivariate Fuzzy Time Series Forecasting Model, Expert Systems with Applications, vol. 36, pp. 7424-7434, 2009.
49. Egrioglu, E., Aladag, C. H., Basaran, M. A., Yolcu, U. and Uslu, V. R. A New Approach Based On The Optimization of The Length of Intervals In Fuzzy Time Series, Journal of Intelligent and Fuzzy Systems, vol. 22, pp. 15-19, 2011.
50. Egrioglu, E., Aladag, C. H., Yolcu, U. Fuzzy Time Series Forecasting With a Novel Hybrid Approach Combining Fuzzy C-Means and Neural Networks, Expert Systems with Applications, vol. 40, pp. 854-857, 2013a.
51. Egrioglu, E., Aladag, C. H., Yolcu, U., Corba, B. S., Cagcag, O. Fuzzy Time Series Method Based on Multiplicative Neuron Model and Membership Values, American Journal of Intelligent Systems, vol. 3, pp. 33-39, 2013b.
52. Egrioglu, E., PSO-based high order time invariant fuzzy time series method-Application to stock exchange data, Economic Modelling, vol. 38, pp. 633-639, 2014.
53. H. Feng, H. Wang, J. Guo, F. Zhang, A modified method of forecasting enrollments based on fuzzy time series. In 2nd International Conference on Soft Computing in Information Communication Technology, Atlantis Press, 2014.
54. Fu, F. P., Chi, K., Che, W. G. and Zhao, Q. J. High-order Difference Heuristic Model of Fuzzy Time Series Based on Particle Swarm Optimization and Information Entropy for Stock Markets, IEEE, International Conference On Computer Design And Appliations, vol. 2, pp. 210-215, 2010.

55. Gangwar, S. S. and Kumar S. Partitions Based Computational Method For High-Order Fuzzy Time Series Forecasting, *Expert Systems with Applications*, vol. 39, pp. 12158-12164, 2012.
56. Gaxiola, F., Melin, P., Valdez, F. and Castillo, O. Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction, *Information Sciences*, vol. 260, pp. 1-14, 2014.
57. Gonzales Chaves, S., Xiberta, B. J. and Llana, C. H. Forecasting of energy production and consumption in Asturias (Northern Spain), *Energy*, vol. 24, pp. 183-198, 1999.
58. Guan, J. W. and Bell, D. A. *Evidence Theory and Its Applications*, Vols. I and II, North-Holland, NY, 1991.
59. Hammond, G. P. and Mackay R. M. Projection of UK oil and gas supply and demand to 2010, *Applied Energy*, vol. 44, pp. 93-112, 1993.
60. Hell, M. B. *Abordagem Neurofuzzy para Modelagem de Sistemas Dinâmicos Não Lineares*, Tese de Doutorado: UNICAMP, Campinas, 2008.
61. Höppner, J. and Wolff, H. The Design of a Fuzzy-Shewhart Control Chart, Technical Report of the Würzburg Research Group on Quality Control, no. 52, 1995.
62. Huarng, K. H. Heuristic Models of Fuzzy Time Series for Forecasting, *Fuzzy Sets and Systems*, vol. 123, pp. 369-386, 2001a.
63. Huarng, K. H. Effective Lengths of Intervals to Improve Forecasting in Fuzzy Time Series, *Fuzzy Sets and Systems*, vol. 123, pp. 387-394, 2001b.
64. Huarng, K. Yu, H. K. and Hsu, Y. W. A Multivariate Heuristic Model for Fuzzy Time-Series Forecasting, *IEEE Transactions on Systems, Man, and Cybernetics Part-B: Cybernetics*, vol. 37, no. 4, pp. 836-846, 2007.
65. Hurvich, C. M. and Tsai, C. L. Regression and Time Series Model Selection in Small Samples. *Biometrika*, vol. 76, no. 2, pp. 297-307, 1989.
66. Hwang, J. R., Chen, S. M., and Lee, C. H. Handling Forecasting Problems Using Fuzzy Time Series. *Fuzzy Sets and System*, vol. 100, pp. 217-228, 1998.
67. Hylleberg, S. Tests For Seasonal Unit Roots General To Specific Or Specific To General? *Journal of Econometrics*, vol. 69, pp. 5-25, 1995.
68. Hyndman, R. J. and Koehler, A. B. Another Look At Measures of Forecast Accuracy, *International Journal of Forecasting*, vol. 22, pp. 679-688, 2006.
69. Jasim, H. T., Salim A. G. J. and Ibraheem, K. I. A Novel Algorithm to Forecast Enrollment Based on Fuzzy Time Series, *Applications and Applied Mathematics: An International Journal*, vol. 1, pp. 385-397, 2012.
