

Orlando Donato Rocha Filho

**Modelagem Baseada em Agrupamento
Nebuloso Evolutivo de Máxima Verossimilhança
Aplicada a Sistemas Dinâmicos Operando em
Ambiente Não-Estacionário**

São Luís – MA

2017

Orlando Donato Rocha Filho

**Modelagem Baseada em Agrupamento Nebuloso
Evolutivo de Máxima Verossimilhança Aplicada a
Sistemas Dinâmicos Operando em Ambiente
Não-Estacionário**

Tese de Doutorado submetido à Coordenação do Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão, como parte dos requisitos necessários à obtenção do grau de Doutor em Engenharia Elétrica na Área de Automação e Controle.

Universidade Federal do Maranhão

Centro de Ciências Exatas e Tecnologia

Programa de Pós-Graduação em Engenharia da Eletricidade

Orientador: Prof. Dr. Ginalber Luiz de Oliveira Serra

São Luís – MA

2017

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Núcleo Integrado de Bibliotecas/UFMA

Rocha Filho, Orlando Donato.

Modelagem Baseada em Agrupamento Nebuloso Evolutivo de
Máxima Verossimilhança Aplicada a Sistemas Dinâmicos
Operando em Ambiente Não-Estacionário / Orlando Donato
Rocha Filho. - 2017.

204 f.

Orientador(a): Ginalber Luiz de Oliveira Serra.

Tese (Doutorado) - Programa de Pós-graduação em
Engenharia de Eletricidade/ccet, Universidade Federal do
Maranhão, São Luís, 2017.

1. Estimacão Paramétrica Recursiva. 2. Máxima
Verossimilhança. 3. Sistemas Nebulosos Evolutivos. 4.
Variável Instrumental Nebulosa. I. Oliveira Serra,
Ginalber Luiz de. II. Título.

Orlando Donato Rocha Filho

**Modelagem Baseada em Agrupamento Nebuloso
Evolutivo de Máxima Verossimilhança Aplicada a
Sistemas Dinâmicos Operando em Ambiente
Não-Estacionário**

Tese de Doutorado submetido à Coordenação do Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão, como parte dos requisitos necessários à obtenção do grau de Doutor em Engenharia Elétrica na Área de Automação e Controle.

Trabalho aprovado. São Luís – MA,

**Prof. Dr. Ginalber Luiz de Oliveira
Serra**
Orientador

Allan Kardec Duailibe Barros
Professor Dr, (UFMA)

Francisco das Chagas Souza
Professor Dr, (UFMA)

Celso José Munaro
Professor Dr, (UFES)

Daniel Furtado Leite
Professor Dr, (UFLA)

*Este trabalho é dedicado às crianças adultas que,
quando pequenas, sonharam em se tornar cientistas.*

AGRADECIMENTOS

A DEUS, por Tudo realizado na minha Vida.

Ao meu filho Marcello e minha esposa Maria das Mercês, pela coragem e incentivo nesta nova jornada.

Ao meu Pai Donato (in memorian) e minha Mãe Ana, meus primeiros grandes mestres.

Ao Prof. Dr. Ginalber Luiz de Oliveira Serra, por acreditar no meu potencial e me orientar nesta tese.

Ao Prof. Dr. João Viana da Fonseca Neto e ao Prof. Dr. Ewaldo Eder Carvalho Santana pelos ensinamentos e compartilhamento de conhecimento.

À Profa. Ma. Danúbia Soares Pires e ao Prof. Dr. Washington Luís Santos Silva pela grande ajuda em todos os momentos de dificuldades e alegrias no desenvolvimento deste Trabalho.

A PPGEE - UFMA por proporcionar a realização deste Trabalho.

Ao Departamento de Eletroeletrônica do IFMA, por acreditar novamente, em especial ao Prof. Me. Marcellino Muniz (in memorian) grande professor e mestre incentivador do Departamento de Eletroeletrônica.

Aos meus alunos do Curso de Engenharia Elétrica Industrial do IFMA, em especial, à turma 2013.1, por todo carinho e pensamento positivo existente.

*“O caráter de um homem não se mede pelas vezes que cai,
mas sim pela elegância com que se levanta”.*
(Charles Chaplin)

RESUMO

Nesta tese é apresentada uma proposta de modelagem baseada máxima verossimilhança aplicada a sistemas dinâmicos operando em ambiente não-estacionário que utiliza a estimação paramétrica recursiva baseada no método de variável instrumental nebulosa, inserido no contexto evolutivo, no sentido de garantir robustez para estimação dos parâmetros diante de dados experimentais corrompidos por ruído. A metodologia é composta por um algoritmo de agrupamento nebuloso evolutivo baseado na similaridade dos dados que emprega uma norma de distância adaptativa baseada no critério de máxima verossimilhança que utiliza uma estratégia de busca adaptativa no experimento para evitar o problema da maldição de dimensionalidade relacionada ao número de regras criadas durante o agrupamento do conjunto de dados. Os resultados computacionais e experimentais para exemplificação da metodologia proposta são: análise estatística da variável instrumental nebulosa inserida no contexto evolutivo; na modelagem caixa preta de uma planta térmica (processo térmico); identificação de um sistema não-linear amplamente divulgado na literatura e a modelagem caixa preta de um helicóptero com dois graus de liberdade que ilustra o desempenho e a eficiência operando ambiente não-estacionário.

Palavras-chave: Sistemas Nebulosos Evolutivos, Variável Instrumental Nebulosa, Estimação Paramétrica Recursiva, Máxima Verossimilhança.

ABSTRACT

This thesis presents a maximum likelihood based modeling approach applied to dynamic systems operating in non-stationary environment that uses recursive parametric estimation based on the method of fuzzy instrumental variable. The context is evolving and the idea is to guarantee a robust for estimation of the parameters of noise-corrupted experimental data. The methodology consists of an evolving fuzzy clustering algorithm based on the similarity of the data which employs an adaptive distance norm based on the maximum likelihood criterion that use an adaptive search strategy on the experiment in order to avoid the *curse of dimensionality* related to the number of rules created during data clustering of the data set. The computational and experimental results to exemplify the proposed methodology are: statistical analysis of the fuzzy instrumental variable inserted in the evolving context; black box modeling of a thermal plant; identification of a benchmark nonlinear system widely published in the literature and the black box modeling of a 2DOF helicopter. These examples are used to illustrate the performance and efficiency by operating in a non-stationary environment.

Keywords: Evolving Fuzzy Systems, Fuzzy Instrumental Variable, Recursive Parameter Estimation, Maximum Likelihood.

LISTA DE ILUSTRAÇÕES

Figura 1 – Diagrama de bloco das diferentes faixas e tipos de modelagem matemática teórica e experimental compreendida pelas modelagens de sistemas dinâmicos reais: caixa branca, caixa cinza e caixa preta (Identificação de Sistemas).	26
Figura 2 – Diagrama geral de sistemas nebulosos evolutivos como uma combinação das características de sistemas nebulosos e de sistemas evolutivos. . . .	28
Figura 3 – O conjunto de dados de flores de íris que é composto por 50 amostras de cada uma das três espécies de íris (íris setosa, íris virgínia e íris versicolor). . . .	30
Figura 4 – O conjunto de dados de flores de íris agrupado de forma eficiente a partir de um agrupamento evolutivo.	30
Figura 5 – Diferentes formas de agrupamento no espaço dimensional \mathbb{R}^2 . O símbolo “•” indica o centro dos grupos após a realização do agrupamento. . . .	37
Figura 6 – Dados Artificiais Cinco Grupos	38
Figura 7 – Dados Artificiais Função Não-Linear	39
Figura 8 – Particionamento possibilístico para cinco grupos com suas respectivas funções de pertinências.	42
Figura 9 – Particionamento probabilístico para cinco grupos com suas respectivas funções de pertinências.	43
Figura 10 – Agrupamento Evolutivo	47
Figura 11 – O Mecanismo de atualização do agrupamento nebuloso evolutivo à medida que os dados de entrada e saída são obtidos dinamicamente. Os parâmetros iniciais do modelo nebuloso evolutivo pode ser extraído de uma modelagem inicial em batelada, de forma <i>online</i> das primeiras amostras do conjunto de dados ou determinados pelo especialista. . . .	48
Figura 12 – Exemplo de um breve processo de agrupamento evolutivo com 3 grupos identificados a partir dos dados no espaço \mathbb{R}^2 . O processo inicia-se com apenas uma amostra e à medida que novas amostras são lidas, as mesmas podem ser incluídas em um grupo existente ou determinar a criação de novos grupos. A amostra $\mathbf{x}^{(1)}$ permite ao agrupamento evolutivo criar um novo grupo C_1^0	50
Figura 13 – Continuação do processo iniciado na Figura 12. A amostra $\mathbf{x}^{(2)}$ atualiza o grupo $C_1^0 \rightarrow C_1^1$, a amostra $\mathbf{x}^{(3)}$ cria um novo grupo C_2^0 e a amostra $\mathbf{x}^{(4)}$ é inserida no grupo mais próximo e a estrutura não é alterada. . . .	50

Figura 14 – A amostra $\mathbf{x}^{(5)}$ atualiza o grupo mais próximo $C_1^1 \rightarrow C_1^2$; a amostra $\mathbf{x}^{(6)}$ é inserida no grupo mais próximo e a estrutura não é alterada; a amostra $\mathbf{x}^{(7)}$ atualiza o grupo $C_2^0 \rightarrow C_2^1$; e a amostra $\mathbf{x}^{(8)}$ cria um novo grupo C_3^0	51
Figura 15 – O processo de agrupamento evolutivo modifica a estrutura a cada nova amostra. A amostra $\mathbf{x}^{(9)}$ atualiza o grupo mais próximo $C_1^2 \rightarrow C_1^3$. . .	51
Figura 16 – Agrupamento de um Conjunto de Dados obtidos da função não linear. Inicialmente tem-se dois grupos selecionados obtidos a partir de um agrupamento nebuloso em batelada.	52
Figura 17 – O terceiro grupo é identificado e criado. Os parâmetros dos agrupamentos existentes também são ajustados em função das novas amostras lidas.	52
Figura 18 – O quarto grupo é identificado e criado. Os parâmetros dos agrupamentos existentes também são ajustados em função das novas amostras lidas. .	53
Figura 19 – O quarto grupo é identificado e criado. Os parâmetros dos agrupamentos existentes também são ajustados em função das novas amostras lidas. .	53
Figura 20 – Mapeamento do Antecedente o Consequente	56
Figura 21 – hipercubo unitário no qual todos os dados são normalizados dentro da faixa de valores de $[0, 1]$	57
Figura 22 – Variável instrumental nebulosa inserida no contexto de máxima verossimilhança aplicada em um sistema nebuloso evolutivo, como uma combinação das metodologias apresentadas em (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015) com o modelo de variável instrumental nebulosa apresentado em (SERRA; BOTTURA, 2007; SERRA, 2012).	67
Figura 23 – Diagrama de bloco da variável instrumental nebulosa inserida no contexto evolutivo com seus principais componentes: etapa de pré-filtragem, modelo nebuloso auxiliar, pesos nebulosos estimação paramétrica do consequente do modelo auxiliar nebuloso.	68
Figura 24 – Agrupamento nebuloso evolutivo de máxima verossimilhança compreendido pela etapa de estimação inicial e estimação evolutiva.	80
Figura 25 – Estimação inicial dos parâmetros do modelo nebuloso baseado no algoritmo de <i>Gustafson-Kessel</i> compreendido pelas etapas de estimação paramétrica do antecedente e estimação paramétrica do consequente (BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007).	82
Figura 26 – Estimação evolutiva: etapa de ajustes da estimação paramétrica evolutiva do antecedente e a estimação paramétrica evolutiva do consequente no instante k	85

Figura 27 – Um breve processo de agrupamento nebuloso evolutivo, representado com amostras em um espaço 2-D. Após a estimativa inicial do agrupamento em batelada, os centros $C_{(1)}^{(0)}$ e $C_{(2)}^{(0)}$ definidos para os grupos 1 e grupos 2, respectivamente; os raios $r_{(1)}^{(0)}$ e $r_{(2)}^{(0)}$ definidos para os grupos 1 e grupo 2, respectivamente. Obtidos a partir da etapa em batelada denominado de instante inicial (0).	92
Figura 28 – Após a leitura da primeira amostra, um novo grupo é criado, $C_{(3)}^{(1)}$, e os demais grupos são atualizados: $C_{(1)}^{(0)} \rightarrow C_{(1)}^{(1)}$ e $C_{(2)}^{(0)} \rightarrow C_{(2)}^{(1)}$. Os parâmetros dos grupos são atualizados: $V_{(i)}^{(0)} \rightarrow V_{(i)}^{(1)}$, $U^{(0)} \rightarrow U^{(1)}$ e $A_{\Sigma(i)}^{(0)} \rightarrow A_{\Sigma(i)}^{(1)}$, para $i = 1, 2, 3$	92
Figura 29 – Após a leitura da $(k - n)$ ésima amostra, os parâmetros do grupos são atualizados: $V_{(i)}^{(k-n-1)} \rightarrow V_{(i)}^{(k-n)}$, $U_{(i)}^{(k-n-1)} \rightarrow U_{(i)}^{(k-n)}$ e $A_{\Sigma(i)}^{(k-n-1)} \rightarrow A_{\Sigma(i)}^{(k-n)}$, para $i = 1, 2, 3$	93
Figura 30 – Após a leitura da $(k - m)$ ésima amostra, um novo grupo é criado. Os parâmetros dos grupos são atualizados: $V_{(i)}^{(k-m-1)} \rightarrow V_{(i)}^{(k-m)}$, $U_{(i)}^{(k-m-1)} \rightarrow U_{(i)}^{(k-m)}$ e $A_{\Sigma(i)}^{(k-m-1)} \rightarrow A_{\Sigma(i)}^{(k-m)}$, para $i = 1, 2, 3, 4$	93
Figura 31 – Após a leitura da (k) ésima amostra, os parâmetros dos grupos são atualizados: $V_{(i)}^{(k-1)} \rightarrow V_{(i)}^{(k)}$, $U_{(i)}^{(k-1)} \rightarrow U_{(i)}^{(k)}$ e $A_{\Sigma(i)}^{(k-1)} \rightarrow A_{\Sigma(i)}^{(k)}$, para $i = 1, 2, 3, 4$	94
Figura 32 – Conjunto de dois grupos com seus respectivos vetores de protótipos e a região de similaridade determinada pela linha tracejada.	98
Figura 33 – Raio do agrupamento determinado para um grau de pertinência limiar alto $u_h \gg 0,5$	98
Figura 34 – Raio do agrupamento determinado para um grau de pertinência limiar baixo $u_h \ll 0,5$	99
Figura 35 – Raio do agrupamento determinado para um grau de pertinência limiar na faixa de $0,4 \leq u_h \leq 0,6$	99
Figura 36 – Grupos compatíveis para atuação do mecanismo de junção de grupos no espaço \mathbb{R}^2	102
Figura 37 – Resultado final do agrupamento de junção entre grupos compatíveis no espaço \mathbb{R}^2	102
Figura 38 – Uma amostra é lida e verifica-se se a condição do fator de credibilidade é satisfeita.	107
Figura 39 – A condição do fator de credibilidade não é satisfeita e a amostra lida é incorporada ao grupo mais próximo.	107
Figura 40 – A condição do fator de credibilidade é satisfeita e um novo grupo é criado a partir da amostra lida.	108
Figura 41 – Diagrama de bloco da análise dos parâmetros do experimento baseado nos métodos dos mínimos quadrados e variável instrumental.	116

Figura 42 – Média da variância $\bar{\sigma}_e^2$ utilizada na equação (5.12h) para cada realização de simulação computacional.	120
Figura 43 – Média da relação de sinal/ruído \overline{SNR} observada para cada realização de simulação computacional.	120
Figura 44 – Índice de correlação da variável instrumental nebulosa com a variável do vetor de regressores. Avaliado pelos índices de correlação médios de <i>Kendall</i> (0,89), <i>Pearson</i> (0,99) e <i>Spearman</i> (0,98); com os respectivos valores de índices médios obtidos a partir das 100 realizações.	121
Figura 45 – Índice de correlação da variável instrumental nebulosa com o erro inserido na variável de saída. Avaliado pelos índices de correlação médios de <i>Kendall</i> (0,03), <i>Pearson</i> (0,04) e <i>Spearman</i> (0,04); com os respectivos valores de índices médios obtidos a partir das 100 realizações.	121
Figura 46 – Resultado da estimação dos parâmetros \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos dos mínimos quadrados não–recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não–recursiva (Batelada) (BFIV) em função do número de realizações.	123
Figura 47 – Resultado da estimação dos parâmetros \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos dos mínimos quadrados não–recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não–recursiva (Batelada) (BFIV) em função do número de realizações.	124
Figura 48 – Resultado da estimação dos parâmetros \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos dos mínimos quadrados não–recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não–recursiva (Batelada) (BFIV) em função do número de realizações.	125
Figura 49 – Resultado da estimação dos parâmetros \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos dos mínimos quadrados não–recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não–recursiva (Batelada) (BFIV) em função do número de realizações.	126
Figura 50 – Resultado da estimação dos parâmetros \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.	127

Figura 51 – Resultado da estimação dos parâmetros \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.	128
Figura 52 – Resultado da estimação dos parâmetros \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.	129
Figura 53 – Resultado da estimação dos parâmetros \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.	130
Figura 54 – Resultado da estimação dos parâmetros \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.	131
Figura 55 – Resultado da estimação dos parâmetros \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.	132
Figura 56 – Resultado da estimação dos parâmetros \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.	133
Figura 57 – Resultado da estimação dos parâmetros \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.	134
Figura 58 – Histograma de parâmetros estimados \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).	135
Figura 59 – Histograma de parâmetros estimados \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).	136

Figura 60 – Histograma de parâmetros estimados \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).	137
Figura 61 – Histograma de parâmetros estimados \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).	138
Figura 62 – Histograma de parâmetros estimados \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).	139
Figura 63 – Histograma de parâmetros estimados \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).	140
Figura 64 – Histograma de parâmetros estimados \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).	141
Figura 65 – Histograma de parâmetros estimados \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).	142
Figura 66 – Processo térmico baseado em uma torradeira elétrica. Processo em tempo real com sinal de entrada (tensão de entrada) e um sinal de saída (temperatura do processo térmico).	143
Figura 67 – Sinais de entrada $u^{(k)}$ e saída $y^{(k)}$ do processo térmico para 1000 amostras.	145
Figura 68 – Sinal de entrada $u^{(k)}$ versus sinal saída $y^{(k)}$ do processo térmico no espaço \mathbb{R}^2 .	145
Figura 69 – Sinal de saída utilizado para realizar a identificação do modelo.	146
Figura 70 – Grupos finais obtidos após a última iteração baseado no fator de aprendizado com taxa de aprendizado fixa, conforme (GEORGIEVA; FILEV, 2009; ANGELOV; FILEV; KASABOV, 2010).	146
Figura 71 – Grupos finais obtidos após a última iteração baseado no fator de aprendizado participativo utilizado nesta tese.	147
Figura 72 – Número de regras evolutivas em função das amostras.	147
Figura 73 – Identificação do Modelo Não-Linear. O eixo da ordenada representa a variação de amplitude das saídas reais e estimadas no processo de identificação, $y^{(k)}$ e $\hat{y}^{(k)}$, respectivamente.	149
Figura 74 – Evolução do número de regras em função do número de amostras lidas.	151
Figura 75 – Evolução do número de amostras/grupo em função do número de amostras lidas.	151

Figura 76 – Coordenadas dos Centros	152
Figura 77 – Evolução dos parâmetros do consequente para as regras das $Regras_{(i)}$ para $i = 1, 2, \dots, 6$	153
Figura 78 – Validação do modelo para o sistema não linear.	155
Figura 79 – Gráfico comparativo do modelo proposto obtido a partir do cálculo da probabilidade <i>a priori</i> conforme proposto em (BABUSKA, 1998).	156
Figura 80 – Helicóptero 2-DOF	157
Figura 81 – Diagrama de corpo livre do Helicóptero 2DOF.	158
Figura 82 – Estrutura utilizada para gerar dados de identificação do modelo.	158
Figura 83 – Estrutura utilizada para gerar dados de validação do modelo.	159
Figura 84 – Diagrama de bloco da análise de desempenho do experimento realizado no Helicóptero 2DOF.	160
Figura 85 – Sinal de entrada para tensão de controle do ângulo de azimute (a), tensão de controle do ângulo de elevação (b).	161
Figura 86 – Sinal de entrada para ângulo de azimute (a), ângulo de elevação (b).	161
Figura 87 – Média móvel do sinal de entrada para tensão de controle do ângulo de azimute em (a) e tensão de controle do ângulo de elevação em (b).	162
Figura 88 – Média móvel do sinal de entrada para ângulo de azimute em (a) e ângulo de elevação em (b).	162
Figura 89 – Variância móvel do sinal de entrada para tensão de controle do ângulo de azimute em (a) e tensão de controle do ângulo de elevação em (b).	163
Figura 90 – Variância móvel do sinal de entrada para ângulo de azimute em (a) e ângulo de elevação em (b).	163
Figura 91 – Diagrama de bloco da análise de desempenho na etapa de validação dos modelos obtidos.	165
Figura 92 – Análise de robustez usada no VAF(%) (a) Ângulo de Azimute e (b) Ângulo de Elevação.	166
Figura 93 – Análise de robustez usada no RMSE (a) Ângulo de Azimute e (b) Ângulo de Elevação.	167
Figura 94 – Análise de robustez usada no NDEI (a) Ângulo de Azimute e (b) Ângulo de Elevação.	168
Figura 95 – Análise de robustez usada no UTHEIL (a) Ângulo de Azimute e (b) Ângulo de Elevação.	169
Figura 96 – Evolução do número de regras do modelo em função do número de amostras.	170
Figura 97 – Convergência dos parâmetros do consequente $\theta_{\varphi_1}^{(i)}$ e $\theta_{\vartheta_1}^{(i)}$ para $regras^{(i)}$, $i =$ $1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	171
Figura 98 – Convergência dos parâmetros do consequente $\theta_{\varphi_2}^{(i)}$ e $\theta_{\vartheta_2}^{(i)}$ para $regras^{(i)}$, $i =$ $1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	172