70. Kai, C. and Wen-Gang, C. N-Order Difference Heuristic Model of Fuzzy Time Series Forecasting, *IEEE*, pp. 358-361, 2009.
71. Kim, C. I., Yu, I. K., and Song, Y. H. Prediction of System Marginal Price of Electricity Using Wavelet Transform Analysis, *Energy Conversion and Management*, vol. 43, pp. 1839-1851, 2002.
72. Klir, G. J., and T. Folger. *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, Englewood Cliffs, NJ, 1988.
73. Klir, G. J., and B. Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1995.
74. Kwakernaak, H. Fuzzy random variables. Part I: Definitions and Theorems, *Information Sciences*, vol. 15, pp. 1-29, 1978.
75. Lago Neto, J. C., Correia, J. C, Parente, R. C. P e Pereira, J. R. Modelo de Projeção para o Consumo de Energia Elétrica e Óleo Diesel no Interior do Estado do Amazonas, *Revista Brasileira de Energia*, vol. 12, no. 1, pp. 59-70, 2006.
76. Lago Neto, J. C. *Modelo de Previsão Fuzzy: Uma Alternativa Metodológica Para Sistemas Eletricamente Isolados na Amazônia*, Tese de Doutorado: UFPA, Belém, 2011.
77. Lee, W. J., Hong, J., A hybrid dynamic and fuzzy time series model for mid-term power load forecasting, *International Journal Electrical Power and Energy Systems*, vol. 64, pp. 1057-1062, 2015.
78. Lee, C. T. and Wang, H. F. A Method for Fuzzy Time Series - An Example for Telecommunication Demands, *IFORS'96*, Canada, 1996.
79. Leone Filho, M. A. *Previsão de Carga de Curto Prazo Usando Ensembles de Previsores Seleccionados e Evoluídos por Algoritmos Genéticos*, Dissertação de Mestrado: UNICAMP, Campinas, 2006.
80. Liu, H. T. An Improved Fuzzy Time Series Forecasting Method Using Trapezoidal Fuzzy Numbers. *Fuzzy Optimization Decision Making*, vol. 6, pp. 63-80, 2007.
81. Lu, W., Chen, X., Pedrycz, W., Liu, X. and Yang, J. Using interval information granules to improve forecasting in fuzzy time series, *International Journal of Approximate Reasoning*, vol. 57, pp. 1-18, 2015.
82. Lu, W., Pedrycz, W., Liu, X., Yang, J. and Li, P. The modeling of time series based on fuzzy information granules, *Expert Systems with Applications*, vol. 41, pp. 3799-3808, 2014.
83. Magalhães, M. H. *Redes Neurais, Metodologias de Agrupamento e Combinação de Previsores Aplicados à Previsão de Vazões Naturais*, Dissertação de Mestrado: UNICAMP, Campinas, 2004.
84. Marszalek, A., Burczynski, T. Modeling and forecasting financial time series with ordered fuzzy candlesticks, *Information Sciences*, vol. 273, pp. 144-155, 2014.
85. Medeiros, L. *Previsão do Preço Spot no Mercado de Energia Elétrica*. Tese de Doutorado, Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil, 2003.

86. Miller, G. A. The Magical Number Seven, Plus or Minus Two: Some Limits On Our Capacity for Processing Information. *Psychological Review*, vol. 2, pp. 343-352, 1994.
87. Morettin, P. A. and Toloi, C. M. C. *Análise de Séries Temporais*. Edgard Blucher, São Paulo, Brasil, 2004.
88. Nasser, A. T., Mikaeily Narsis, K. A. and Rasol, S. M. Signed Decomposition of Fully Fuzzy Linear Systems, *Applications and Applied Mathematics: An International Journal* 1, 2008.
89. Ortega, N. R. S. *Aplicação da Teoria de Conjuntos Fuzzy a Problemas da Biomedicina*. Tese de Doutorado, Instituto de Física, Universidade de São Paulo, São Paulo, Brasil, 2001.
90. Pucciarelli, A. J. *Modelagem de Séries Temporais Discretas Utilizando Modelo Nebuloso Takagi-Sugeno*, Dissertação de Mestrado: UNICAMP, Campinas, 2005.
91. Qiu, W., Liu, X., and Wang, L. Forecasting shanghai composite index based on fuzzy time series and improved C-fuzzy decision trees, *Expert Systems with Applications*, vol. 39, pp. 7680-7689, 2012.
92. Reston Filho, J. C., Affonso, C. M. and Oliveira R. C. L. Energy price prediction multi-step ahead using hybrid model in the Brazilian market, *Electric Power Systems Research*, vol. 117, pp. 115-122, 2014.
93. Rubinstein, S., Goor, A., Rotshtein, A. Time series forecasting of crude oil consumption using neuro-fuzzy inference, *Journal of Industrial and Intelligent Information*, vol. 3, pp. 84-90, 2015.
94. Sadaei, H. J., Enayatifar, R., Abdullah, A. H. and Gani, A. Short-term load forecasting using a hybrid model with a refined exponentially weighted fuzzy time series and an improved harmony search, *Electrical Power and Energy Systems*, vol. 62, pp. 118-129, 2014.
95. Sah, M. and Degtiarev, Y. K. Forecasting Enrollment Model Based on First-Order Fuzzy Time Series, *World Academy of Science, Engineering and Technology, The Board of Trustees The University of Alabama*, vol. 1, pp. 375-378, 2005.
96. Salgado, R. M. *Um Modelo de Previsão de Carga Por Barramento*, Dissertação de Mestrado: UNICAMP, Campinas, 2004.
97. Shafer, G. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.
98. Singh, S. R. A Simple Method of Forecasting Based On Fuzzy Time Series, *Applied Mathematics and Computation* vol. 186, pp. 330-339, 2007b.
99. Singh, S. R. A simple time variant method for fuzzy time series forecasting, *Cybernetics and Systems: An International Journal*, vol. 38, pp. 305-321, 2007c.
100. Singh, S. R. A Computational Method of Forecasting Based On High-Order Fuzzy Time Series, *Expert Systems with Applications*, vol. 36, pp. 10551-10559, 2009.
101. Singh, P. and Borah, B. High-order fuzzy-neuro expert system for time series forecasting, *Knowledge-Based Systems*, vol. 46, pp. 12-21, 2013a.
102. Singh, P. and Borah, B. An efficient time series forecasting model based on fuzzy time series, *Engineering Applications of Artificial Intelligence*, vol. 26, pp. 2443-2457, 2013b.
103. Singh, P. and Borah, B. Forecasting stock index price based on M-factors fuzzy time series and particle swarm optimization, *International Journal of Approximate Reasoning*, vol. 55, pp. 812-833, 2014.
104. Sivasamy, R. and Ama, N. O. Mixed Average-Based Fuzzy Time Series Models for Forecasting Future Civilian Fatalities by Terrorist Attacks in South Asia, *International Journal of Physics and Mathematical Sciences*, vol. 4, pp. 20-25, 2014.
105. Song, Q. A Note On Fuzzy Time Series Model Selection With Sample Autocorrelation Functions, *Cybernetics and Systems*, vol. 34, pp. 93-107, 2003.
106. Song, Q. and Chissom, B. S. Fuzzy Time Series and Its Models, *Fuzzy Sets and Systems*, vol. 54, pp. 269-277, 1993a.
107. Song, Q. and Chissom, B. S. Forecasting Enrollments with Fuzzy Time Series - Part C, *Fuzzy Sets and Systems*, vol. 54, pp. 1-10, 1993b.
108. Song, Q. and Chissom, B. S. Forecasting Enrollments with Fuzzy Time Series - Part C, *Fuzzy Sets and Systems*, vol. 62, pp. 1-8, 1994.
109. Song, Q., Leland, R. P. and Chissom, B. S. Fuzzy Stochastic Fuzzy Time Series and Its Models, *Fuzzy Sets and Systems*, vol. 88, pp. 333-341, 1997.
110. Sugeno M. *Theory of Fuzzy Integrals and Its Applications*. Tese de Doutorado, Instituto de Tecnologia de Tokio, Tokio, Japão, 1974.
111. Sullivan, J., and Woodall, W. H. A Comparison of Fuzzy Forecasting and Markov Modeling, *Fuzzy Sets and Systems*, vol. 64, no. 3, pp. 279-293, 1994.
112. Sun, B. Q., Guo, H., Karimi, H. R., Ge, Y., Xiong, S. Prediction of stock index futures prices based on fuzzy sets and multivariate fuzzy time series, *Neurocomputing*, vol. 151, pp. 1528-1536, 2015.
113. Tanaka, H. and Ishibuchi, H. Possibility Regression Analysis Based On Linear Programming, *Fuzzy Regression Analysis*, in: J. Kacprzyk, M. Fedrizzi (Eds.), *Omnitech Press, Warsaw and Physica-Verlag, Heidelberg*, pp. 47-60, 1992.
114. Tannock, J. D. T. A Fuzzy Control Charting Method for Individuals, *International Journal of Production Research*, vol.41, no. 5, pp. 1017-1032, 2003.
115. Tsaur, R. C., Yang, J. C., Wang, H. F. Fuzzy Relation Analysis in Fuzzy Time Series Model, *Computers & Mathematics*, vol. 49, pp. 539-548, 2005.

116. Tiao, G. C. and Tsay, R. S. Consistency Properties of Least Squares Estimates of Autoregressive Parameters in Arma Models. *The Annals of Statistics*, vol. 11, no. 3, pp. 856-871, 1983.