Figura 99 – Convergência dos parâmetros do consequente $\theta_{\varphi_3}^{(i)}$ e $\theta_{\vartheta_3}^{(i)}$ para $regras^{(i)}$, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	173
Figura 100 – Convergência dos parâmetros do consequente $\theta_{\varphi_4}^{(i)}$ e $\theta_{\vartheta_4}^{(i)}$ para $regras^{(i)}$, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	174
Figura 101 – Convergência dos parâmetros do consequente $\theta_{\varphi_5}^{(i)}$ e $\theta_{\vartheta_5}^{(i)}$ para $regras^{(i)}$, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	175
Figura 102 – Convergência dos parâmetros do consequente $\theta_{\varphi_6}^{(i)}$ e $\theta_{\vartheta_6}^{(i)}$ para $regras^{(i)}$, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	176
Figura 103 – Convergência dos parâmetros do consequente $\theta_{\varphi_7}^{(i)}$ e $\theta_{\vartheta_7}^{(i)}$ para $regras^{(i)}$, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	177
Figura 104 – Convergência dos parâmetros do consequente $\theta_{\varphi_8}^{(i)}$ e $\theta_{\vartheta_8}^{(i)}$ para $regras^{(i)}$, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.	178
Figura 105 – Validação do modelo para o Ângulo de Azimute φ com SNR de 5 dB.	180
Figura 106 – Validação do modelo para o Ângulo de Elevação ϑ com SNR de 5 dB.	180
Figura 107 – Processo estocástico produzido pelo mapeamento das amostras em funções de um parâmetro k	188

LISTA DE TABELAS

Tabela 1 – Força da correlação entre duas séries de dados	115
Tabela 2 – Análise comparativa da média (desvio-padrão) dos parâmetros estimados $\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4$ para os métodos de mínimos quadrados não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV)	122
Tabela 3 – Análise comparativa da média (desvio-padrão) dos parâmetros estimados $\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4$ para os métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).	122
Tabela 4 – Análise Comparativa das Taxas de Aprendizado utilizadas no Algoritmo de Agrupamento Nebuloso Evolutivo	148
Tabela 5 – Análise Comparativa para o Problema da Identificação	155
Tabela 6 – Análise Comparativa para o Problema da Identificação em Função do Cálculo da Probabilidade <i>a Priori</i>	155
Tabela 7 – Teste de Hipótese nula do Teste <i>Dickey-Fuller</i> (MORETTIN, 2006).	164
Tabela 8 – Evolução do número de regras em função do número de amostras.	179
Tabela 9 – Análise Comparativa das Métricas VAF, NDEI, RMSE e UTHEIL com SNR de 5 dB.	181

LISTA DE ABREVIATURAS E SIGLAS

ARX	AutoRegressive with eXogenous onput
BLS	Batch Least Squares
BFIV	Batch Fuzzy Instrumental Variable
IV	Instrumental Variable
FIV	Fuzzy Instrumental Variable
RLS	Recursive Least Squares
RFIV	Recursive Fuzzy Instrumental Variable
NARX	Nonlinear AutoRegressive with eXogenous onput
NDEI	Non-Dimensional Error Index
\mathbb{R}	Conjunto dos Números Reais
RMSE	Root Mean Square Error
VAF	Variance Accounted For

CONTRIBUÇÕES DO TRABALHO

Alguns trabalhos científicos derivados deste texto já foram publicados, sendo um capítulo de livro, 13 artigos em anais de congressos (9 internacionais e 4 nacionais) e três artigos em periódicos. Os trabalhos são apresentados a seguir:

Artigos Publicados em Periódicos

1. Rocha, O and Serra, G. Journal of Control, Automation and Electrical Systems. Adaptive Neuro-Fuzzy Black Box Modeling Based on Instrumental Variable Evolving Algorithm. V. 28, No. 1. pp. 50–67. ISSN 2195–3899. DOI: [10.1007/s40313-016-0285-8](https://doi.org/10.1007/s40313-016-0285-8). (2017). (Qualis B1).
2. Rocha Filho, Orlando Donato e Serra, Ginalber Luiz de Oliveira. Journal of Intelligence and Fuzzy Systems. Evolving Neuro-Fuzzy Network Modeling Approach Based on Recursive Fuzzy Instrumental Variable. V. 32, No. 6. pp. 4159–4172. ISSN 1064–1246 (Impresso). ISSN 1875-8967 (*online*). DOI: [10.3233/IFS-16569](https://doi.org/10.3233/IFS-16569). (2017). (Qualis A2, JCR 1.004).
3. Orlando Donato Rocha Filho and Ginalber Luiz de Oliveira Serra. Fuzzy Set and Systems. Recursive Fuzzy Instrumental Variable Based Evolving Neuro-Fuzzy Identification for Non-Stationary Dynamic System in a Noisy Environment. ISSN 0165-0114. DOI:[10.1016/j.fss.2017.05.016](https://doi.org/10.1016/j.fss.2017.05.016). (2017). (Qualis A1, JCR 2.098).

Capítulo de Livro Publicado

4. Filho, Orlando Donato Rocha; de Oliveira Serra, Ginalber Luiz. Online Evolving Fuzzy Clustering Algorithm Based on Maximum Likelihood Similarity Distance. Lecture Notes in Computer Science. 1ed.: Springer International Publishing, 2014, v. 1, p. 269-280. DOI [10.1007/978-3-319-12027-0-22](https://doi.org/10.1007/978-3-319-12027-0-22). Print ISBN 978-3-319-12026-3. Online ISBN 978-3-319-12027-0.

Artigos Publicados em Congressos, Simpósios e Conferências

5. ROCHA FILHO, O. D.; SERRA, G. L. O. Análise de Variável Instrumental Nebulosa Inserida no Contexto de Modelagem Nebulosa Evolutiva Baseada em Máxima Verossimilhança. In: XIII Simpósio Brasileiro de Automação Inteligente - SBAI 2017. Porto Alegre, RS, Brasil. 2017.

6. ROCHA FILHO, O. D.; SERRA, G. L. O. Evolving Fuzzy Clustering Algorithm based on Maximum Likelihood with Participatory Learning. In: IEEE Conference on Evolving and Adaptive Intelligent Systems (IEEE EAIS 2016), Natal, RN, Brazil. Pages 65–72. DOI:10.1109/EAIS.2016.7502493. IEEE Conference on Evolving and Adaptive Intelligent Systems, 2016.
7. ROCHA FILHO, O. D.; SERRA, G. L. O. Algoritmo Evolutivo para Identificação de Sistemas Dinâmicos Estocásticos Não Lineares Multivariáveis Não Estacionários. In: XXI Congresso Brasileiro de Automática - CBA 2016, Vitória/ES. p. 3458-463, 2016.
8. ROCHA FILHO, O. D.; SERRA, G. L. O. Online Identification Based on Instrumental Variable Evolving Neuro-Fuzzy Model for Stochastic Dynamic Systems. In: IEEE World Congress on Computational Intelligence (IEEE WCCI 2016), Vancouver, Canadá. Pages 9–16. ISBN (online) 978-1-5090-0626-7. Proceedings of the 2016 IEEE World Congress on Computational Intelligence. DOI:10.1109/FUZZ-IEEE.2016.7737661. 2016.
9. ROCHA FILHO, O. D.; SERRA, G. L. O. Evolving Fuzzy Modeling Approach based on Maximum Likelihood with Participatory Learning. In: : 13th International Conference on Informatics in Control, Automation and Robotics (ICINCO), Lisboa. Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics (ICINCO), 2016.
10. Donato Rocha Filho O. and Serra G. Evolving Black Box Recursive Modeling Algorithm. In Proceedings of the 7th International Joint Conference on Computational Intelligence, V. 2: FCTA(IJCCI–2015). Pages 89-96. ISBN: 978-989-758-157-1, SCITER-PRESS, DOI:10.5220/0005582300890096. 2015.
11. ROCHA FILHO, Orlando Donato; de Oliveira Serra, Ginalber Luiz. Algoritmo de Agrupamento Nebuloso Evolutivo de Máxima Verossimilhança Baseado em Variável Instrumental para Identificação de Sistemas Não-Lineares. In: XII Simpósio Brasileiro de Automação Inteligente, SBAI 2015, Natal, Brasil, 2015.
12. ROCHA FILHO, Orlando Donato; de Oliveira Serra, Ginalber Luiz. On Line Fuzzy Learning Maximum Likelihood-Instrumental Variable Evolving Algorithm. IN: 2º Latin-American Congress on Computational Intelligence – LACCI 2015. Pages 1–6. Curitiba, Brasil. DOI: 10.1109/LA-CCI.2015.7435931. 2015.
13. ROCHA FILHO, Orlando Donato; SERRA, G. L. O. Instrumental Variable based Maximum Likelihood Evolving Fuzzy Algorithm for Nonlinear System Identification. In: IEEE 24th - International Symposium on Industrial Electronics, Búzios/RJ. p. 83-88. 2015. DOI: 10.1109/ISIE.2015.7281448. June, 2015.
14. ROCHA FILHO, Orlando Donato; SERRA, G. L. O. MIMO Evolving Learning based on Maximum Likelihood Algorithm Applied to Black Box Fuzzy Modeling for Systems Identification Design. In: ICINCO–2015 - Proceedings of the 12th International

Conference on Informatics in Control, Automation and Robotics, Colmar, Alsace, France, Lisboa: SCITEPRESS Science and Technology Publications, 2015. v. I. p. 199-206. ISBN: 978-989-758-122-9. July, 2015.

15. ROCHA FILHO, O. D.; SERRA, G. L. O. Algoritmo Online de Agrupamento Nebuloso Evolutivo Baseado em Critério de Máxima Verossimilhança. In: XVI Congresso Latinoamericano de Control Automático, CLCA, Cancún, Quintana Roo. XVI Congresso Latinoamericano de Control Automático, CLCA, 2014.

16. ROCHA FILHO, Orlando Donato; de Oliveira Serra, Ginalber Luiz. Evolving maximum likelihood clustering algorithm. In: 2014 IEEE Symposium on Evolving and Autonomous Learning Systems (EALS), 2014, Orlando. 2014 IEEE Symposium on Evolving and Autonomous Learning Systems (EALS). p. 109-115. DOI: 10.1109/EALS.2014.7009511. 2014.

17. ROCHA FILHO, O. D. ; SERRA, G. L. O. Análise Numérica da Predição de Séries Temporais Utilizando Lógica Nebulosa. In: II Congresso Brasileiro de Sistemas Fuzzy, Natal. II Congresso Brasileiro de Sistemas Fuzzy, 2012.

SUMÁRIO

	Trabalhos Publicados	19
1	INTRODUÇÃO	25
1.1	Motivação e Relevância	32
1.2	Justificativa	33
1.3	Contribuição	34
1.4	Organização do Trabalho	34
2	FUNDAMENTAÇÃO TEÓRICA	36
2.1	Agrupamento	36
2.1.1	Partições Nebulosas, Possibilísticas e Rígidas	40
2.1.1.1	Partição Rígida	40
2.1.1.2	Partição Nebulosa	41
2.1.1.3	Partição Possibilística	41
2.1.2	Agrupamento com Matriz de Covariância Nebulosa	43
2.1.2.1	Algoritmo <i>Gustafson–Kessel</i>	43
2.1.2.2	Algoritmo de Máxima Verossimilhança	45
2.1.3	Agrupamento Evolutivo	46
2.1.4	Agrupamento Nebuloso Evolutivo	47
2.2	Modelo Nebuloso Takagi–Sugeno	55
2.2.1	Estrutura Nebulosa <i>Takagi–Sugeno</i>	55
2.2.2	Inferência Nebulosa <i>Takagi–Sugeno</i>	55
2.3	Normalização e Padronização	57
2.3.1	Normalização	57
2.3.2	Padronização	58
3	ESTIMAÇÃO PARAMÉTRICA COM VARIÁVEL INSTRUMENTAL NEBULOSA	59
3.1	Estrutura do Modelo	59
3.2	Estimadores de Mínimos Quadrados	60
3.2.1	Método dos Mínimos Quadrados Não-Recursivo (Batelada)	60
3.2.2	Método dos Mínimos Quadrados Recursivo	61
3.2.3	Problema da Estimação do Método dos Mínimos Quadrados	62
3.2.4	Vetor de Variável Instrumental	63
3.2.5	Escolha dos Instrumentos	65
3.3	Variável Instrumental Nebulosa inserida no Contexto Evolutivo	66

3.3.1	Modelo Nebuloso Auxiliar	68
3.3.2	Modelo de Pré-Filtragem	70
3.3.3	Matriz de Pesos Nebulosos	71
3.3.4	Parâmetros do Consequente do Modelo Nebuloso Auxiliar	72
3.3.4.1	Esquema Não-Recursivo (Batelada)	72
3.3.4.2	Esquema Recursivo	72
3.3.5	Composição da Variável Instrumental Nebulosa	74
3.4	Análise de Convergência da Variável Instrumental Nebulosa	75
3.4.1	Análise de Convergência - Lema 1	76
3.4.2	Análise de Convergência - Lema 2	77
3.4.3	Análise de Convergência - Lema 3	77
3.4.4	Análise de Convergência - Teorema 1	78
4	AGRUPAMENTO NEBULOSO EVOLUTIVO BASEADO EM MÁ-	
	XIMA VEROSSIMILHANÇA	80
4.1	Estimação Inicial	81
4.1.1	Estimação Inicial a partir de Dados em Batelada	81
4.2	Estimação Evolutiva	83
4.2.1	Estimação Paramétrica Evolutiva do Antecedente	83
4.2.2	Estimação Paramétrica Evolutiva do Consequente	84
4.3	Formulação do Algoritmo de Agrupamento Nebuloso em Batelada:	
	estimação inicial	85
4.3.1	Estimação Paramétrica em Batelada do Antecedente	86
4.3.2	Passos do Algoritmo do Antecedente para Gustafson–Kessel	88
4.3.3	Estimação Paramétrica em Batelada do Consequente	88
4.3.4	Passos do Algoritmo do Consequente Gustafson–Kessel	90
4.4	Algoritmo de Agrupamento Nebuloso Recursivo: estimação evolutiva	91
4.4.1	Estimação Paramétrica Evolutiva do Antecedente	91
4.4.1.1	Estimação da Probabilidade a Priori	94
4.4.1.2	Estimação das Distâncias	96
4.4.1.3	Estimação do Grupo Mais Próximo	96
4.4.1.4	Estimação do Raio do Grupo Mais Próximo	96
4.4.1.5	Atualização dos Parâmetros dos Grupos	98
4.4.1.6	Fator de Aprendizado Participativo	100
4.4.1.7	Atualização e Junção de Grupos	101
4.4.1.8	Passos do Algoritmo de Junção dos Grupos	106
4.4.1.9	Fator de Credibilidade	106
4.4.2	Estimação Paramétrica Evolutiva do Consequente	108
4.4.2.1	Modelo Nebuloso Takagi–Sugeno	110
4.4.3	Passos do Algoritmo de Estimação Evolutiva–Recursiva	110

5	RESULTADOS COMPUTACIONAIS E EXPERIMENTAIS	113
5.1	Métricas de Avaliação	113
5.2	Coefficiente de Correlação	114
5.3	Análise da Variável Instrumental Nebulosa	116
5.4	Identificação de um Processo Térmico	143
5.5	Identificação de Sistemas Não Lineares	148
5.6	Identificação de um Helicóptero - 2DOF	156
6	CONCLUSÕES E PROPOSTAS FUTURAS	182
6.1	Conclusões	182
6.2	Propostas Futuras	184
	 APÊNDICES	 186
	APÊNDICE A – PROCESSOS ESTOCÁSTICOS	187
A.1	Definição	187
A.2	Classificação de Processos Estocásticos	188
A.3	Especificação de Processos Estocásticos	188
A.4	Momentos de Processos Estocásticos	189
A.5	Estacionariedade de Processos Estocásticos	190
A.6	Sistema Dinâmico Estocástico	190
A.7	Sistemas Dinâmicos Não-Estacionários	191
	 APÊNDICE B – MÁXIMA VEROSSIMILHANÇA	 193
B.1	Verossimilhança	193
B.2	Métodos e Estimadores de Máxima-Verossimilhança	194
B.3	Variância	194
	 APÊNDICE C – PARÂMETROS INICIAIS DOS ALGORITMOS	 196
	 REFERÊNCIAS	 197

1 INTRODUÇÃO

O comportamento temporal de sistemas dinâmicos das áreas de engenharia, bem como das diversas áreas da biologia, medicina, economia e afins, pode ser descrito por meio de **modelos matemáticos**. Denomina-se **modelagem** o processo de elaboração de um **modelo matemático** que seja adequado a um sistema dinâmico no sentido de representar com certa fidelidade a interação entre as variáveis envolvidas. De acordo com a Teoria de Sistemas, observa-se que há duas abordagens gerais para modelagem, denominadas de: **modelagem teórica** e **modelagem experimental** (JUANG, 1994; LJUNG, 1999; KEESMAN, 2011; ISERMANN; NCHHOF, 2011; WANG; GARNIER, 2012; AGUIRRE, 2015).

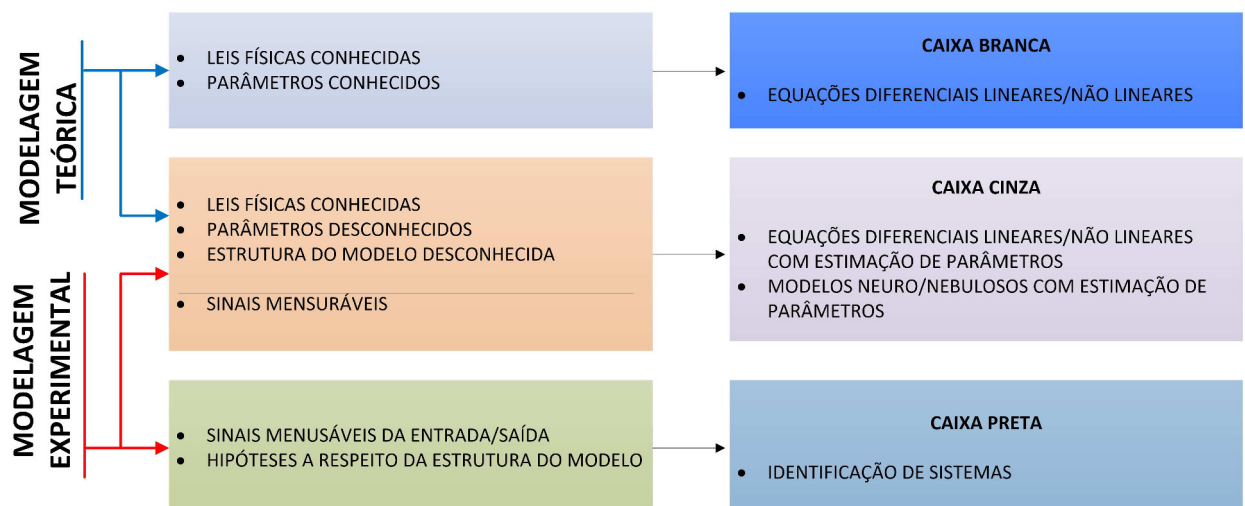
De acordo com (JUANG, 1994; LJUNG, 1999; KEESMAN, 2011; ISERMANN; NCHHOF, 2011; WANG; GARNIER, 2012; AGUIRRE, 2015), a **modelagem teórica**, também conhecida como análise teórica, permite a obtenção de um modelo a partir das equações físicas do estado, equações fenomenológicas e equações de equilíbrio de sistemas dinâmicos. Ao aplicar métodos de modelagem, obtém-se um conjunto de equações matemáticas, que finalmente conduzirá ao modelo teórico final representativo da estrutura e parâmetros do sistema dinâmico. Na **modelagem experimental**, também conhecida como análise experimental, obtém-se um modelo matemático que é derivado das medidas realizadas no sistema dinâmico, isto é, a partir dos sinais de entrada, estados e sinais de saída deste.

Na figura Figura 1 é ilustrado o diagrama de bloco das diferentes faixas e tipos de modelagem matemática teórica e experimental compreendida pelas três abordagens para a modelagem de sistemas dinâmicos reais são descritas como segue:

- i) **Modelagem Caixa–Branca:** neste tipo de modelagem, requer-se que o especialista conheça profundamente e esteja bem familiarizado com o sistema dinâmico em questão. Além disso, faz-se necessário conhecer bem as relações matemáticas fundamentais que descrevem os fenômenos envolvidos, tendo em vista, que a imprecisão nos valores dos parâmetros e a complexidade do modelo matemático final, torna a modelagem caixa–branca uma tarefa complicada. Esta modelagem é também conhecida como modelagem pela física ou natureza do processo ou modelagem fenomenológica; por essa razão, o conhecimento prévio e entendimento completo do comportamento físico do sistema dinâmico é bastante complexo em aplicações reais (BAHREMAND, 2015; CAPONETTO et al., 2014; LIN, 2013).

- ii) **Modelagem Caixa–Preta:** nem sempre é viável o equacionamento do processo físico para que seja realizado a modelagem caixa–branca. Neste caso surge a necessidade de uma modelagem que não necessite conhecimento prévio ou requeira pouco conhecimento sobre os princípios fundamentais que regem o comportamento do sistema. Esta área é denominada de modelagem caixa–preta ou modelagem empírica e desenvolve técnicas e algoritmos para obter (identificar) modelos matemáticos de sistemas dinâmicos a partir dos dados gerados pelo próprio sistema. O principal problema desta modelagem está na escolha adequada da estrutura para o modelo, de forma a obter o máximo de informação sobre o comportamento do sistema dinâmico a ser modelado (MAGERL; CEPERIC; BARIC, 2015; BOSCAINO et al., 2015; HASSANEIN et al., 2015).
- iii) **Modelagem Caixa–Cinza:** modelagem híbrida em que a interpretação física pode ser utilizada dependendo do conhecimento prévio do sistema, além de se utilizar o conjunto de dados para compor o modelo matemático. Esta modelagem híbrida combina as vantagens de ambos os modelos, caixa–branca e caixa–preta (DADHE et al., 2001; MANGOUROVA; RINGWOOD, 2006; THORDARSON; DANMARKS, 2012).

Figura 1: Diagrama de bloco das diferentes faixas e tipos de modelagem matemática teórica e experimental compreendida pelas modelagens de sistemas dinâmicos reais: caixa branca, caixa cinza e caixa preta (Identificação de Sistemas).



Fonte: Adaptado de (ISERMANN; NCHHOF, 2011).

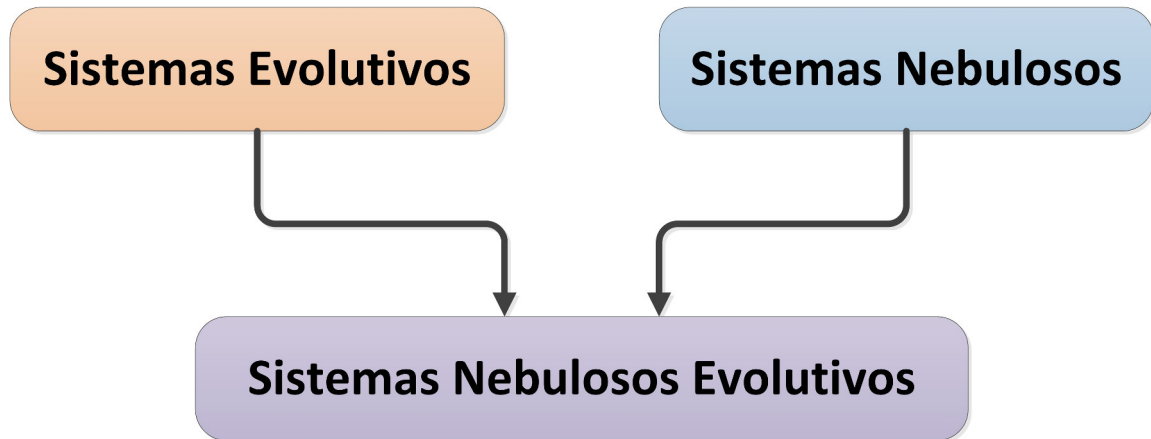
A modelagem experimental tem encontrado cada vez mais atenção ao longo dos últimos anos por se tratar de uma modelagem menos complexa e sem necessidade de conhecimento das leis físicas envolvidas no modelo, principalmente se os parâmetros derivados do modelo não forem suficientemente precisos.

A identificação de sistemas é a determinação experimental do comportamento temporal de um sistema dinâmico a partir de medições dos sinais de entrada e saída do modelo dinâmico. Opcionalmente, estes sinais de medição podem estar inseridos em um ambiente cujas medidas podem estar corrompidas por um ruído autocorrelacionado. Portanto, procedimentos e técnicas são necessários para que se possa identificar o modelo diante do ambiente com ruído para uma posterior validação do modelo matemático obtido (JUANG, 1994; LJUNG, 1999; KEESMAN, 2011; ISERMANN; NCHHOF, 2011; WANG; GARNIER, 2012; AGUIRRE, 2015).

No final do século XX, os avanços relacionados às arquiteturas computacionais em nível de *hardware* e *software* permitiram a captura e análise de modelos orientados a dados em diversas áreas da Engenharia. Áreas como manufatura industrial, controle de processos, predição de séries temporais, processamento de sinais e identificação de sistemas buscam novas metodologias que sejam capazes de se adaptar à dinâmica evolutiva dos seus respectivos conjuntos de dados (KEESMAN, 2011; ISERMANN; NCHHOF, 2011; WANG; GARNIER, 2012; AGUIRRE, 2015).

De acordo com (LUGHOFER, 2011; PRATAMA; ANAVATTI; LUGHOFER, 2013; LUGHOFER et al., 2015), os modelos orientados a dados se adaptam automaticamente à dinâmica evolutiva do sistema quando equipados com algoritmos de aprendizado incrementais. Estes algoritmos são capazes de alterar a estrutura do modelo à medida que uma nova amostra do conjunto de dados é lida. Inserido no conceito de sistemas evolutivos, há os sistemas nebulosos evolutivos que combinam as melhores características dos sistemas nebulosos com os sistemas evolutivos, conforme ilustrado na Figura 2 (ANGELOV, 2013; COSTA; ANGELOV; GUEDES, 2015).

Figura 2: Diagrama geral de sistemas nebulosos evolutivos como uma combinação das características de sistemas nebulosos e de sistemas evolutivos.



Fonte: Adaptado de (LUGHOFFER, 2011).

Inicialmente, a tarefa de modelagem era realizada a partir de um sistema nebuloso baseado no conhecimento adquirido do especialista a partir de modelos nebulosos linguísticos, ou seja, sistemas de inferência nebuloso *Mamdani*, como forma de capturar o conhecimento qualitativo em um determinado sistema dinâmico, inserido em conjunto de regras linguísticas do tipo **SE** < *antecedente* > **ENTÃO** < *consequente* > com suas respectivas proposições nebulosas (WANG, 1997; LI, 2006; ASSAWINCHAICHOTE; NGUANG; SHI, 2007; PEDRYCZ; GOMIDE, 2007). Na década de 90, foram realizados estudos de modelos nebulosos baseados em dados, ou seja, modelos nebulosos funcionais do tipo *Takagi-Sugeno*, no qual as informações do próprio modelo eram extraídas a partir de um conjunto de dados, também inseridos em um conjunto de regras linguísticas do tipo **SE** < *antecedente* > **ENTÃO** < *consequente* >, porém no consequente é utilizado um funcional das variáveis linguísticas do antecedente ao invés de um conjunto nebuloso; conseqüentemente, as informações do especialista eram utilizadas de forma complementar durante o processo de modelagem a partir da inferência nebulosa *Takagi-Sugeno* (BABUSKA, 1998; LI, 2006; PEDRYCZ; GOMIDE, 2007; DU; ZHANG, 2008; BABUSKA B. SCHUTTER; GUERRA, 2010; GUO et al., 2011; JEEVA; PONNUSAMY; SANGAPILLAI, 2013; BENZAOUIA; HAJJAJI, 2014).

Uma abordagem eficaz para identificação de sistemas, no contexto nebuloso funcional, inicia-se a partir do particionamento do conjunto de dados, isto é, o conjunto de dados é subdividido em grupos menores determinados a partir de uma similaridade existente entre uma dada amostra e seu respectivo grupo, realizado a partir de um algoritmo de agrupamento nebuloso em batelada. Este particionamento realizado em batelada em um

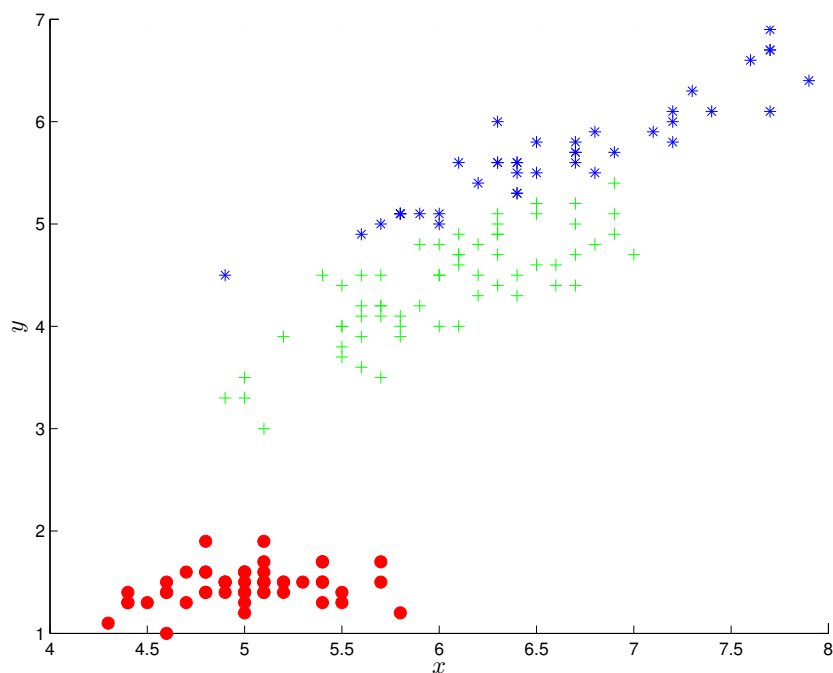
conjunto de dados que já está disponível antes de ser efetuado qualquer tipo de análise, necessita de um conhecido prévio a respeito do número de grupos utilizados no particionamento e o total de grupos existentes definem o total de regras nebulosas do modelo de *Takagi–Sugeno*. Os algoritmos de agrupamentos nebulosos em batelada necessitam de uma informação prévia do número de grupos, que é fornecido pelo especialista, além do mais, o total de regras nebulosas são mantidas constantes durante o processo de modelagem de um sistema dinâmico operando em ambiente não-estacionário (BABUSKA, 1998; PEDRYCZ; GOMIDE, 2007; MAAN; AHMAD, 2008; ANGELOV, 2013; BARUAH; ANGELOV, 2014).

Outro desafio na área de identificação de sistemas está no contexto evolutivo, que consiste em desenvolver modelos inteligentes que sejam capazes de alterar sua estrutura e ajustar seus parâmetros internos à medida que as amostras são lidas de um conjunto de dados. Nesta topologia, o número de grupos existentes e os conjuntos nebulosos das variáveis linguísticas do antecedente são extraídos a partir de um agrupamento realizado no conjunto de dados, de modo que, as amostras existentes serão agrupadas conforme a similaridade existente entre elas (KASABOV; FILEV, 2006; KASABOV, 2007; WATTS, 2009; PEDRYCZ; GOMIDE, 2007; MAAN; AHMAD, 2008; ANGELOV, 2013; BARUAH; ANGELOV, 2014).

Na Figura 3 é ilustrado um conjunto de dados de flores de íris que é composto por 50 amostras de cada uma das três espécies de íris (íris setosa, íris virgínia e íris versicolor), cujas as dimensões x e y são: comprimento sepal e comprimento da pétala, respectivamente. A identificação de sistemas que se baseia em um agrupamento eficiente no contexto evolutivo deve ser capaz de extrair destes conjuntos de dados o total de grupos distintos existentes. Na Figura 4 é ilustrado o desempenho de um agrupamento evolutivo realizado neste conjunto de dados e se observou que este foi capaz de identificar e agrupar três grupos de flores de íris com características distintas.

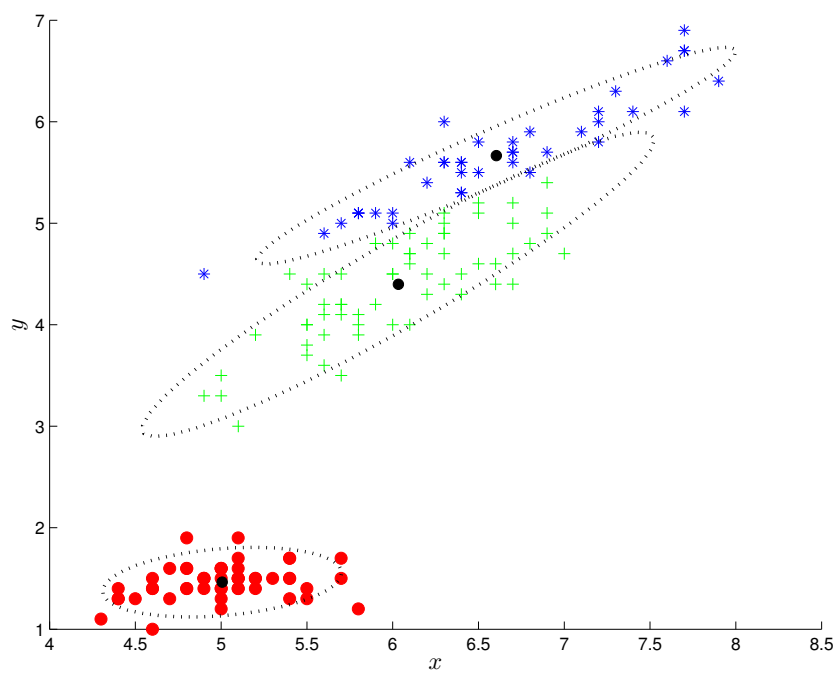
Para alcançar tal objetivo, com os estudos no início do século XXI, as primeiras pesquisas voltadas para agrupamentos nebulosos evolutivos baseavam-se no critério da distância Euclidiana, e estes algoritmos empregavam métricas de distância não adaptativas e, conseqüentemente, por vezes, tinham seu agrupamento nebuloso comprometido (KASABOV; FILEV, 2006; KASABOV, 2007; WATTS, 2009). Deste então, surgiram pesquisas de agrupamentos nebulosos evolutivos que se baseavam em métricas de distância adaptativas, tais como as metodologias baseadas na distância de *Mahalanobis*. Esta métrica de distância permite detectar diferentes formas e orientações dos grupos de dados similares (ANGELOV; FILEV; KASABOV, 2010; GEORGIEVA; FILEV, 2009; ANGELOV, 2013; BARUAH; ANGELOV, 2014). Atualmente pesquisas vem sendo desenvolvidas a fim de aprimorar a capacidade de alterar a estrutura e parâmetros de modelos nebulosos baseada em algoritmos de aprendizado incrementais (CAMINHAS; LEMOS; GOMIDE, 2013; ANGELOV, 2013; BORDIGNON; GOMIDE, 2014; BENZAOUIA; HAJJAJI, 2014). Também, citam-se

Figura 3: O conjunto de dados de flores de íris que é composto por 50 amostras de cada uma das três espécies de íris (íris setosa, íris virgínia e íris versicolor).



Fonte: Elaborado pelo autor.

Figura 4: O conjunto de dados de flores de íris agrupado de forma eficiente a partir de um agrupamento evolutivo.



Fonte: Elaborado pelo autor.

estruturas evolutivas como: FLEXFIS (*Flexible Fuzzy Inference System*), detalhado em (LUGHOFER, 2008); SOFMLS (*Self-Organizing Fuzzy Modified Least-Squares Network*) em (RUBIO, 2009); eFuMO (*Evolving Fuzzy Model*) em (DOVZAN; LOGAR; SKRJANC, 2015); eGNN (*Evolving Granular Neural Networks*) e FBeM (*Fuzzy Set Based Evolving Modeling*) estruturas granulares evolutivas detalhadas em (LEITE, 2012; LEITE; COSTA; GOMIDE, 2013).

Um problema característico de sistemas nebulosos evolutivos é o crescimento exponencial do número de regras nebulosas em função do aumento da dimensão do seu espaço de entrada e saída. Este é um dos principais empecilhos à aplicação de sistemas nebuloso evolutivos, considerando que o número de regras nebulosas podem crescer à medida que as amostras são lidas, e é conhecido na literatura como *maldição da dimensionalidade* (do inglês *curse of dimensionality*). Em (KASABOV; FILEV, 2006; KASABOV, 2007; WATTS, 2009; ANGELOV; FILEV; KASABOV, 2010; GEORGIEVA; FILEV, 2009) verifica-se várias estratégias para a minimização deste problema, porém, há um campo aberto para pesquisa de estratégias de junção de regras similares, de modo que, o algoritmo do sistema nebuloso evolutivo seja capaz de buscar e agrupar grupos de elementos de dados baseando-se na similaridade entre eles, conseqüente o problema da *maldição da dimensionalidade* será minimizado.

A identificação de sistemas dinâmicos usando modelos nebulosos evolutivos inseridos na inferência de *Takagi-Sugeno* está dividida basicamente em duas etapas principais: identificação da estrutura evolutiva e estimação paramétrica recursiva. A primeira etapa consiste na definição do número de regras nebulosas de acordo com o número de partições das variáveis linguísticas de entrada definidas no antecedente. A topologia é extraída a partir do agrupamento realizado no conjunto de dados *online* à medida que as amostras são lidas. A segunda etapa consiste na estimação dos parâmetros do antecedente e do conseqüente em função das regras nebulosas identificadas pelo algoritmo nebuloso evolutivo, onde os parâmetros do conseqüente são os coeficientes das expressões função das variáveis linguísticas do antecedente (SERRA, 2012). Quando aplicado em problemas reais de engenharia, a estimação paramétrica dos coeficientes do conseqüente deve ser eficiente diante de dados experimentais corrompidos por ruído (BABUSKA, 1998; BABUSKA B. SCHUTTER; GUERRA, 2010; SERRA, 2012). Os estimadores paramétricos do tipo mínimos quadrados podem produzir excelentes resultados mediante dados corrompidos por um ruído branco. Caso o ruído ou erro na equação seja autocorrelacionado, o estimador de mínimos quadrados e os demais estimadores podem ser usados, mas são dependentes da precisão do modelo de ruído (LJUNG, 1999; ISERMANN; NCHHOF, 2011; WANG; GARNIER, 2012; AGUIRRE, 2015).

Uma estimação paramétrica consistente e robusta sem a necessidade da modelagem do ruído, com excelentes propriedades de estimação, pode ser utilizada a partir do uso da

variável instrumental (VI) e os métodos de estimação de parâmetros com base na variável instrumental têm uma longa história nos problemas de engenharia e são amplamente discutidos em (SODERSTROM; STOICA, 1983; YOUNG, 1984; SODERSTROM; STOICA, 1989; YOUNG, 2011; ISERMANN; NCHHOF, 2011; AGUIRRE, 2015).

Em (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015), é apresentado a abordagem de uma variável instrumental ótima refinada estendida no contexto da máxima verossimilhança. Esta metodologia proporcionou uma estimação paramétrica recursiva dos parâmetros do consequente em um ambiente cujos dados podem estar corrompidos por ruído, sem a necessidade de conhecimento prévio do modelo de ruído adotado.

Em (SERRA; BOTTURA, 2007; SERRA, 2012), tem-se o conceito da variável instrumental nebulosa, com aplicações ao processo de identificação, recursivo ou não-recursivo, associada a uma estrutura do tipo ARX nebulosa *Takagi-Sugeno*. Esta metodologia é estatisticamente não correlacionada com o ruído do sistema, a fim de minimizar o problema da estimação paramétrica do consequente em um ambiente ruidoso. No que se segue, a abordagem da variável instrumental nebulosa que emprega as metodologias propostas em (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015), inserida num contexto evolutivo, permitirão uma estimação paramétrica recursiva consistente e robusta diante de dados corrompidos por ruído operando em ambiente não-estacionário.

1.1 Motivação e Relevância

O desenvolvimento de técnicas de identificação de sistemas que sejam capazes de lidar com sistemas dinâmicos em ambientes cujos dados podem ser corrompidos por ruído, motivam projetistas a desenvolverem metodologias que sejam capazes de adaptar modelos para acompanhar o fluxo de dados de carácter não-estacionário contendo aleatoriedade.

Modelos que se baseiam em métricas de distância adaptativa, devem ser capazes de acompanhar a dinâmica das diferentes formas e orientações dos grupos. Neste contexto, observa-se que o princípio da máxima verossimilhança demonstra ser um interessante procedimento usado para se trabalhar com distâncias adaptativas diante de conjunto de dados com características não-estacionária (GEORGIEVA; FILEV, 2009; SKRJANC; DOVZAN; GOMIDE, 2014; SILVA et al., 2014; SUN et al., 2014; PETELIN; KOCIJAN, 2014; BARUAH; ANGELOV, 2014). Também, verifica-se que a máxima verossimilhança demonstra ser um interessante procedimento, devido ao fato da matriz de agrupamento nebuloso deste algoritmo utilizar uma métrica exponencial e não está sujeita as condições restritivas de volume dos grupos (BABUSKA, 1998; ABONYI et al., 2005).