117. Tsay, R. S. and Tiao, G. C. Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models. *Journal of the American Statistical Association*, vol. 79, no. 385, pp. 84-96, 1984.
118. Tsay, R. S. and Tiao, G. C. Use of Canonical Analysis in Time Series Model Identification. *Biometrika*, vol. 72, pp. 299-315, 1985.
119. Tseng, F. M., Tzeng, G. H., Yu, H. C. and Yuan, B. J. C. Fuzzy ARIMA Model for Forecasting the Foreign Exchange Market, *Fuzzy Sets and Systems* vol. 118, pp. 9-19, 2001.
120. Tseng, F. M and Tzeng, G. H. A Fuzzy Seasonal ARIMA Model for Forecasting, *Fuzzy Sets and Systems*, vol.126, pp. 367-376, 2002.
121. Uri, N. D. Forecasting Peak System Load Using a Combined Time Series and Econometric Model, *Applied Energy*, vol. 4, pp. 219-227, 1978.
122. Uri, N. D and Flanagan, S. P. Short-Term Forecasting of Crude Petroleum and Natural Gas Production, *Applied Energy*, vol. 5, pp. 297-310, 1979.
123. Uslu, V. R., Bas, E., Yolcu, U. and Egrioglu, E. A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations, *Swarm and Evolutionary Computation* vol. 15, pp. 19-26, 2014.
124. Vargens, J. M., Vellasco, M. M. B. R., Tanscheit, R. Previsão de Produção Agrícola Baseada Em Regras Linguísticas e Lógica Fuzzy, *Revista Controle & Automação*, vol. 14, no.2, pp. 114-120, 2003.
125. Wang, L., Zeng, Y., and Chen, T. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting, *Expert Systems with Applications* vol. 42, pp. 855-863, 2015.
126. Wang, G. and Zhang Y. *The Theory of Fuzzy Stochastic Processes*, *Fuzzy Sets and Systems*, vol. 51, pp. 161-178, 1992.
127. Wang, Li-Xin. *A Course in Fuzzy Systems and Control*, 1997.
128. Wang, L. R. and Rowlands, H. An Approach of Fuzzy Logic Evaluation and Control in SPC, *Quality and Reliability Engineering International*, vol.16, no. 2, pp. 91-98, 2000.
129. Wang, L., Liu, X. and Pedrycz, W. Effective Intervals Determined by Information Granules to Improve Forecasting in Fuzzy Time Series, *Expert Systems with Applications*, vol. 40, pp. 5673-5679, 2013.
130. Wang, L., Liu, X., Pedrycz, W. and Shao, Y. Determination of temporal information granules to improve forecasting in fuzzy time series, *Expert Systems with Applications*, vol. 41, pp. 3134-3142, 2014.
131. Wei, W. W. S. *Time Series Analysis: Univariate and Multivariate Methods*. Boston, Pearson Education, vol. 2, 2005.
132. Whittle, P. Prediction and Regulation by Linear Least Squares Methods, *Journal of the Royal Statistical Society, Serie A*, vol. 132, pp. 320-332, 1969.
133. Wong, W. K. Bai, E. and Chu, A. W. C. Adaptive time variant models for fuzzy time series forecasting, *IEEE Transaction on Systems, Man and Cybernetics-Part B: Cybernetics*, vol. 40, no. 6, pp. 1531-1542, 2010.
134. Wu, H. C. Statistical Confidence Intervals for Fuzzy Data, *Expert Systems with Applications*, vol. 36, pp. 2670-2676, 2009.
135. Yan-Hua, Y. and Li-Xia, S. On Fuzzy Time Series Method, *IEEE 3rd International Symposium on Knowledge Acquisition and Modeling*, TBD Wuhan, China, pp. 297-300, 2010.
136. Yu, H. K. Weighted Fuzzy Time Series Models for TAIEX Forecasting. *Physica A*, vol. 349, pp. 609-624, 2005.
137. Yu, T. H. K. and Huarng, K. H. A Bivariate Fuzzy Time Series Model to Forecast the TAIEX. *Expert Systems with Applications*, vol. 34, pp. 2945-2952, 2008.
138. Zadeh, L.A. Fuzzy sets, *Information and Control* vol. 8, pp. 338-353, 1965.
139. Zadeh, L.A. The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning I, *Information Science* vol. 8, pp. 199-249, 1975a.
140. Zadeh, L.A. The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning II, *Information Science* vol. 8, pp. 301-357, 1975b.
141. Zadeh, L.A. The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning III, *Information Science*, vol. 9, pp. 43-80, 1975c.
142. Zadeh, L. A. Fuzzy Sets As a Basis for a Theory of Possibility, *Fuzzy Sets and Systems*, vol. 1, no. 1, pp. 3-28, 1978.