Algoritmos de agrupamento nebuloso em batelada desempenham um papel importante na modelagem de dados experimentais. Contudo, estes algoritmos requerem uma condição inicial fornecida pelo especialista, isto é, o número de grupos deve ser

conhecido (BABUSKA, 1998; BABUSKA B. SCHUTTER; GUERRA, 2010). As diversas áreas da engenharia, como manufatura, controle, predição, processamento de sinais e identificação, buscam novas metodologias para modelagem computacional que se baseiam em agrupamentos de dados (DOVZAN; LOGAR; SKRJANC, 2015). Um problema importante em agrupamento de dados é a estimativa do número de grupos (*clusters*) a partir de um conjunto de dados oriundos de um sistema dinâmico operando em um ambiente não-estacionário. Vários trabalhos lidam com sistemas dinâmicos adaptativos e evolutivos para detecção de grupos com volume variável, formas incertas, tamanhos desiguais e densidades diferentes (ANGELOV, 2013; DOVZAN; LOGAR; SKRJANC, 2015). Atualmente, pesquisas sobre agrupamentos evolutivos inteligentes, especialmente sobre sistemas nebulosos, tem-se dado para melhorar a capacidade de alterar a estrutura e os parâmetros de modelos a medida que novos dados são mensurados (LIU; ZHANG; ZHANG, 2013; ANGELOV, 2013; COSTA; ANGELOV; GUEDES, 2015; LUGHOFFER et al., 2015; SKRJANC, 2015; DOVZAN; LOGAR; SKRJANC, 2015).

Existem vários trabalhos que propõem modelos funcionais nebulosos evolutivos para lidar com problemas de identificação de sistemas, previsão de séries temporais, detecção e diagnóstico adaptativo de falhas que utilizam métodos de estimação paramétrica recursiva. Porém, a pesquisa sobre sistemas evolutivos que se baseiam em metodologias que garantem estimação paramétrica robusta diante de ambiente com ruído é uma área de estudo em aberto (LIMA et al., 2010; ANGELOV; FILEV; KASABOV, 2010; LEITE et al., 2011; MACIEL; GOMIDE; BALLINI, 2012; PRATAMA; ANAVATTI; LUGHOFFER, 2013; CAMINHAS; LEMOS; GOMIDE, 2013; ANGELOV, 2013; BARUAH; ANGELOV, 2014).

1.2 Justificativa

O uso de agrupamento nebuloso evolutivo com fator de aprendizagem participativo se aplica a sistemas dinâmicos cujo fluxo de dados apresentam características não-estacionárias, conseqüentemente a dinâmica do agrupamento deve ser capaz de adaptar os parâmetros e sua estruturas à medida que uma nova informação surge no fluxo de dados. Tal metodologia aplica-se a identificação de sistemas dinâmicos operando em um ambiente não-estacionário (PRATAMA; ANAVATTI; LUGHOFFER, 2013; CAMINHAS; LEMOS; GOMIDE, 2013; MACIEL; GOMIDE; BALLINI, 2012; ANGELOV, 2013; BARUAH; ANGELOV, 2014).

Assim, há a necessidade de desenvolver algoritmos para modelagem de sistemas dinâmicos operando em ambiente não-estacionário, que sejam capazes de alterar a estrutura do modelo e seus parâmetros, à medida que os dados são obtidos, sem a necessidade de fornecer uma condição inicial por parte do especialista (supervisor) e permita, dinamicamente, detectar as diferentes formas e orientações dos diversos conjuntos de dados.

Portanto, uma metodologia de identificação baseada em agrupamento evolutivo baseado em máxima verossimilhança com fator de aprendizagem participativo, inserido em uma estrutura de regras nebulosas do tipo *Takagi–Sugeno*, com estimação paramétrica recursiva baseada em variável instrumental, para identificação de sistemas dinâmicos na presença de dado não-estacionários é de importância fundamental.

1.3 Contribuição

As contribuições da metodologia proposta são:

- i) Desenvolver e analisar um algoritmo de agrupamento evolutivo participativo baseado na similaridade dos dados de acordo com a métrica de máxima verossimilhança;
- ii) Apresentar uma estratégia de busca adaptativa ao experimento no sentido de evitar o problema da *madição da dimensionalidade* relacionada ao número de regras criadas;
- iii) Utilizar a estimação paramétrica recursiva baseada no método de variável instrumental nebulosa, inserido no contexto evolutivo, no sentido de garantir robustez para estimação dos parâmetros diante de dados experimentais corrompidos por ruído. Isto significa que não é necessária informação detalhada sobre as características estatísticas do ruído;
- iv) Apresentar uma abordagem para aprendizagem recursiva do consequente que incorpora variáveis instrumentais nebulosas pré-filtradas no projeto de estimadores recursivos de mínimos quadrados;
- v) Apresentar uma nova metodologia para cálculo da probabilidade *a priori* das distâncias segundo a métrica de máxima verossimilhança.

1.4 Organização do Trabalho

O presente trabalho de tese de doutorado objetiva apresentar uma modelagem baseada em agrupamento nebuloso evolutivo de máxima verossimilhança aplicada a sistemas dinâmicos operando em ambiente não-estacionário. O capítulo 2 são apresentados conceitos preliminares de agrupamento nebuloso evolutivo; o modelo nebuloso funcional do tipo Takagi–Sugeno e suas principais características. No capítulo 3 apresenta os conceitos envolvidos na estimação paramétrica com variável instrumental nebulosa inserida no contexto evolutivo baseada em máxima verossimilhança, além da análise de convergência da variável instrumental nebulosa. O formalismo matemático e o pseudocódigo da metodologia proposta são discutidos no capítulo 4. Resultados computacionais e experimentais para ilustrar a eficiência e aplicabilidade da metodologia proposta tem como base: análise

estatística da variável instrumental nebulosa inserida no contexto evolutivo, na modelagem caixa preta de uma planta térmica (processo térmico), identificação de um sistema não-linear amplamente citado na literatura, e modelagem do tipo caixa preta de um helicóptero com dois graus de liberdade com dados não-estacionários. Estes são discutidos no Capítulo 5. No capítulo 6 são apresentadas as conclusões e propostas futuras de trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, serão apresentados os principais conceitos gerais e noções básicas envolvidos na elaboração desta tese. A fim de atingir este objetivo, uma breve introdução aos conceitos de: agrupamento nebuloso; modelo nebuloso Takagi–Sugeno; normalização e padronização, são apresentados.

2.1 Agrupamento

A principal ideia na análise de agrupamento é a distribuição por classes ou grupos de objetos que mantêm entre eles uma determinada similaridade, além de garantir a organização dos dados pertencentes aos seus respectivos grupos. Esta tarefa deve garantir que os dados pertencentes a seus respectivos grupos sejam homogêneos entre si, e heterogêneos aos demais grupos existentes. Os algoritmos de agrupamento permitem revelar as estruturas base (grupos com elementos semelhantes) nos conjuntos de dados, de tal forma que podem ser explorados não somente para classificação e reconhecimento de padrões, como também, utilizados para a redução da complexidade envolvida na modelagem e otimização de sistemas dinâmicos (BABUSKA, 1998; LI, 2006; BABUSKA B. SCHUTTER; GUERRA, 2010; BENZAOUIA; HAJJAJI, 2014; SUN et al., 2014).

As técnicas de agrupamento são aplicadas a conjuntos de dados discretos ou contínuos que assumem valores numéricos, isto é, dados quantitativos; também podem ser aplicadas aos conjuntos de dados que representam características de qualidade (rótulos) associadas a um item específico, isto é, dados qualitativos. Além disso, técnicas de agrupamento podem ser aplicadas a junção de ambos os tipos de dados. Nesta tese, será considerado o agrupamento de dados quantitativos. O conjunto de dados é uma coletânea de observações que devem refletir a dinâmica de um processo físico qualquer. Cada observação deste processo está de acordo com p variáveis de medida, o que define a sua dimensionalidade.

A observação dos dados no instante de tempo k é agrupada em um vetor coluna de dimensionalidade p , dado por:

$$\mathbf{z}^{(k)} = \begin{bmatrix} z_{(1)}^{(k)} & \dots & z_{(p)}^{(k)} \end{bmatrix}^T, \quad \mathbf{z}^{(k)} \in \mathbb{R}^p \quad \forall k. \quad (2.1)$$

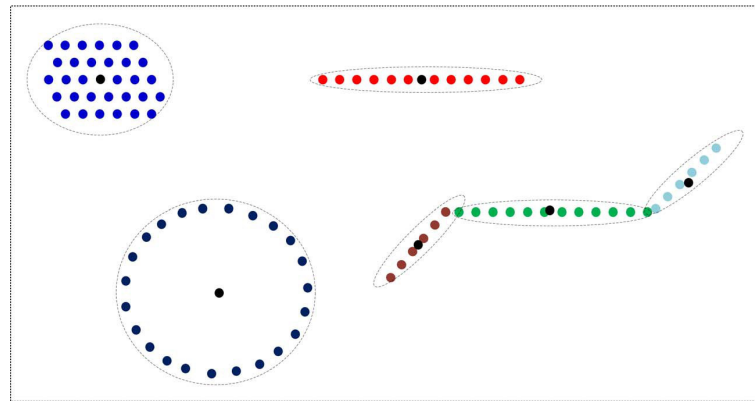
O conjunto dos dados relativos às N observações é denotado por $\mathbf{Z} = \{\mathbf{z}^{(k)} | k = 1, 2, \dots, N\}$,

que representa a matriz característica do processo físico, dada por:

$$\mathbf{Z} = \begin{bmatrix} z_{(1)}^{(1)} & z_{(1)}^{(2)} & \dots & z_{(1)}^{(N)} \\ z_{(2)}^{(1)} & z_{(2)}^{(2)} & \dots & z_{(2)}^{(N)} \\ \vdots & \vdots & \vdots & \vdots \\ z_{(p)}^{(1)} & z_{(p)}^{(2)} & \dots & z_{(p)}^{(N)} \end{bmatrix}. \quad (2.2)$$

O principal objetivo da técnica de agrupamento é a divisão do conjunto de dados em grupos (do inglês *cluster*). Entenda-se grupo como sendo um aglomerado de objetos que são mais semelhantes entre si do que aos outros membros dos demais grupos. A medida de similaridade adotada, neste contexto de agrupamento, é definida como sendo uma métrica de distância, ou seja, os dados são similares se os mesmos estiverem dentro de um certo limiar de distância em relação ao centro de um determinado grupo. Após a tarefa de agrupamento, os grupos podem assumir diferentes formas geométricas, devido ao fato das amostras individuais do conjunto de dados apresentarem características de similaridades que podem diferenciar entre as demais, conseqüentemente as amostras similares são agrupadas entre si, conforme ilustrado na Figura 5 (BABUSKA, 1998; BABUSKA B. SCHUTTER; GUERRA, 2010; BENZAOUIA; HAJJAJI, 2014).

Figura 5: Diferentes formas de agrupamento no espaço dimensional \mathbb{R}^2 . O símbolo “●” indica o centro dos grupos após a realização do agrupamento.



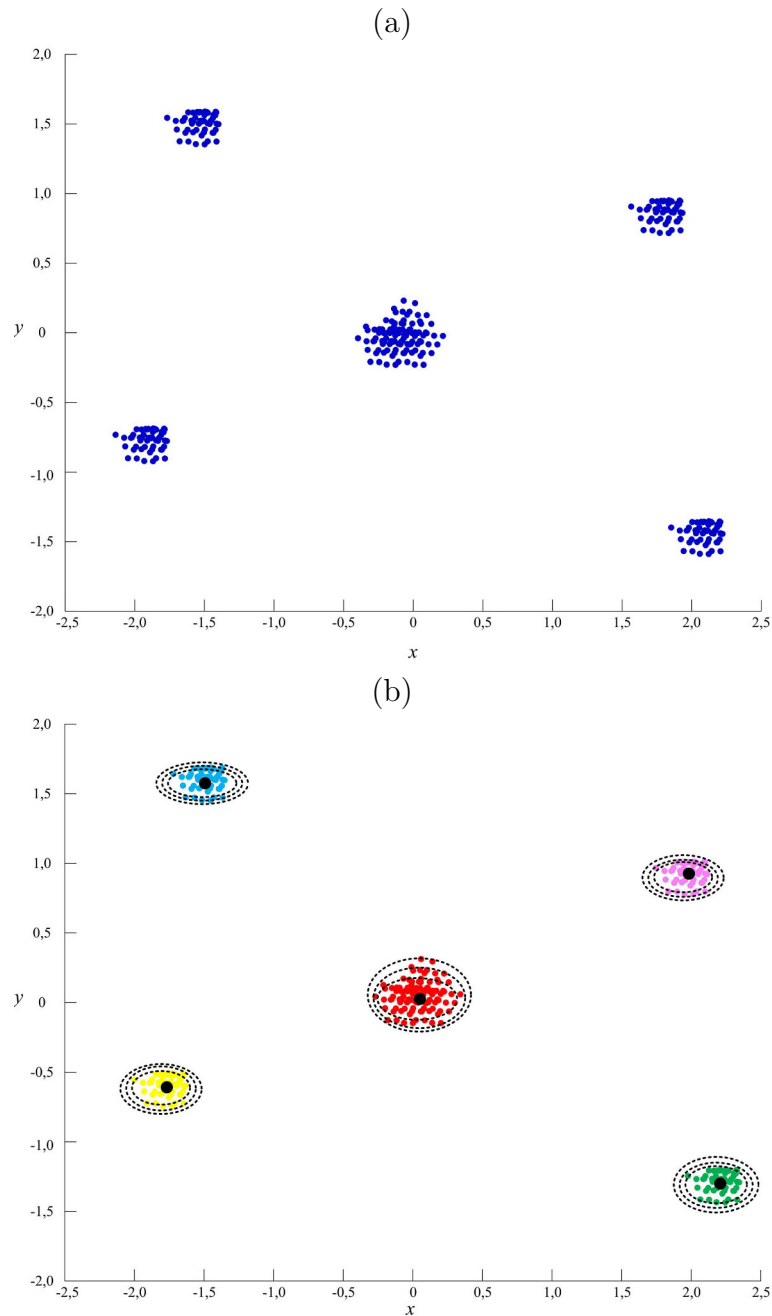
Fonte: Elaborada pelo autor

No sentido de ilustrar a relação existente entre um grupo e suas respectivas amostras, após a realização da etapa de agrupamento, considera-se o exemplo de dois conjuntos de dados artificiais utilizados para exemplificar agrupamento nebuloso, como segue:

- a) **Cinco-grupos** (PAL; SARKAR, 2014): trata-se de um conjunto de dados com um disco interno de 100 amostras, e quatro grupos pequenos de 50 amostras, compondo os quatro lados do conjunto de dados central, conforme mostrado na Figura 6a. A matriz

característica deste conjunto de dados é dada por $\mathbf{Z} = [x, y]$, $-2,0 < x, y < 2,0$. Para um número de grupos definido previamente e após a realização do algoritmo de agrupamento, observa-se o resultado final neste conjunto de dados, conforme a Figura 6b.

Figura 6: Conjunto de dados artificiais denominados de “cinco grupos”. Definiu-se que o algoritmo deveria encontrar 5 grupos distintos. O símbolo “•” indica o centro dos agrupamento. (a) Dados originais, (b) Dados agrupados.



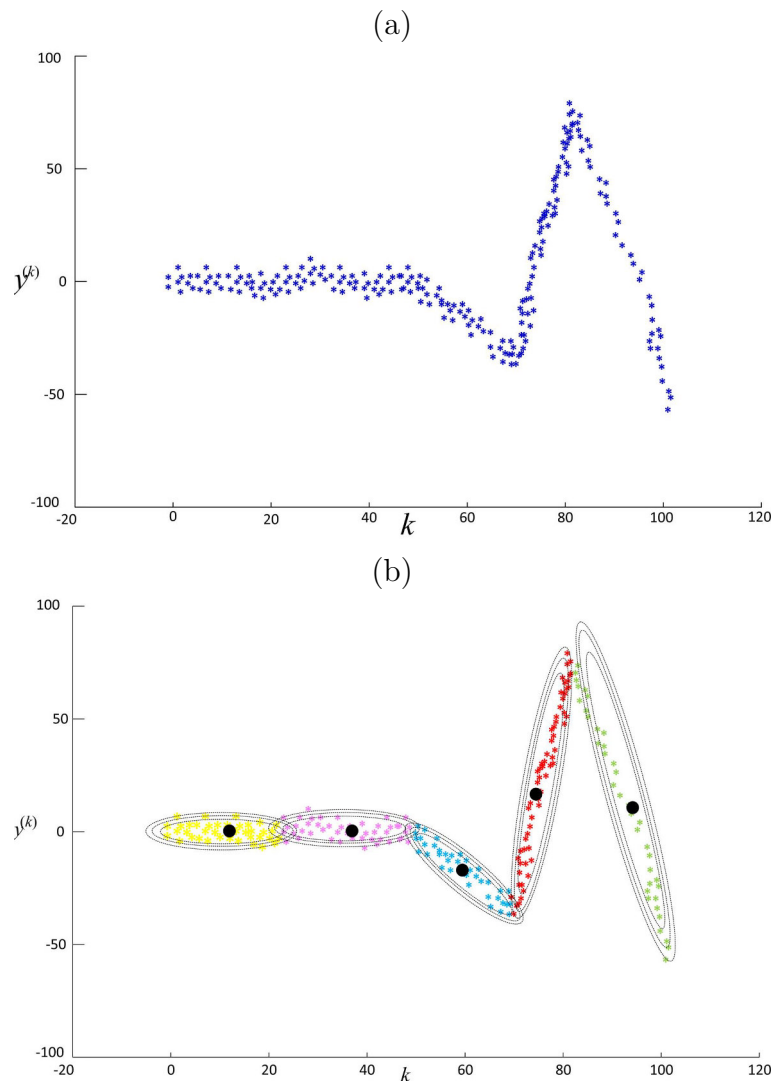
Fonte: Elaborada pelo autor

- b) **Função não linear** (ANGELOV, 2013): trata-se de um conjunto de dados artificiais, descrito pela equação a seguir:

$$y^{(k)} = 0,0001\sin(0,001k^2)k^3 + \varepsilon, \quad k \in [0, 100] \quad (2.3)$$

onde $\varepsilon \in N(0, 25)$ é um ruído aleatório, conforme mostrado na Figura 7a. O conjunto de dados contém 200 amostras e a matriz característica deste conjunto de dados é dada por $\mathbf{Z} = [k, y^{(k)}]$, $0 < k < 100$, $-50 < y < 100$. Também, após a realização do agrupamento, em que o número de grupos foi definido previamente, observa-se na Figura 7b o resultado final de um agrupamento.

Figura 7: Conjunto de dados artificiais de uma função não linear (2.3). Definiu-se que o algoritmo deveria encontrar 5 grupos distintos. O símbolo “•” indica o centro dos agrupamento. (a) Dados originais, (b) Dados agrupados.



Fonte: Elaborada pelo autor

2.1.1 Partições Nebulosas, Possibilísticas e Rígidas

O entendimento do particionamento nebuloso, possibilístico e rígido torna-se necessário para a formulação de problemas voltados para técnicas de identificação de sistemas baseadas em agrupamentos. Os particionamentos nebulosos e possibilísticos são vistos como uma generalização do particionamento rígido que é formulado em termos da teoria clássica de conjuntos (BABUSKA, 1998; ABONYI; FEIL, 2007; PEDRYCZ, 2007; KEESMAN, 2011).

2.1.1.1 Partição Rígida

O principal objetivo de um particionamento é a divisão de um conjunto \mathbf{Z} de dados em c grupos, onde c pode ser considerado como informação *a priori*. Baseado na teoria clássica de conjuntos, uma partição rígida de \mathbf{Z} , pode ser definida como uma família de conjuntos $\{\mathbf{T}_i | i = 1, 2, \dots, c\} \subset \mathbf{Z}$, com as seguintes propriedades:

$$\bigcup_{i=1}^c \mathbf{T}_i = \mathbf{Z} \quad (2.4)$$

para as seguintes restrições,

$$\mathbf{T}_i \cap \mathbf{T}_j = 0, \quad 1 \leq i \neq j \leq c \quad (2.5)$$

e

$$0 \subset \mathbf{T}_i \subset \mathbf{Z}, \quad 1 \leq i \leq c \quad (2.6)$$

A equação (2.4) significa que o conjunto de dados \mathbf{Z} contém todo os subconjuntos $\mathbf{T}_i, i = 1, 2, \dots, c$. Todos os subconjuntos devem ser disjuntos, como se observa na restrição (2.5) e nenhum dos subconjuntos contidos no conjuntos de dados é vazio. Isto é dado pela restrição (2.6).

Em termos de funções de pertinências nebulosas, μ , as equações (2.4), (2.5) e (2.6) podem ser expressas da seguinte forma:

$$\bigvee_{i=1}^c \mu_{\mathbf{T}_i} = 1 \quad (2.7)$$

para

$$\mu_{\mathbf{T}_i} \wedge \mu_{\mathbf{T}_j} = 0, \quad i \leq i \neq j \leq c \quad (2.8)$$

e

$$0 \leq \mu_{\mathbf{T}_i} \leq 1, \quad 1 \leq i \leq c \quad (2.9)$$

Dada uma amostra de observação $\mathbf{z}^{(k)}, k = 1, 2, \dots, N$, denota-se sua pertinência por $\mu_{(i)}(\mathbf{z}^{(k)})$ a ser representado neste trabalho por $\mu_{(i)}^{(k)}$. Dada uma matriz de partição,

com $c \times N$ elementos, a matriz $\tilde{\mathbf{U}}_\mu = [\mu_{(i)}^{(k)}]$, é dita rígida se, e somente se, esta satisfaz as seguintes condições:

$$\mu_{(i)}^{(k)} \in \{0, 1\}, \quad 1 \leq i \leq c, \quad 1 \leq k \leq N \quad (2.10)$$

$$\sum_{i=1}^c \mu_{(i)}^{(k)} = 1, \quad 1 \leq k \leq N \quad (2.11)$$

$$0 < \sum_{k=1}^N \mu_{(i)}^{(k)} < N, \quad 1 \leq i \leq c \quad (2.12)$$

Assim, define-se particionamento rígido, M_{hc} , do conjunto de dados \mathbf{Z} , como segue:

$$M_{hc} = \left\{ \tilde{\mathbf{U}}_\mu \in \mathbb{R}^{c \times N} \mid \mu_{(i)}^{(k)} \in \{0, 1\}, \forall i, k; \sum_{i=1}^c \mu_{(i)}^{(k)} = 1; \forall k; 0 < \sum_{k=1}^N \mu_{(i)}^{(k)} < N, \forall i \right\} \quad (2.13)$$

2.1.1.2 Partição Nebulosa

Dada uma matriz de partição, com $c \times N$ elementos, a matriz $\tilde{\mathbf{U}}_\mu = [\mu_{(i)}^{(k)}]$, é dita nebulosa se, e somente se, esta satisfaz as seguintes condições:

$$\mu_{(i)}^{(k)} \in [0, 1], \quad 1 \leq i \leq c, \quad 1 \leq k \leq N \quad (2.14)$$

$$\sum_{i=1}^c \mu_{(i)}^{(k)} = 1, \quad 1 \leq k \leq N \quad (2.15)$$

$$0 < \sum_{k=1}^N \mu_{(i)}^{(k)} < N, \quad 1 \leq i \leq c \quad (2.16)$$

A equação (2.14) significa que a pertinência é um valor situado no intervalo de 0 a 1 atribuído a uma amostra de observação $\mathbf{z}^{(k)}$, $k = 1, 2, \dots, N$. A condição de normalização dos graus de pertinências em relação aos centros dos grupos é definida na equação (2.15). A soma dos graus de pertinências de todas as observações não deve ser superior ao total de observações, conforme equação (2.16).

Assim, define-se particionamento nebuloso, M_{fc} , do conjunto de dados \mathbf{Z} , como segue:

$$M_{fc} = \left\{ \tilde{\mathbf{U}}_\mu \in \mathbb{R}^{c \times N} \mid \mu_{(i)}^{(k)} \in [0, 1], \forall i, k; \sum_{i=1}^c \mu_{(i)}^{(k)} = 1; \forall k; 0 < \sum_{k=1}^N \mu_{(i)}^{(k)} < N, \forall i \right\} \quad (2.17)$$

2.1.1.3 Partição Possibilística

Dada uma matriz de partição, com $c \times N$ elementos, a matriz $\tilde{\mathbf{U}}_\mu = [\mu_{(i)}^{(k)}]$, é dita possibilística se, e somente se, esta satisfaz as seguintes condições:

$$\mu_{(i)}^{(k)} \in [0, 1], \quad 1 \leq i \leq c, \quad 1 \leq k \leq N \quad (2.18)$$

$$\exists i, \mu_{(i)}^{(k)} > 0, \forall k \quad (2.19)$$

$$0 < \sum_{k=1}^N \mu_i^{(k)} < N, \quad 1 \leq i \leq c \quad (2.20)$$

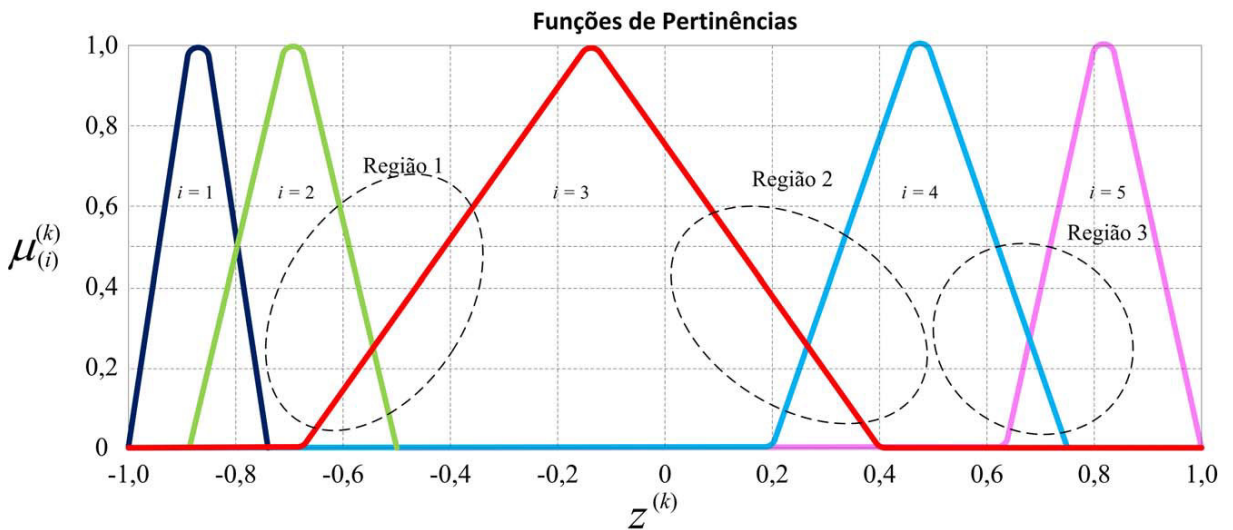
A equação (2.18) significa que a pertinência é um valor situado no intervalo de 0 a 1 atribuído a uma amostra de observação $z^{(k)}, k = 1, 2, \dots, N$. A condição de normalização para os graus de pertinências em relação aos centros dos grupos não é garantida de acordo com a equação (2.19). A soma dos graus de pertinências de todas as observações não deve ser superior ao total de observações, de acordo com a equação (2.20).

Assim, define-se particionamento possibilístico, M_{pc} , do conjunto de dados Z , como segue:

$$M_{pc} = \left\{ \tilde{U}_\mu \in \mathbb{R}^{c \times N} \mid \mu_{(i)}^{(k)} \in [0, 1], \forall i, k; \forall k, \exists i, \mu_{(i)}^{(k)} > 0; 0 < \sum_{k=1}^N \mu_{(i)}^{(k)} < N, \forall i \right\} \quad (2.21)$$

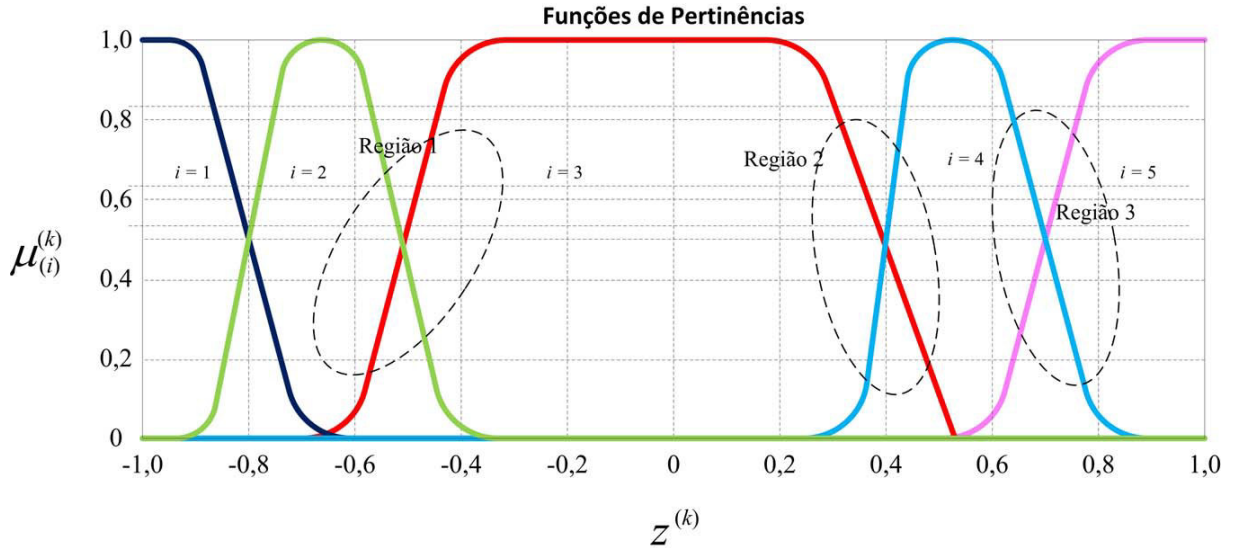
A Figura 8 e Figura 9 são ilustrados um conjunto de dados com o particionamento de cinco grupos com suas respectivas funções de pertinências. Observe o efeito do particionamento possibilístico na região 1, região 2 e região 3 da Figura 8: para uma dada amostra qualquer com seus respectivos graus de pertinências, percebe-se que a soma de todos os graus de pertinências é menor que 1. O efeito do particionamento probabilístico é verificado Figura 9 e observe que a condição de normalização da equação (2.15) é satisfeita em todas as regiões.

Figura 8: Particionamento possibilístico para cinco grupos com suas respectivas funções de pertinências.



Fonte: Elaborada pelo autor

Figura 9: Particionamento probabilístico para cinco grupos com suas respectivas funções de pertinências.



Fonte: Elaborada pelo autor

Nesta tese, adota-se o particionamento nebuloso e dentre os algoritmos de agrupamento nebuloso existentes, utilizar-se-á um algoritmo de distância adaptativa matriz de covariância nebulosa de máxima verossimilhança.

2.1.2 Agrupamento com Matriz de Covariância Nebulosa

Os algoritmos de agrupamento nebuloso do tipo *Gustafson–Kessel* e Máxima Verossimilhança são considerados algoritmos de agrupamento de distância adaptativa e podem ser utilizados em conjuntos de dados independentes de os mesmos estarem normalizados.

2.1.2.1 Algoritmo *Gustafson–Kessel*

De acordo com (BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007), este algoritmo emprega uma norma de distância adaptativa que é capaz de detectar conjunto de dados de diferentes formas geométricas. Cada grupo tem sua própria matriz de norma induzida $\mathbf{A}_{(i)}$, $i = 1, 2, \dots, c$, o qual determina uma norma de produto interno para distância entre uma dada amostra no instante k e os respectivos i ésimos grupos existentes, expresso por

$$D_{(i)k\mathbf{A}_{(i)}}^2 = \left(\mathbf{z}^{(k)} - \mathbf{v}_{(i)}\right)^T \mathbf{A}_{(i)} \left(\mathbf{z}^{(k)} - \mathbf{v}_{(i)}\right), \quad (2.22)$$

onde $\mathbf{z}^{(k)}$ e $\mathbf{v}_{(i)}$ são: a amostra lida no instante k e o vetor de protótipos do grupo i , respectivamente. O vetor de protótipos representa o vetor dos centros dos grupos existentes e matriz de vetor de protótipos é expressa por

$$\mathbf{V} = [\mathbf{v}_{(1)}, \mathbf{v}_{(2)}, \dots, \mathbf{v}_{(c)}], \mathbf{v}_{(i)} \in \mathbb{R}^p. \quad (2.23)$$

A função objetivo do algoritmo de *Gustafson–Kessel* é definida por:

$$j(\mathbf{Z}; \tilde{\mathbf{U}}_0, \mathbf{V}, \mathbf{A}) = \sum_{i=1}^c \sum_{k=1}^N \left(\mu_{(i)}^{(k)} \right)^m D_{(i)k\mathbf{A}_{(i)}}^2, \quad (2.24)$$

onde $\tilde{\mathbf{U}}_0 \in M_{fc}$ é a matriz de partição nebulosa de inicialização, $\mathbf{V} \in \mathbb{R}^{p \times c}$ é o vetor de protótipos e o coeficiente de incerteza $m > 1$. A solução é dada por,

$$(\tilde{\mathbf{U}}_\mu, \mathbf{V}, \mathbf{A}) = \underset{M_{fc} \times \mathbb{R}^{p \times c} \times PDP}{\arg \min} j(\mathbf{Z}; \tilde{\mathbf{U}}_0, \mathbf{V}, \mathbf{A}), \quad (2.25)$$

De acordo com (BABUSKA, 1998; ABONYI et al., 2005), a expressão para a matriz $\mathbf{A}_{(i)}$ é obtida a partir de:

$$\mathbf{A}_{(i)} = [\det(\mathbf{A}_{\Sigma_i})]^{1/p} (\mathbf{A}_{\Sigma_i})^{-1}, \quad (2.26)$$

onde \mathbf{A}_{Σ_i} é matriz de covariância nebulosa para o i ésimo grupo definido por

$$\mathbf{A}_{\Sigma_i} = \frac{\sum_{k=1}^N \left[\left(\mu_{(i)}^{(k)} \right)^m (\mathbf{z}^{(k)} - \mathbf{v}_{(i)})^T (\mathbf{z}^{(k)} - \mathbf{v}_{(i)}) \right]}{\sum_{k=1}^N \left(\mu_{(i)}^{(k)} \right)^m}. \quad (2.27)$$

O vetor de protótipos do i ésimo grupo é dado por

$$\mathbf{v}_{(i)} = \frac{\sum_{k=1}^N \left[\left(\mu_{(i)}^{(k)} \right)^m \mathbf{z}^{(k)} \right]}{\sum_{k=1}^N \left(\mu_{(i)}^{(k)} \right)^m}, 1 \leq i \leq c. \quad (2.28)$$

A estrutura da matriz de covariância dos grupos fornece informações a respeito da forma e orientação do grupo. A proporção dos comprimentos dos eixos hiper-elipsóides do grupo é dada pela razão das raízes quadradas dos autovalores da matriz de covariância. As direções dos eixos hiper-elipsóides é dado pela razão das raízes quadradas dos autovalores da matriz de covariância.

Os elementos da matriz de partição nebulosa são calculados para $1 \leq i \leq c$ e $1 \leq k \leq N$, de acordo com a seguinte expressão

$$\mu_{(i)}^{(k)} = \frac{1}{\sum_{j=1}^c \left(\frac{(\mathbf{z}^{(k)} - \mathbf{v}_{(i)})^T [\det(\mathbf{A}_{\Sigma_i})]^{1/p} (\mathbf{A}_{\Sigma_i})^{-1} (\mathbf{z}^{(k)} - \mathbf{v}_{(i)})}{(\mathbf{z}^{(k)} - \mathbf{v}_{(j)})^T [\det(\mathbf{A}_{\Sigma_j})]^{1/p} (\mathbf{A}_{\Sigma_j})^{-1} (\mathbf{z}^{(k)} - \mathbf{v}_{(j)})} \right)^{2/(m-1)}}. \quad (2.29)$$

2.1.2.2 Algoritmo de Máxima Verossimilhança

De acordo com (BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007), os estimadores de máxima verossimilhança empregam uma norma de distância exponencial expressa por

$$D_{(i)k\Sigma(i)}^2 = \frac{[\det(\Sigma(i))]^{1/2}}{P(i)} \exp^{\frac{1}{2}(\mathbf{z}^{(k)} - \mathbf{v}_{(i)})^T (\Sigma(i))^{(-1)} (\mathbf{z}^{(k)} - \mathbf{v}_{(i)})}, \quad (2.30)$$

De acordo com (GATH; GEVA, 1989; ABONYI et al., 2005; SOLEIMANI-B; C.; BABAK, 2010), a matriz de covariância do i ésimo grupo é dada por:

$$\Sigma(i) = \frac{\sum_{k=1}^N \left[(\mu_{(i)}^{(k)})^{w_g} (\mathbf{z}^{(k)} - \mathbf{v}_{(i)})^T (\mathbf{z}^{(k)} - \mathbf{v}_{(i)}) \right]}{\sum_{k=1}^N (\mu_{(i)}^{(k)})^{w_g}}, \quad (2.31)$$

onde $w_g \in [1, \infty)$ representa o peso das pertinências.

A probabilidade *a priori* de selecionar o i ésimo grupo, é dada por:

$$P(i) = \frac{1}{N} \sum_{k=1}^N \mu_{(i)}^{(k)}. \quad (2.32)$$

O vetor de protótipos do i ésimo grupo é dado por

$$\mathbf{v}_{(i)} = \frac{\sum_{k=1}^N \left[(\mu_{(i)}^{(k)})^{w_g} \mathbf{z}^{(k)} \right]}{\sum_{k=1}^N (\mu_{(i)}^{(k)})^{w_g}}, \quad 1 \leq i \leq c. \quad (2.33)$$

Os elementos da matriz de partição nebulosa são calculados para $1 \leq i \leq c$ e $1 \leq k \leq N$, de acordo com a seguinte expressão

$$\mu_{(i)}^{(k)} = \frac{1}{\sum_{j=1}^c \left(\frac{[\det(\Sigma(i))]^{1/2}}{P(i)} \exp^{\frac{1}{2}(\mathbf{z}^{(k)} - \mathbf{v}_{(i)})^T (\Sigma(i))^{(-1)} (\mathbf{z}^{(k)} - \mathbf{v}_{(i)})} \right)^{2/(m-1)}, \quad 1 \leq i \leq c. \quad (2.34)$$

De acordo com (GATH; GEVA, 1989; BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007; ZHU, 2011), o algoritmo de máxima verossimilhança demonstra ser um interessante algoritmo para agrupamento nebuloso, tendo em vista que sua matriz de agrupamento nebuloso utiliza uma métrica exponencial e não está sujeito as condições restritivas impostas pelo volume dos grupos, portanto, é capaz de detectar

diferentes formas geométricas, densidades e volumes dos grupos de dados. Para uma convergência mais rápida na determinação dos parâmetros do algoritmo, este necessita de uma boa inicialização, de acordo com a literatura atual, uma boa inicialização para este é algoritmo pode ser obtida a partir do algoritmo de agrupamento de *Gustafson-Kessel*.

2.1.3 Agrupamento Evolutivo

De acordo com (LUGHOFER, 2011), os modelos podem ser divididos em quatro classes:

- a) Modelos Analíticos;
- b) Modelos Baseados em Conhecimento;
- c) Modelos Orientados por Dados;
- d) Modelos híbridos que representam uma combinação dos modelos acima citados.

Os modelos analíticos são deduzidos analiticamente a partir de conhecimento específico a respeito das leis físicas, químicas, mecânicas e/ou comportamento do processo biológico inerente aos sistemas dinâmicos. Estes modelos podem ser representados de diferentes formas, tais como equações diferenciais e integrais, ou como equações de estados.

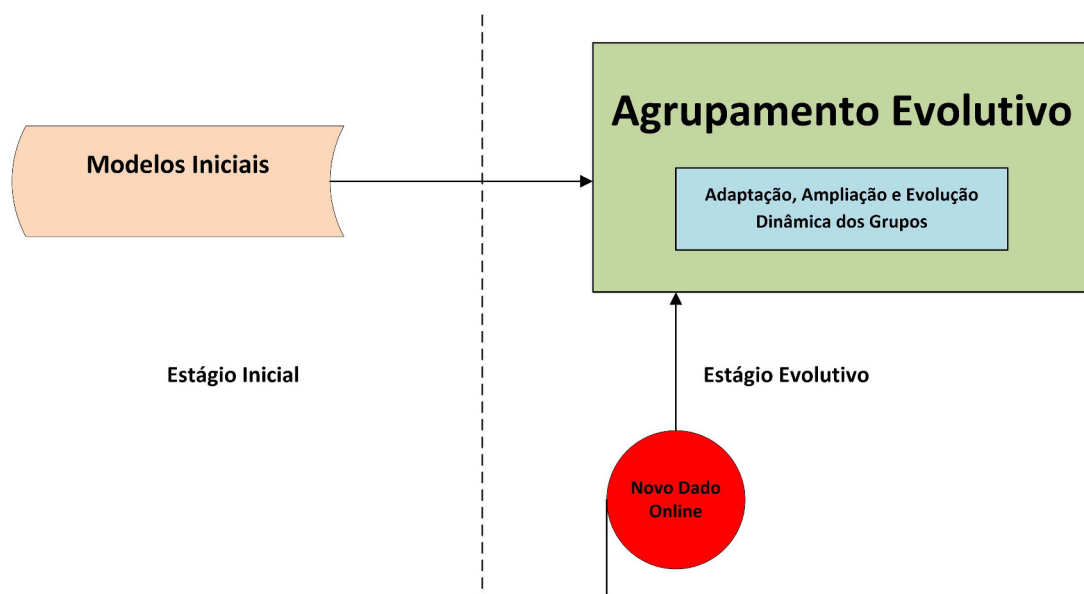
Os modelos baseados em conhecimento são obtidos a partir do conhecimento e experiência que o especialista obteve sobre o sistema dinâmico em longo prazo. Esta experiência de longo prazo resulta em um amplo conhecimento sobre o sistema dinâmico e as condições necessárias que foram utilizadas para sua formulação.

Os modelos orientados por dados, que são adotados nesta tese, entendem-se como qualquer forma de modelos matemático que são totalmente projetados, extraídos ou aprendidos a partir de um conjunto de dados. Contrapondo-se aos modelos baseados em conhecimento, nestes os dados normalmente representam as observações obtidas do sistema dinâmico.

Agrupamentos evolutivos baseiam-se em modelos orientados por dados, que são automaticamente adaptados, ampliados e evoluídos dinamicamente em tempo real com base nas novas amostras. Como tal, os agrupamentos evolutivos são capazes de suportar quaisquer cenários de modelagem de sistemas dinâmicos para dados de fluxo contínuo e medições dinâmicas em tempo real por meio da modificação de sua estrutura evolutiva (PEDRYCZ; GOMIDE, 2007; LUGHOFER, 2011). Portanto, os agrupamento evolutivos são mais adaptativos e flexíveis quando comparados às estruturas tradicionais de agrupamento existentes, haja vista que estes apresentam características de aprendizado *online* e são capazes de permitir a extração de novas características dos dados à medida que estes são adquiridos, ou seja, permitem um aprendizado gradual contínuo, conforme ilustrado

na Figura 10. Os modelos iniciais do agrupamento evolutivo podem ser construídos a partir de um modelo inicial denominado de modelo em batelada e realizará o treinamento inicial destes. Caso não utilize modelos iniciais em batelada, a modelagem evolutiva pode ser inicializada do zero, ou seja, a modelagem será aplicada às primeiras amostras do conjunto de dados *online* (PEDRYCZ; GOMIDE, 2007), (LUGHOFER, 2011; SKRJANC; DOVZAN; GOMIDE, 2014).

Figura 10: Mecanismo de atualização do agrupamento evolutivo à medida que um novo dado é lido. O agrupamento evolutivo pode ser inicializado a partir de um modelo em batelada ou inicializado a partir das primeiras amostras *online*.



Fonte: Adaptado de (LUGHOFER, 2011)

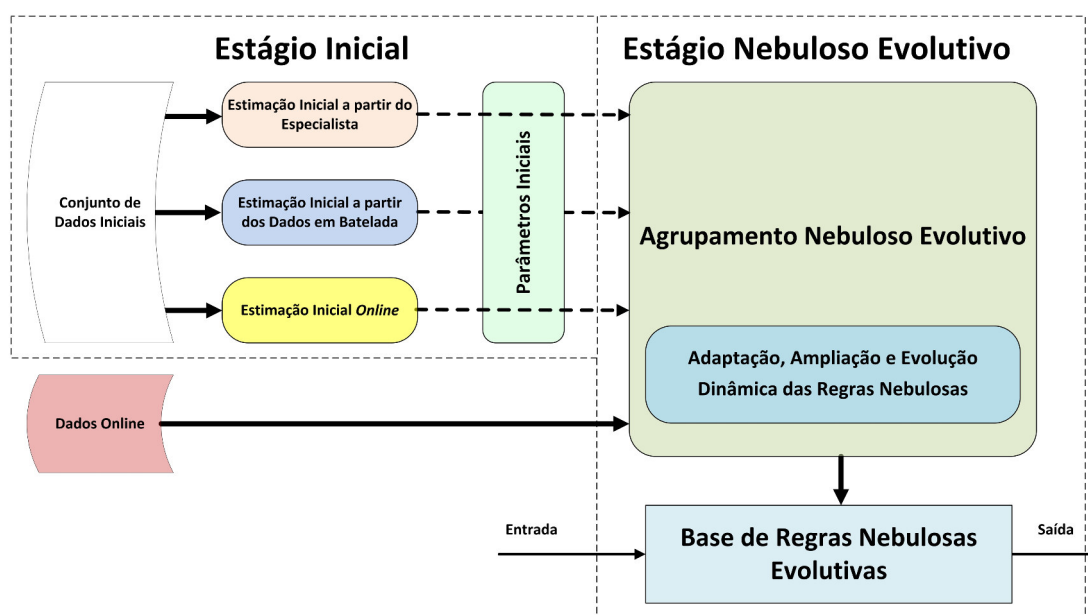
2.1.4 Agrupamento Nebuloso Evolutivo

Modelagem funcional nebulosa evolutiva é uma abordagem com vasta literatura que propõe diversos modelos funcionais nebulosos evolutivos para solução de problemas de engenharia, computação, economia, meteorologia, biomedicina, entre outros, utilizando-se de agrupamentos nebulosos evolutivos (LUGHOFER, 2008; RUBIO, 2009; LEITE, 2012; LEITE; COSTA; GOMIDE, 2013; CAMINHAS; LEMOS; GOMIDE, 2013; ANGELOV, 2013; BORDIGNON; GOMIDE, 2014; BENZAOUIA; HAJJAJI, 2014; DOVZAN; LOGAR; SKRJANC, 2015).

Basicamente, um agrupamento nebuloso evolutivo pode ter seus parâmetros iniciais obtidos a partir de uma estimativa inicial dos dados, no qual é realizado a etapa de agrupamento em batelada do conjunto de dados iniciais (estimativa inicial a partir dos dados

em batelada) (ANGELOV; FILEV; KASABOV, 2010; ANGELOV, 2013). Se não houver disponível um conjunto de dados para realizar a estimação inicial e então disponibilizar os parâmetros iniciais do agrupamento evolutivo, verifica-se que os parâmetros iniciais podem ser obtidos a partir do conhecimento prévio estabelecido pelo especialista (estimação inicial a partir do especialista), ou a partir das condições iniciais determinadas pela primeira amostra lida do fluxo de dados contínuo (estimação inicial *online*) (KASABOV, 2002; KASABOV; FILEV, 2006; KASABOV, 2007; ANGELOV; FILEV; KASABOV, 2010; ANGELOV, 2013), conforme ilustrado na Figura 11. A base de regras nebulosas evolutivas são atualizadas à medida que o modelo nebuloso evolutivo é reconfigurado, ou seja, após a adaptação, ampliação e evolução dinâmicas das regras nebulosas atualize os parâmetros do consequente final. À medida que a estrutura do modelo nebuloso baseado em agrupamento evolutivo se modifica, a partir das amostras (dados) lidas, observa-se uma adaptação dos protótipos já existentes e/ou criação de novos protótipos. Por conseguinte, as amostras lidas são readaptadas para a nova estrutura modificada e então fornecem informações para estimação paramétrica evolutiva do consequente da base de regras nebulosas evolutivas (KASABOV, 2002; KASABOV; FILEV, 2006; KASABOV, 2007; PEDRYCZ; GOMIDE, 2007; ANGELOV; FILEV; KASABOV, 2010; LUGHOFFER, 2011; ANGELOV, 2013).

Figura 11: O Mecanismo de atualização do agrupamento nebuloso evolutivo à medida que os dados de entrada e saída são obtidos dinamicamente. Os parâmetros iniciais do modelo nebuloso evolutivo pode ser extraído de uma modelagem inicial em batelada, de forma *online* das primeiras amostras do conjunto de dados ou determinados pelo especialista.



Fonte: Elaborada pelo autor

A fim de se garantir o desenvolvimento de grupos, há duas perspectivas importantes a serem observadas durante o agrupamento nebuloso evolutivo gradual a partir do fluxo de dados. A primeira perspectiva baseia-se no fato de que o desempenho do agrupamento nebuloso evolutivo deve ser o mais próximo possível ou melhor do que o desempenho obtido através de agrupamento em batelada. A segunda perspectiva é que o agrupamento evolutivo deve ser capaz de convergir, de acordo com um critério de otimização determinado (LUGHOFER, 2011).

Os algoritmos que realizam agrupamento evolutivo baseiam-se na forma em que os protótipos representativos de grupos e parâmetros associados são atualizados a cada nova amostra lida, de acordo com os seguintes princípios básicos:

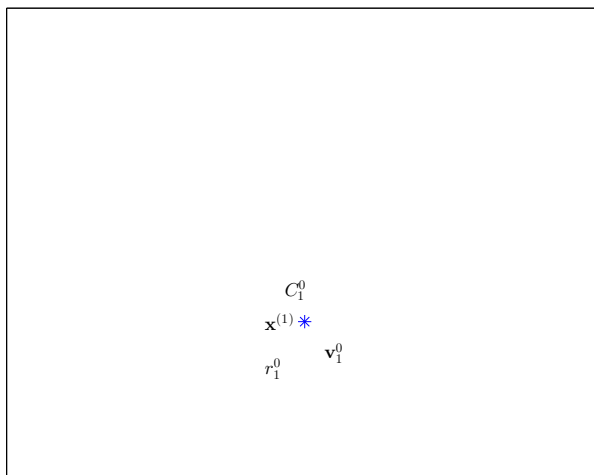
- a) Um novo grupo é criado, onde o centro é determinado a partir do nova amostra lida. O raio deste novo centro é iniciado com zero, quando não há etapa batelada inicial, haja vista que nesta etapa inicial são definidos dois centros inicialmente. Entenda-se raio como sendo a maior distância entre uma amostra e seu respectivo centro de grupo similar em relação às demais amostras existentes neste grupo. Se o grupo contém apenas uma amostra, o seu raio tem valor zero, pois a amostra é o centro do grupo;
- b) À medida que os grupos são atualizados, os protótipos representativos de grupos e parâmetros associados dos mesmos são atualizados; além disso, as amostras pertencentes aos grupos deverão ser reorganizadas à medida que a estrutura é modificada;
- c) Os grupos não serão alterados se o valor da amostra lida estiver dentro dos limites de um grupo próximo;
- d) Um novo grupo será criado se a amostra lida não pertencer aos demais grupos existentes, segundo um critério de similaridade.

No sentido de ilustrar o mecanismo de agrupamento evolutivo, observa-se da Figura 12 a Figura 15, $(* \mathbf{x}^{(k)})$ representa a amostra no instante k ; $(\bullet \mathbf{v}_j^{\text{iteração}})$ representa o centro do agrupamento na iteração corrente; $(r_j^{\text{iteração}})$ representa o raio do agrupamento na iteração corrente; o j -ésimo grupo da iteração corrente é representado por $(C_j^{\text{iteração}})$. O agrupamento evolutivo sem a etapa inicial em batelada é descrito a seguir:

- a) a amostra lida, $\mathbf{x}^{(1)}$, permite ao agrupamento evolutivo criar um novo grupo, C_1^0 ;
- b) a amostra $\mathbf{x}^{(2)}$ atualiza o grupo $C_1^0 \rightarrow C_1^1$; a amostra $\mathbf{x}^{(3)}$ cria um novo grupo C_2^0 ; e a amostra $\mathbf{x}^{(4)}$ não interfere nos grupos existentes;
- c) a amostra lida, $\mathbf{x}^{(5)}$, atualiza o grupo $C_1^1 \rightarrow C_1^2$; a amostra lida, $\mathbf{x}^{(6)}$, não interfere nos grupos existentes; a amostra $\mathbf{x}^{(7)}$ atualiza o grupo $C_2^0 \rightarrow C_2^1$; e a amostra $\mathbf{x}^{(8)}$ cria um novo grupo C_3^0 ;

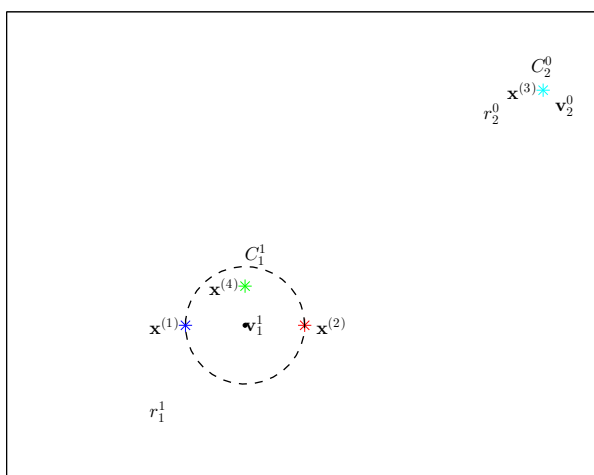
d) a amostra $\mathbf{x}^{(9)}$ atualiza o grupo $C_1^2 \rightarrow C_1^3$.

Figura 12: Exemplo de um breve processo de agrupamento evolutivo com 3 grupos identificados a partir dos dados no espaço \mathbb{R}^2 . O processo inicia-se com apenas uma amostra e à medida que novas amostras são lidas, as mesmas podem ser incluídas em um grupo existente ou determinar a criação de novos grupos. A amostra $\mathbf{x}^{(1)}$ permite ao agrupamento evolutivo criar um novo grupo C_1^0 .



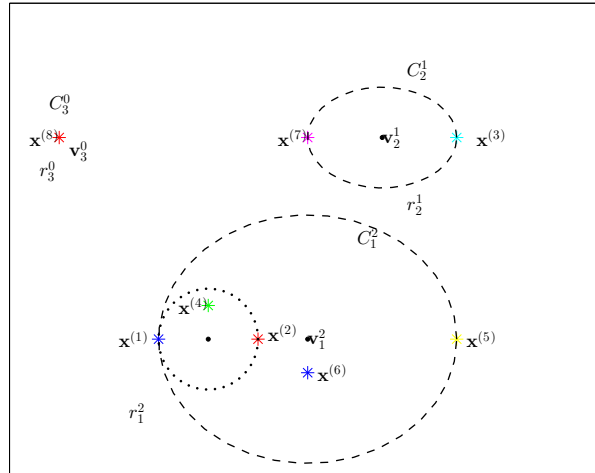
Fonte: Elaborado pelo autor.

Figura 13: Continuação do processo iniciado na Figura 12. A amostra $\mathbf{x}^{(2)}$ atualiza o grupo $C_1^0 \rightarrow C_1^1$, a amostra $x^{(3)}$ cria um novo grupo C_2^0 e a amostra $\mathbf{x}^{(4)}$ é inserida no grupo mais próximo e a estrutura não é alterada.



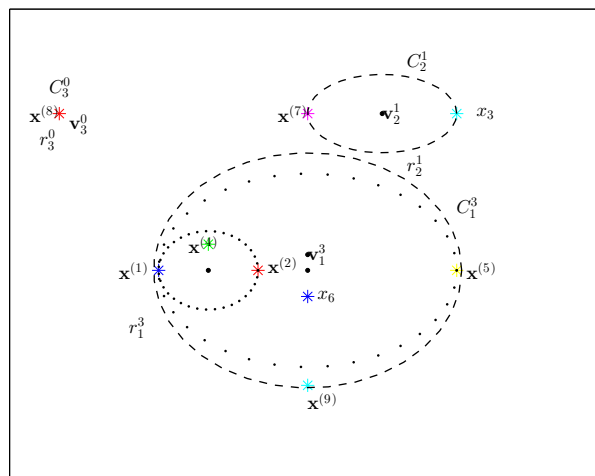
Fonte: Elaborado pelo autor.

Figura 14: A amostra $\mathbf{x}^{(5)}$ atualiza o grupo mais próximo $C_1^1 \rightarrow C_1^2$; a amostra $\mathbf{x}^{(6)}$ é inserida no grupo mais próximo e a estrutura não é alterada; a amostra $\mathbf{x}^{(7)}$ atualiza o grupo $C_2^0 \rightarrow C_2^1$; e a amostra $\mathbf{x}^{(8)}$ cria um novo grupo C_3^0 .



Fonte: Elaborado pelo autor.

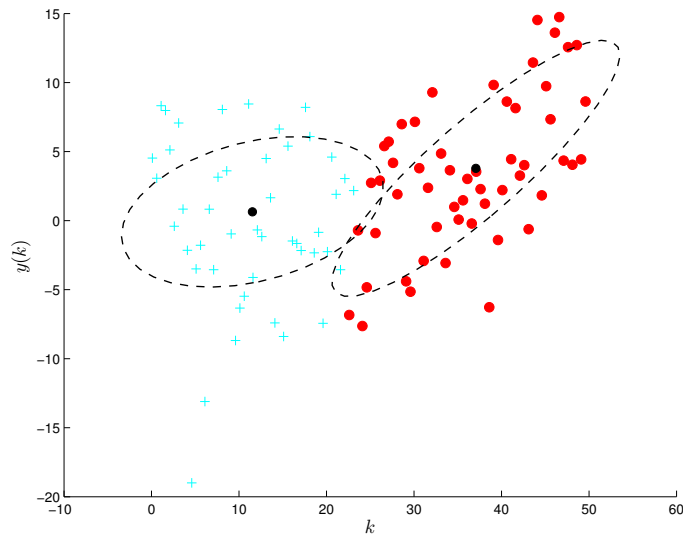
Figura 15: O processo de agrupamento evolutivo modifica a estrutura a cada nova amostra. A amostra $\mathbf{x}^{(9)}$ atualiza o grupo mais próximo $C_1^2 \rightarrow C_1^3$.



Fonte: Elaborado pelo autor.

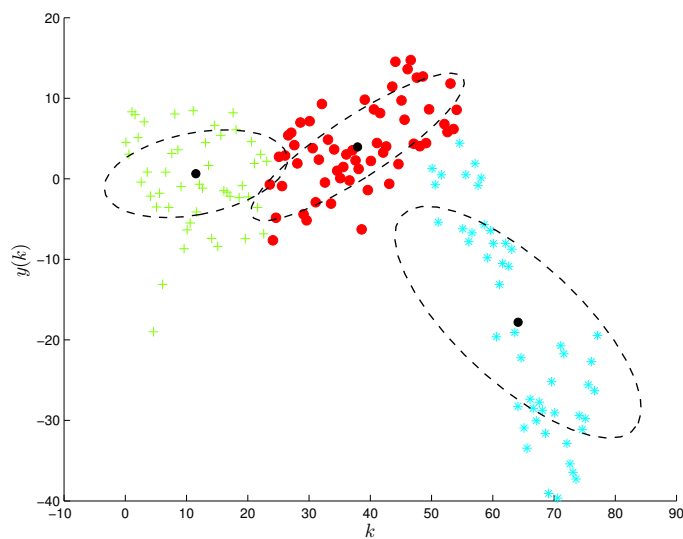
Também, no sentido de ilustrar o agrupamento evolutivo a partir de uma etapa inicial em batelada observa-se da Figura 16 a Figura 19. Este mecanismo é inicializado a partir de uma modelagem inicial em batelada realizada no conjunto de dados em batelada do estágio inicial.

Figura 16: Agrupamento de um Conjunto de Dados obtidos da função não linear. Inicialmente tem-se dois grupos selecionados obtidos a partir de um agrupamento nebuloso em batelada.



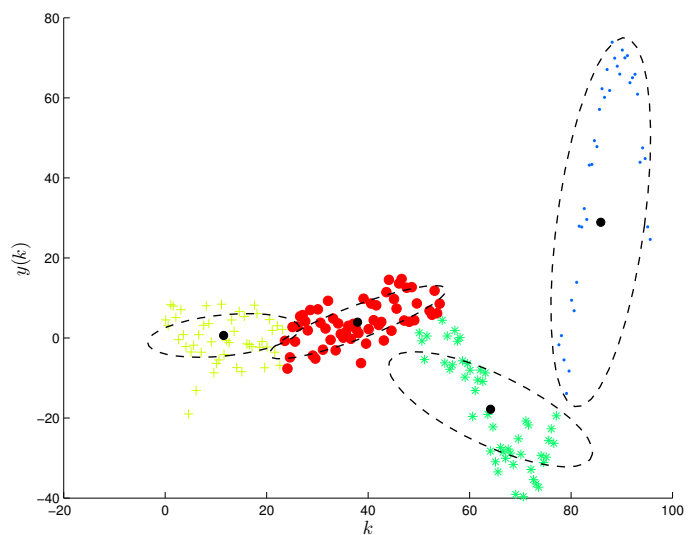
Fonte: Elaborado pelo autor.

Figura 17: O terceiro grupo é identificado e criado. Os parâmetros dos agrupamentos existentes também são ajustados em função das novas amostras lidas.



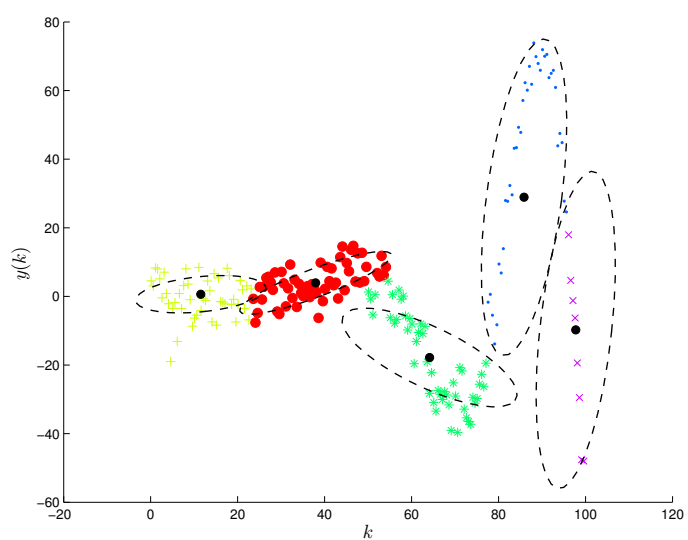
Fonte: Elaborado pelo autor.

Figura 18: O quarto grupo é identificado e criado. Os parâmetros dos agrupamentos existentes também são ajustados em função das novas amostras lidas.



Fonte: Elaborado pelo autor.

Figura 19: O quarto grupo é identificado e criado. Os parâmetros dos agrupamentos existentes também são ajustados em função das novas amostras lidas.



Fonte: Elaborado pelo autor.

Os passos do algoritmo para realização do agrupamento evolutivo são descritos como segue:

Passo 0: São determinadas as condições iniciais de implementação do algoritmo de agrupamento evolutivo. No agrupamento evolutivo ECM (do inglês *Evolving Clustering Method*) baseado na métrica Euclidiana, proposto por (KASABOV, 2002), cria-se o primeiro centro a partir da primeira amostra lida. No agrupamento evolutivo baseado em (Gustafson–Kessel) eGK (do inglês *Evolving Clustering Based on GK Similarity Distance*) apresentado por (GEORGIEVA; FILEV, 2009), determina-se os parâmetros iniciais dos vetores de protótipos, matriz de covariância e partição nebulosa a partir do agrupamento em batelada baseado no método GK.

Passo 1: No instante de tempo k é lida uma nova amostra. Logo, é mensurada a distância desta amostra em relação aos centros dos grupos existentes.

Passo 2: Calcula-se a distância da amostra lida em relação aos centros já existentes. No agrupamento evolutivo ECM, a distância é mensurada a partir da métrica Euclidiana (KASABOV, 2002). A distância é mensurada a partir da norma de *Mahalanobis* quando se utiliza o agrupamento evolutivo eGK (ANGELOV; FILEV; KASABOV, 2010; GEORGIEVA; FILEV, 2009). Nesta tese, adota-se o cálculo das distâncias via critério de máxima verossimilhança.

Passo 3: Encontra-se o grupo mais próximo à amostra lida.

Passo 4: Atualiza-se o grupo ECM mais próximo sujeito à condição restrita de que a distância do grupo mais próximo é menor que um limiar pré-estabelecido. No agrupamento evolutivo eGK, atualiza-se o grupo mais próximo sujeito a condição restrita de que a distância é menor que o raio do grupo mais próximo. Nesta tese, atualizam-se todos os grupos existentes, haja vista que o grupo mais próximo indica que não há necessidade de ser criado um novo grupo, nas mesmas condições pré-estabelecidas pelo agrupamento eGK.

Passo 5: Cria-se um novo grupo se a condição restrita do **Passo 4** não for satisfeita. Nos agrupamentos evolutivos ECM e eGK o grupo é criado com raio zero. Nesta tese, ao criar um novo grupo, todas as amostras já lidas são atualizadas de acordo com os grupos já criados, e então calcula-se um novo raio para os grupos existentes.

Passo 6: Atualiza-se os parâmetros dos agrupamentos evolutivos.

Passo 7: Se todas amostras existentes em um fluxo de dados já foram processadas, encerra-se o agrupamento evolutivo. Caso contrário, retorna ao **Passo 1**.

2.2 Modelo Nebuloso Takagi–Sugeno

O modelo nebuloso Takagi–Sugeno é descrito por um conjunto de regras do tipo **SE** < antecedente > **ENTÃO** < consequente > o qual representa as relações locais dos sinais de entrada/saída de um sistema dinâmico. A principal característica de um modelo nebuloso *Takagi–Sugeno* é a aproximação de sistemas dinâmicos não lineares a partir de modelos locais lineares. Este é ideal para se trabalhar com identificação nebulosa de sistemas dinâmicos com aplicações nas áreas de modelagem e controle, como pode ser visto na literatura que aborda o referido tema: (TAKAGI; SUGENO, 1985; BABUSKA, 1998; PEDRYCZ; GOMIDE, 2007; ABONYI; FEIL, 2007; SERRA; BOTTURA, 2007; ANGELOV; FILEV; KASABOV, 2010; LUGHOFFER, 2011; SERRA, 2012; ANGELOV, 2013; JEEVA; PONNUSAMY; SANGAPILLAI, 2013; BENZAOUIA; HAJJAJI, 2014).

2.2.1 Estrutura Nebulosa *Takagi–Sugeno*

Seja $\mathbf{x}^{(k)} \in \mathbb{R}^p = [x_{(1)}^{(k)} \ x_{(2)}^{(k)} \ \dots \ x_{(p)}^{(k)}]^T$, para $k = 1, 2, \dots, N$, as regras nebulosas do modelo *Takagi–Sugeno* são definidas como segue:

$$\text{Regra}_{i|i=1,2,\dots,c} : \text{SE } x_{(1)}^{(k)} \text{ É } \Gamma_{(i)}^{(1)} \text{ E } \dots \text{ E } x_{(p)}^{(k)} \text{ É } \Gamma_{(i)}^{(p)} \text{ ENTÃO } y_{(i)}^{(k)} = f_{(i)}(\mathbf{x}^{(k)}). \quad (2.35)$$

O vetor $\mathbf{x}^{(k)} \in \mathbb{R}^p$ tem seu universo de discurso particionado por regiões nebulosas; a variável $x_{(j)}^{(k)}$, com $j = 1, 2, \dots, p$ pertence a um conjunto nebuloso $\Gamma_{(i)}^{(j)}$, com $i = 1, 2, \dots, c$, e $j = 1, 2, \dots, p$, definido por uma função de pertinência dada por:

$$\mu_{\Gamma_{(i)}^{(j)}}(\mathbf{x}^{(k)}) : \mathbb{R}^p \longrightarrow [0, 1]^p. \quad (2.36)$$

2.2.2 Inferência Nebulosa *Takagi–Sugeno*

O mecanismo de inferência do modelo nebuloso *Takagi–Sugeno* define o grau de ativação de cada regra para um vetor de entrada característico e este, por sua vez, determina o grau de disparo de uma regra conforme os valores das variáveis de entrada. O grau de ativação, β , da i -ésima regra é calculada a partir de operadores de agregação, ou seja, de uma norma T, operador E, denotado por $\otimes : [0, 1] \times [0, 1] \longrightarrow [0, 1]$:

$$\beta^i(\mathbf{x}^{(k)}) = \mu_{\Gamma_{(i)}^{(1)}}(x_{(1)}^{(k)}) \otimes \mu_{\Gamma_{(i)}^{(2)}}(x_{(2)}^{(k)}) \otimes \dots \otimes \mu_{\Gamma_{(i)}^{(p)}}(x_{(p)}^{(k)}) \quad (2.37)$$

De acordo com (WANG, 1997) uma norma T, $t : [0, 1] \times [0, 1] \longrightarrow [0, 1]$ mapeia as funções de pertinências dos conjuntos nebulosos $\Gamma_{(i)}^{(j)}$, $j = 1, 2, \dots, p$ no espaço definido pela funções de pertinências da interseção de $\Gamma_{(i)}^{(1)} \text{ E } \dots \text{ E } \Gamma_{(i)}^{(p)}$, isto é,

$$t[\mu_{\Gamma_{(i)}^{(j)}(x)}, \mu_{\Gamma_{(i)}^{(k)}(x)}] = \mu_{\Gamma_{(i)}^{(j)}(x) \cap \Gamma_{(i)}^{(k)}(x)}, \quad i \neq j. \quad (2.38)$$

Uma norma T deve satisfazer os seguintes axiomas:

Axioma t1: $t(0, 0) = 0; t(\mu_{\Gamma^{(i)}}(x), 1) = t(1, \mu_{\Gamma^{(i)}}(x)) = \mu_{\Gamma^{(i)}}(x)$, condições de contorno;

Axioma t2: $t(\mu_{\Gamma^{(i)}}(x), \mu_{\Gamma^{(j)}}(x)) = t(\mu_{\Gamma^{(j)}}(x), \mu_{\Gamma^{(i)}}(x))$ $i \neq j$, comutatividade.

Axioma t3: Se $\mu_{\Gamma^{(i)}}(x) \leq \mu'_{\Gamma^{(i)}}(x)$ e $\mu_{\Gamma^{(j)}}(x) \leq \mu'_{\Gamma^{(j)}}(x)$, então $t(\mu_{\Gamma^{(j)}}(x), \mu_{\Gamma^{(i)}}(x)) \leq t(\mu'_{\Gamma^{(i)}}(x), \mu'_{\Gamma^{(j)}}(x))$, não-decrescente.

Axioma t4: $t[t(\mu_{\Gamma^{(i)}}(x), \mu_{\Gamma^{(j)}}(x)), \mu_{\Gamma^{(l)}}(x)] = t[\mu_{\Gamma^{(i)}}(x), t(\mu_{\Gamma^{(j)}}(x), \mu_{\Gamma^{(l)}}(x))]$ $i \neq j \neq l$, associatividade.

Qualquer função que satisfaça os axiomas t1–t4 é considerada uma norma T. Neste trabalho, o produto algébrico será utilizado como norma T. Ele é expresso por

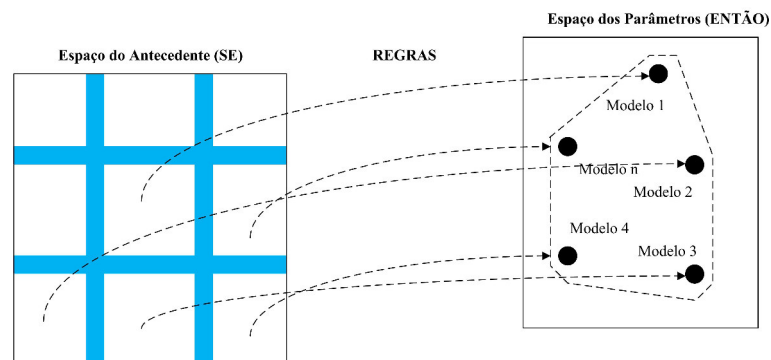
$$\beta^{(i)}(\mathbf{x}^{(k)}) = \prod_{j=1}^p \mu_{\Gamma^{(j)}}^{(i)}(\mathbf{x}^{(k)}). \quad (2.39)$$

A saída *defuzzificada* do sistema nebuloso é dada pela média ponderada das funções consequentes:

$$y^{(k)} = \frac{\sum_{i=1}^c \beta^{(i)}(\mathbf{x}^{(k)}) f_{(i)}(\mathbf{x}^{(k)})}{\sum_{j=1}^c \beta^{(j)}(\mathbf{x}^{(k)})} \quad k = 1, 2, \dots, N. \quad (2.40)$$

Na Figura 20 observa-se como o modelo Takagi–Sugeno pode ser considerado como um mapeamento do espaço do antecedente (entrada) para uma região de convergência no espaço dos submodelos locais definidos pelos parâmetros do consequente e além disto, verifica-se que o mapeamento da composição de cada regra está relacionado a um submodelo consequente. Esta propriedade dos modelos de inferência do tipo *Takagi–Sugeno* de estrutura linear no consequente permite modelagens robustas em diversos sistemas dinâmicos (BABUSKA, 1998; SERRA; BOTTURA, 2009).

Figura 20: Mapeamento do conjunto de regras definidas no Antecedente ao conjunto de funções ou submodelos definidos no Consequente.



Fonte: Modificado de (BABUSKA, 1998)

2.3 Normalização e Padronização

A maioria dos problemas reais em engenharia envolvem dados com diferentes características e faixa de valores de amplitude variadas. A fim de aplicar corretamente as métricas de distâncias utilizadas nas normas de *Mahalanobis* e máxima verossimilhança, verifica-se que as diferentes características de um conjunto de dados precisam de uma normalização ou padronização. A padronização e/ou normalização garante que as diversas características de um conjunto de dados estarão na mesma faixa de valores e não serão influenciadas durante o agrupamento quando utilizado com normas adaptativas (BABUSKA, 1998; ANGELOV, 2013).

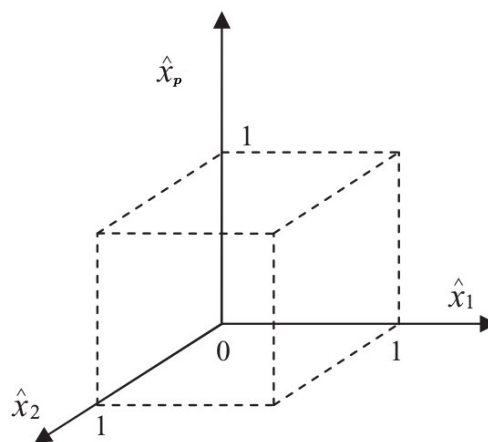
2.3.1 Normalização

A normalização tem como principal objetivo mapear um conjunto de dados com diferentes faixa de valores e amplitude para dentro da faixa de valores de $[0, 1]$. Isto é realizado a partir da equação que se segue (ANGELOV, 2013):

$$\hat{x}^{(k)} = \frac{x^{(k)} - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})} \quad k = 1, 2, \dots, N, \quad (2.41)$$

onde $\hat{x}^{(k)}$ representa o valor da amostra de $x^{(k)}$ estimada no instante k , $\max(\mathbf{x})$ e $\min(\mathbf{x})$ os valores máximos e mínimos, respectivamente do conjunto de dados \mathbf{x} . Na Figura 21 é ilustrado um hipercubo unitário na qual todas amostradas de um conjunto de dados são transformadas.

Figura 21: hipercubo unitário no qual todos os dados são normalizados dentro da faixa de valores de $[0, 1]$.



Fonte: Modificado de (ANGELOV, 2013).

Obviamente, percebe-se que a fórmula (2.41) requer os valores máximos e mínimos do conjunto de dados, bastante utilizada em problemas de modelagem quando se trabalha com dados em batelada; porém para dados em processamento *online* ou em tempo real, a normalização torna-se um problema, considerando que não se tem os valores máximos e mínimos durante o processo de leitura.

2.3.2 Padronização

Esta forma de normalização não é indicada para dados *online*. A padronização de dados é indicada para problemas de dados *online* (BABUSKA, 1998; ANGELOV, 2013). A equação para padronização, no caso de fluxo de dados é:

$$\hat{x}^{(k)} = \frac{x^{(k)} - \bar{x}^{(k)}}{\sigma_x^{(k)}} \quad k = 1, 2, \dots, N. \quad (2.42)$$

De acordo com (ANGELOV, 2013), a média e variância não serão as mesmas para cada elemento do vetor de características; além disso, a distribuição estatística de cada elemento não será a mesma para os demais, mas pelo Teorema *Chebyshev* (A.PAPOULIS; PILAI, 2002), observa-se que neste processo de padronização, as amostras estarão dentro do intervalo de $[-3, +3]$, com probabilidade muito pequena de existir amostras com valores fora desta faixa.

A média e o desvio-padrão no instante k , $\bar{x}^{(k)}$ e $\sigma_x^{(k)}$, respectivamente, são obtidos a partir de,

$$\bar{x}^{(k)} = \frac{k-1}{k} \bar{x}^{(k-1)} + \frac{1}{k} x^{(k)} \quad (2.43)$$

e

$$\left(\sigma_x^{(k)}\right)^2 = \frac{k-1}{k} \left(\sigma_x^{(k-1)}\right)^2 + \frac{1}{k} \left(Dm^{(k)} - Dn^{(k)}\right)^2 \quad (2.44)$$

onde as densidades da amostra no instante k , $Dm^{(k)}$ e $Dn^{(k)}$ são calculadas a partir de

$$Dm^{(k)} = \frac{k-1}{k} Dm^{(k-1)} + \frac{1}{k} Dn^{(k)} \quad (2.45)$$

de modo que

$$Dn^{(k)} = \frac{1}{1 + (x^{(k)})^2 - 2x^{(k)} \frac{1}{k} \sum_{j=1}^k x^{(j)} + \frac{1}{k} \sum_{j=1}^k (x^{(j)})^2}. \quad (2.46)$$

A padronização é bastante conveniente para dados *online*, porque esta possibilita a atualização da média e desvio-padrão de forma *online* para cada instante k de observação (ANGELOV, 2013).

Este capítulo apresentou as principais definições relacionadas à composição desta tese, destacando-se os principais conceitos para o desenvolvimento da metodologia proposta.

3 ESTIMAÇÃO PARAMÉTRICA COM VARIÁVEL INSTRUMENTAL NEBULOSA

Os modelos nebulosos do tipo *Takagi–Sugeno* representam uma possível solução para problemas de identificação de sistemas não lineares e incertos, permitindo a construção de uma função do consequente na forma linear, onde o vetor de regressão é representado por um número finito de entradas e saídas do sistema dinâmico a ser modelado. Na etapa da estimação paramétrica verifica-se a necessidade de estimadores que não sejam polarizados e tendenciosos, basta ver que há, possivelmente, a presença de ruído autocorrelacionado (perturbação do sistema) em problemas reais de engenharia e os métodos tradicionais de estimação paramétrica falham nestes casos. Uma provável solução para contornar este problema, é o uso de estimadores de variáveis instrumentais (GOODWIN; PAYNE, 1977; YOUNG, 1984; SODERSTROM; STOICA, 1989; LJUNG, 1999; YOUNG, 2011; WANG; GARNIER, 2012; YOUNG, 2015; AGUIRRE, 2015).

Neste capítulo, será apresentado a metodologia de variável instrumental nebulosa inserida no contexto evolutivo, baseada em máxima verossimilhança, abordada no contexto não–recursivo (batelada) e recursivo; também será realizado a análise de convergência da variável instrumental nebulosa utilizada nesta tese.

3.1 Estrutura do Modelo

Nesta tese, a estrutura de modelo ARX será usada. Este modelo é aplicado na maioria das metodologias de identificação de sistemas não lineares e incertos inseridos no consequente de modelos nebulosos do tipo *Takagi–Sugeno*.

Seja o seguinte modelo ARX a ser descrito,

$$y^{(k)} = \left(\boldsymbol{\psi}^{(k-1)} \right)^T \boldsymbol{\Theta} + e^{(k)} \quad (3.1)$$

e

$$\boldsymbol{\psi}^{(k-1)} = \left[y^{(k-1)} \ y^{(k-2)} \ \dots \ y^{(k-n_y)} \ \dots \ u^{(k-1)} \ u^{(k-2)} \ \dots \ u^{(k-n_u)} \right] \quad (3.2)$$

onde no instante k de tempo discreto, $\boldsymbol{\psi}^{(k-1)}$ é o vetor de regressores; n_y e n_u a dimensionalidade das variáveis \mathbf{u} e \mathbf{y} , respectivamente; $\boldsymbol{\Theta}$ é o vetor de parâmetros; e $e^{(k)}$ o resíduo do sistema.

3.2 Estimadores de Mínimos Quadrados

A estimação paramétrica do vetor de parâmetros Θ pode ser obtida a partir do método dos mínimos quadrados não-recursivo (batelada) ou de forma recursiva (SODERSTROM; STOICA, 1983; SODERSTROM; STOICA, 1989; YOUNG, 2011; YOUNG, 2015; AGUIRRE, 2015). As formulações destes métodos são descritas a seguir:

3.2.1 Método dos Mínimos Quadrados Não-Recursivo (Batelada)

O problema real é estimar um vetor de parâmetros $\hat{\Theta}$ do vetor de parâmetros real Θ a partir dos N pares do conjunto de dados $\{y^{(k)}, \psi^{(k)}\}$, $k = 1, 2, \dots, N$ (SODERSTROM; STOICA, 1989; AGUIRRE, 2015). Dados os N pares de medidas, obtém-se o seguinte sistema linear, expresso por

$$\begin{aligned} y(1) &= (\psi^{(1)})^T \Theta, \\ y(2) &= (\psi^{(2)})^T \Theta, \\ &\vdots \\ y(N) &= (\psi^{(N)})^T \Theta. \end{aligned}$$

As equações acima podem ser reescritas na forma matricial como

$$\mathbf{Y} = \Psi \Theta, \quad (3.3)$$

onde

$$\mathbf{Y} = \begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{pmatrix} \quad (3.4)$$

e

$$\Psi = \begin{pmatrix} (\psi^{(1)})^T \\ (\psi^{(2)})^T \\ \vdots \\ (\psi^{(N)})^T \end{pmatrix}. \quad (3.5)$$

De forma similar pode-se escrever o resíduo do sistema da seguinte forma:

$$\mathbf{e} = \begin{pmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{pmatrix}. \quad (3.6)$$

De acordo com (SODERSTROM; STOICA, 1989; AGUIRRE, 2015), a solução do método dos mínimos quadrados consiste em determinar um vetor de parâmetros estimado $\hat{\Theta}$

de modo a minimizar a função de custo (GOODWIN; PAYNE, 1977; SODERSTROM; STOICA, 1989):

$$V(\Theta) = \frac{1}{2} \sum_{k=1}^N (e^{(k)})^2. \quad (3.7)$$

A minimização da função de custo (3.7) permite encontrar o vetor de parâmetros estimado, a minimização é dada por

$$\min_{\Theta} V(\Theta) = V(\hat{\Theta}). \quad (3.8)$$

Fazendo o gradiente da equação (3.7) igual a zero,

$$\frac{d}{d\Theta} V(\Theta) = 0. \quad (3.9)$$

Obtém-se a solução para o método dos mínimos quadrados não-recursivo (SODERSTROM; STOICA, 1989; AGUIRRE, 2015), expressa por

$$\hat{\Theta} = (\Psi \Psi^T)^{-1} \Psi^T \mathbf{Y}. \quad (3.10)$$

3.2.2 Método dos Mínimos Quadrados Recursivo

Na identificação de sistemas cujos dados variam com tempo, percebe-se que as observações mais recentes devem influenciar na estimação dos parâmetros, considerando-se que estas observações contêm informações mais atualizadas do conjunto de dados, do modelo a ser identificado, no instante de tempo da observação.

Neste caso, no método dos mínimos quadrados recursivo é inserido um fator de esquecimento e este pode ser atualizado a cada nova iteração, de maneira que, os valores baixos, indicam que os dados atuais serão esquecidos em relação aos dados anteriores, ou seja, dados atuais tem menos influência na estimação paramétrica do vetor de parâmetros (GOODWIN; PAYNE, 1977; SODERSTROM; STOICA, 1989; ISERMANN; NCHHOF, 2011; WANG; GARNIER, 2012; AGUIRRE, 2015).

O método dos mínimos quadrados recursivo com fator de esquecimento é expresso por,

$$\hat{\Theta}^{(k)} = \hat{\Theta}^{(k-1)} + \mathbf{K}_{\hat{\Theta}}^{(k)} \left(y^{(k)} - \psi^{(k)} \hat{\Theta}^{(k-1)} \right), \quad (3.11)$$

onde o ganho de adaptação do vetor estimado, referido como ganho de *Kalman* é dado por,

$$\mathbf{K}_{\hat{\Theta}}^{(k)} = \frac{\mathbf{P}_{\hat{\Theta}}^{(k-1)} \psi^{(k)}}{\delta_{\hat{\Theta}}^{(k)} + \psi^{(k)T} \mathbf{P}_{\hat{\Theta}}^{(k-1)} \psi^{(k)}}. \quad (3.12)$$

A matriz de covariância dos estimadores paramétricos recursivos é expressa por

$$\mathbf{P}_{\hat{\Theta}}^{(k)} = \frac{1}{\delta_{\hat{\Theta}}^{(k)}} \left(\mathbf{P}_{\hat{\Theta}}^{(k-1)} - \frac{\mathbf{P}_{\hat{\Theta}}^{(k-1)} \psi^{(k)} \psi^{(k)T} \mathbf{P}_{\hat{\Theta}}^{(k-1)}}{\delta_{\hat{\Theta}}^{(k)} + \psi^{(k)T} \mathbf{P}_{\hat{\Theta}}^{(k-1)} \psi^{(k)}} \right). \quad (3.13)$$

onde, $0 < \delta_{\Theta}^{(k)} < 1$ é o fator de esquecimento cujo valor típico é da ordem de 0,99 (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015).

3.2.3 Problema da Estimação do Método dos Mínimos Quadrados

Considerando os problemas reais de engenharia, observa-se que os valores dos sinais de entrada e saída estão sujeitos a ruído, ou seja, o vetor de regressores com erro nas variáveis será considerado na equação (3.1), de modo que esta é reescrita da seguinte forma

$$y^{(k)} = \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right)^T \boldsymbol{\Theta} + \eta^{(k)} \quad (3.14)$$

e

$$\boldsymbol{\xi}^{(k-1)} = \left[\xi^{(k-1)} \quad \xi^{(k-2)} \quad \dots \quad \xi^{(k-n_y)} \quad \dots \quad \xi^{(k-1)} \quad \xi^{(k-2)} \quad \dots \quad \xi^{(k-n_u)} \right], \quad (3.15)$$

onde $\boldsymbol{\xi}^{(k)}$ representa a perturbação do sistema no instante k , e $\eta^{(k)}$ é o ruído.

De acordo com (SODERSTROM; STOICA, 1989; YOUNG, 2011; WANG; GARNIER, 2012; AGUIRRE, 2015), as equações normais são expressas por:

$$\left[\sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right)^T \right] \hat{\boldsymbol{\Theta}} = \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) y^{(k)}. \quad (3.16)$$

Na equação (3.16) multiplica-se ambos os lados por $\frac{1}{N}$, para se obter:

$$\left[\frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right)^T \right] \hat{\boldsymbol{\Theta}} = \frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) y^{(k)}. \quad (3.17)$$

Substituindo a equação (3.14) na equação (3.17) obtém-se

$$\begin{aligned} & \left[\frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right)^T \right] \hat{\boldsymbol{\Theta}} = \dots \\ & \frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \left[\left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right)^T \boldsymbol{\Theta} + \eta^{(k)} \right]. \end{aligned} \quad (3.18)$$

Reorganizando a equação (3.18), obtém-se

$$\begin{aligned} & \left[\frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right)^T \right] (\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}) = \dots \\ & \frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \eta^{(k)}. \end{aligned} \quad (3.19)$$

O termo $\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}$ é o erro paramétrico no instante k . Isolando-o na equação (3.19) obtém-se

$$\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta} = \left[\frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right)^T \right]^{-1} \frac{1}{N} \sum_{k=1}^N \left(\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)} \right) \eta^{(k)}. \quad (3.20)$$

Obtendo-se o limite no sentido da probabilidade quando $N \rightarrow \infty$, os termos da equação (3.20) tendem para seus valores esperados $\tilde{E}\{\bullet\}$ (SODERSTROM; STOICA, 1989; WANG; GARNIER, 2012). Os resultados finais são expressos da seguinte forma

$$p.\lim_{N \rightarrow \infty} \left[\frac{1}{N} \sum_{k=1}^N (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)}) (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)})^T \right]^{-1} = \dots$$

$$\tilde{E} \left\{ (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)}) (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)})^T \right\}, \quad (3.21)$$

$$p.\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)}) \eta^{(k)} = \tilde{E} \left\{ (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)}) \eta^{(k)} \right\}. \quad (3.22)$$

Diante de dados completamente livre de ruído ou considerando a independência estatística entre as variáveis $\boldsymbol{\psi}^{(k)}$, $\boldsymbol{\xi}^{(k)}$ e $\eta^{(k)}$, observa-se que a equação (3.22) anula-se, ou seja,

$$\tilde{E} \left\{ (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)}) \eta^{(k)} \right\} = 0. \quad (3.23)$$

Porém, na presença de dados que não são completamente livres de ruído e/ou as variáveis $\boldsymbol{\psi}^{(k)}$, $\boldsymbol{\xi}^{(k)}$ e $\eta^{(k)}$ são correlacionadas, a equação (3.23) não será verdadeira, portanto o método dos mínimos quadrados não-recursivo e recursivo não garantirá a veracidade do cálculo do vetor de parâmetros estimado, conforme (SODERSTROM; STOICA, 1983; SODERSTROM; STOICA, 1989; YOUNG, 2011; WANG; GARNIER, 2012; SERRA, 2012; YOUNG, 2015; AGUIRRE, 2015).

3.2.4 Vetor de Variável Instrumental

Para superar o problema da estimação do método de mínimos quadrados diante de problemas de polarização e inconsistência causados por ruído autocorrelacionado, a geração de um vetor de variáveis independentes estatisticamente do ruído e correlacionado com o vetor de regressores $\boldsymbol{\psi}^{(k)}$ a partir do sistema dinâmico, se faz necessária. Inicialmente, propõem-se uma variável $\tilde{\boldsymbol{\psi}}^{(k)}$ que seja descorrelacionada com a perturbação do sistema conforme (SODERSTROM; STOICA, 1983; SODERSTROM; STOICA, 1989; YOUNG, 2011; WANG; GARNIER, 2012; YOUNG, 2015; AGUIRRE, 2015), e modificando a equação (3.16), obtém-se as seguintes equações normais,

$$\left[\sum_{k=1}^N \left(\tilde{\boldsymbol{\psi}}^{(k-1)} \right) (\boldsymbol{\psi}^{(k-1)} + \boldsymbol{\xi}^{(k-1)})^T \right] \hat{\boldsymbol{\theta}} = \sum_{k=1}^N \tilde{\boldsymbol{\psi}}^{(k-1)} y^{(k)}, \quad (3.24)$$

onde $\tilde{\boldsymbol{\psi}}^{(k)}$ é um vetor de variáveis instrumentais. Este vetor pode ser escolhido de diferentes maneiras (a serem exemplificadas nas próximas seções) de modo a garantir as condições de consistência na equação (3.24). Fazendo-se as modificações utilizando o vetor de variável instrumental nas considerações adotadas para as equações (3.17)–(3.20), obtém-se

o seguinte termo para o erro paramétrico,

$$\hat{\Theta} - \Theta = \left[\frac{1}{N} \sum_{k=1}^N \left(\tilde{\psi}^{(k-1)} \right) \left(\psi^{(k-1)} + \xi^{(k-1)} \right)^T \right]^{-1} \frac{1}{N} \sum_{k=1}^N \left(\tilde{\psi}^{(k-1)} \right) \eta^{(k)}. \quad (3.25)$$

Obtendo-se o limite no sentido de probabilidade quando $N \rightarrow \infty$, os termos da equação (3.25) tendem para seus valores esperados $\tilde{E}\{\bullet\}$ (SODERSTROM; STOICA, 1989; WANG; GARNIER, 2012). Os resultados finais são expressos da seguinte forma

$$p.\lim_{N \rightarrow \infty} \left[\frac{1}{N} \sum_{k=1}^N \left(\tilde{\psi}^{(k-1)} \right) \left(\psi^{(k-1)} + \xi^{(k-1)} \right)^T \right]^{-1} = \dots$$

$$\tilde{E} \left\{ \left(\tilde{\psi}^{(k-1)} \right) \left(\psi^{(k-1)} + \xi^{(k-1)} \right)^T \right\}, \quad (3.26)$$

$$p.\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \left(\tilde{\psi}^{(k-1)} \right) \eta^{(k)} = \tilde{E} \left\{ \left(\tilde{\psi}^{(k-1)} \right) \eta^{(k)} \right\}. \quad (3.27)$$

Os elementos do vetor $\tilde{\psi}^{(k)}$ são usualmente denominados de instrumentos ou variáveis instrumentais, e são determinados de acordo o problema em que o método do mínimos quadrados não-recursivo ou recursivo está inserido; além disso, estes devem ser totalmente descorrelacionados com qualquer perturbação do sistema e altamente correlacionados com as variáveis do vetor de regressores (SODERSTROM; STOICA, 1983; SODERSTROM; STOICA, 1989; YOUNG, 2011; WANG; GARNIER, 2012; YOUNG, 2015; AGUIRRE, 2015). As variáveis $\tilde{\psi}^{(k)}$, $\xi^{(k)}$ e $\eta^{(k)}$ são descorrelacionadas, portanto, a equação (3.27) para o limite no sentido da probabilidade quando $N \rightarrow \infty$, é expressa por

$$\tilde{E} \left\{ \left(\tilde{\psi}^{(k-1)} \right) \eta^{(k)} \right\} = 0. \quad (3.28)$$

A variável instrumental estendida é uma generalização da equação (3.24) com a inclusão de uma etapa de pré-filtragem (SODERSTROM; STOICA, 1983; WANG; GARNIER, 2012; YOUNG, 2015), dada por:

$$\hat{\Theta} = \arg \min_{\theta} \left\| \left[\sum_{k=1}^N \tilde{\psi}^{(k)} \mathbf{F}(z)^{-1} \left(\psi^{(k)} \right)^T \right] \Theta - \left[\sum_{k=1}^N \tilde{\psi}^{(k)} \mathbf{F}(z)^{-1} y^{(k)} \right] \right\|_{\mathbf{W}}^2. \quad (3.29)$$

A variável instrumental estendida é $\tilde{\psi}^{(k)}$; $\mathbf{F}(z^{-1})$ é um filtro assintótico estável; e $\|x\|_{\mathbf{W}}^2 = x^T \mathbf{W} x$, onde \mathbf{W} é uma matriz de pesos definida positiva. A solução é obtida a partir de:

$$\hat{\Theta} = \left(\tilde{\mathbf{R}}_N^T \mathbf{W} \tilde{\mathbf{R}}_N \right)^{-1} \tilde{\mathbf{R}}_N^T \mathbf{W} r_N \quad (3.30)$$

onde

$$\tilde{\mathbf{R}}_N = \frac{1}{N} \sum_{k=1}^N \tilde{\psi}^{(k)} \mathbf{F}(z^{-1}) \left(\psi^{(k)} \right)^T \quad (3.31)$$

e

$$r_N = \frac{1}{N} \sum_{k=1}^N \tilde{\boldsymbol{\psi}}^{(k)} \mathbf{F}(z^{-1}) y^{(k)}. \quad (3.32)$$

A escolha do vetor de variável instrumental estendida permitirá que as variáveis $\tilde{\boldsymbol{\psi}}^{(k)}$, $\boldsymbol{\xi}^{(k)}$ e $\eta^{(k)}$ sejam descorrelacionadas. Portanto, a equação (3.27) para o limite no sentido da probabilidade quando $N \rightarrow \infty$, em termos de variável instrumental estendida é expressa por

$$\tilde{E} \left\{ \left(\tilde{\boldsymbol{\psi}}^{(k-1)} \right) \mathbf{F}(z^{-1}) \eta^{(k)} \right\} = 0. \quad (3.33)$$

3.2.5 Escolha dos Instrumentos

A escolha da variável instrumental $\tilde{\boldsymbol{\psi}}$ deve ser realizada de tal forma que esta esteja estatisticamente correlacionada com os regressores do modelo $\boldsymbol{\psi}$ e que satisfaça a condição imposta pela equação (3.28) e que seja descorrelacionada da perturbação do sistema (GOODWIN; PAYNE, 1977; SODERSTROM; STOICA, 1989; LJUNG, 1999; YOUNG, 2011; WANG; GARNIER, 2012; YOUNG, 2015; AGUIRRE, 2015).

Inicialmente, seja o modelo ARX expresso da seguinte forma:

$$y^{(k)} = \left(\boldsymbol{\psi}^{(k-1)} \right)^T \hat{\boldsymbol{\Theta}}_{MQ} + e^{(k)}. \quad (3.34)$$

Para um vetor de regressores denotado por,

$$\boldsymbol{\psi}^{(k-1)} = \left[y^{(k-1)} \dots y^{(k-n_y)} \ u^{(k-1)} \dots u^{(k-n_u)} \right]^T. \quad (3.35)$$

De acordo com (YOUNG, 2011; YOUNG, 2015; AGUIRRE, 2015), uma sugestão de variável instrumental é descrita na forma:

$$\tilde{\boldsymbol{\psi}}^{(k-1)} = \left[\hat{y}^{(k-1)} \dots \hat{y}^{(k-n_y)} \ u^{(k-1)} \dots u^{(k-n_u)} \right]^T, \quad (3.36)$$

onde $\hat{y}^{(k)}$ é obtido a partir da equação (3.34) e os parâmetros $\hat{\boldsymbol{\Theta}}_{MQ}$ são obtidos a partir do método dos mínimos quadrados não-recursivo ou recursivo.

Uma outra forma de escolher a variável instrumental é utilizar a entrada $\mathbf{u}^{(k)}$ como próprio instrumento (YOUNG, 1984; YOUNG, 2011), ou seja:

$$\tilde{\boldsymbol{\psi}}^{(k-1)} = \left[u^{(k-1)} \ u^{(k-2)} \ u^{(k-3)} \dots u^{(k-n_u-n_y)} \right]^T \quad (3.37)$$

Outra sugestão de uso de variável instrumental, apresentada em (YOUNG, 2011; YOUNG, 2015), é utilizar a entrada $\mathbf{u}^{(k)}$ atrasada em conjunto com a própria entrada não atrasada, dada por:

$$\tilde{\boldsymbol{\psi}}^{(k-1)} = \left[-\vartheta^{(k-1)} \dots -\vartheta^{(k-n)} \ u^{(k)} \dots u^{(k-n)} \right]^T \quad (3.38)$$

e

$$\hat{\boldsymbol{\vartheta}}^T = \left[-u^{(k-\tau)} \dots -u^{(k-\tau-n)} \ u^{(k)} \dots u^{(k-n)} \right], \quad (3.39)$$

onde τ é o atraso de tempo utilizado para deslocar o sinal de entrada em um determinado instante k em relação aos demais instantes.

A utilização de um vetor de variáveis instrumentais pode comprometer o desempenho da estimação paramétrica e também deve ser considerado que não há uma única variável instrumental para solução dos problemas de estimação paramétrica baseados no método dos mínimos quadrados não-recursivo ou recursivo.

Também é observado que a variável instrumental deve ser, idealmente pouco correlacionada com a perturbação do sistema e bastante correlacionada com os regressores do modelo utilizados no sistema dinâmico. Atingido-se este objetivo obtém-se o melhor desempenho da estimação paramétrica baseada no método dos mínimos quadrado com variável instrumental (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015; AGUIRRE, 2015).

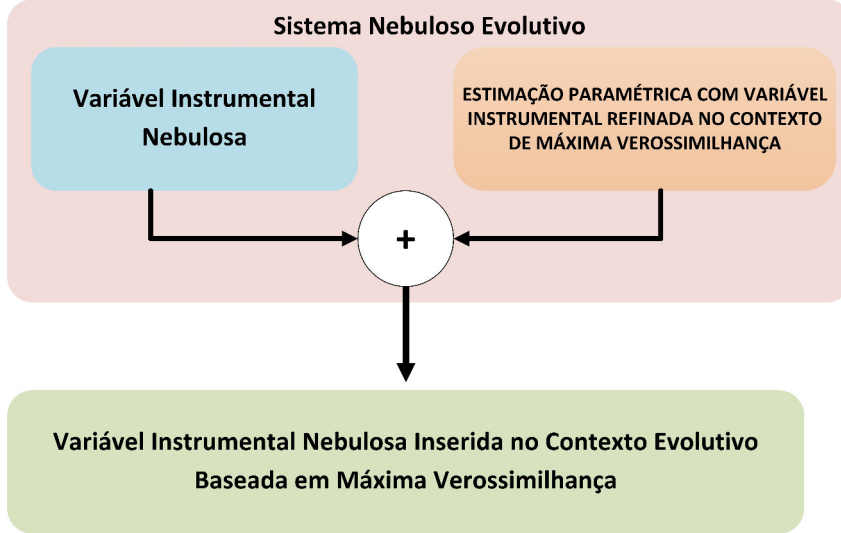
3.3 Variável Instrumental Nebulosa inserida no Contexto Evolutivo

Nesta tese, para superar a influência do ruído autocorrelacionado, possivelmente, presente nos dados experimentais, adota-se uma variável instrumental estendida que seja eficiente quanto a garantia de remoção da polarização na estimação paramétrica do consequente (GOODWIN; PAYNE, 1977; SODERSTROM; STOICA, 1989; LJUNG, 1999; YOUNG, 2011; WANG; GARNIER, 2012; YOUNG, 2015; AGUIRRE, 2015). A variável instrumental nebulosa utilizada, inserida no contexto evolutivo, foi obtida a partir da composição da estimação paramétrica com variável instrumental refinada no contexto de máxima verossimilhança (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015) e variável instrumental nebulosa (SERRA; BOTTURA, 2007; SERRA, 2012), conforme ilustrado no diagrama de blocos da Figura 22.

De acordo com (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015), tem-se uma abordagem da variável instrumental ótima refinada estendida no contexto da máxima verossimilhança. Esta metodologia proporciona uma estimação paramétrica recursiva dos parâmetros do consequente em um ambiente cujos dados podem estar corrompidos por ruído, sem a necessidade de conhecimento prévio do modelo de ruído adotado.

De acordo com (SERRA; BOTTURA, 2007; SERRA, 2012), tem-se o conceito da variável instrumental nebulosa, com aplicações ao processo de identificação, recursivo ou não-recursivo, associada a uma estrutura do tipo ARX nebulosa *Takagi-Sugeno*. Esta metodologia é estatisticamente não correlacionada com o ruído do sistema, a fim de minimizar o problema da estimação paramétrica do consequente em um ambiente ruidoso.

Figura 22: Variável instrumental nebulosa inserida no contexto de máxima verossimilhança aplicada em um sistema nebuloso evolutivo, como uma combinação das metodologias apresentadas em (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015) com o modelo de variável instrumental nebulosa apresentado em (SERRA; BOTTURA, 2007; SERRA, 2012).



Fonte: Elaborado pelo autor.

Os principais componentes que compõem a variável instrumental nebulosa inserida no contexto evolutivo baseada em máxima verossimilhança são ilustrados no diagrama de bloco da Figura 23, os componentes principais são: modelo nebuloso auxiliar, pré-filtragem, matriz de pesos nebulosos e submodelo algoritmo de identificação.

Inicialmente, sejam os vetores de entrada e saída com erros nas variáveis.

$$\mathbf{u}_e^{(k)} = \mathbf{u}^{(k)} + \boldsymbol{\xi}^{(k)} \quad k = 1, 2, \dots, N \quad (3.40)$$

e

$$\mathbf{y}_e^{(k)} = \mathbf{y}^{(k)} + \boldsymbol{\xi}^{(k)} \quad k = 1, 2, \dots, N. \quad (3.41)$$

O vetor de regressores de dados do sistemas dinâmico e o vetor de regressores com erros nas variáveis, $\boldsymbol{\psi}^{(k)}$ e $\boldsymbol{\psi}_e^{(k)}$, respectivamente, são expressos por

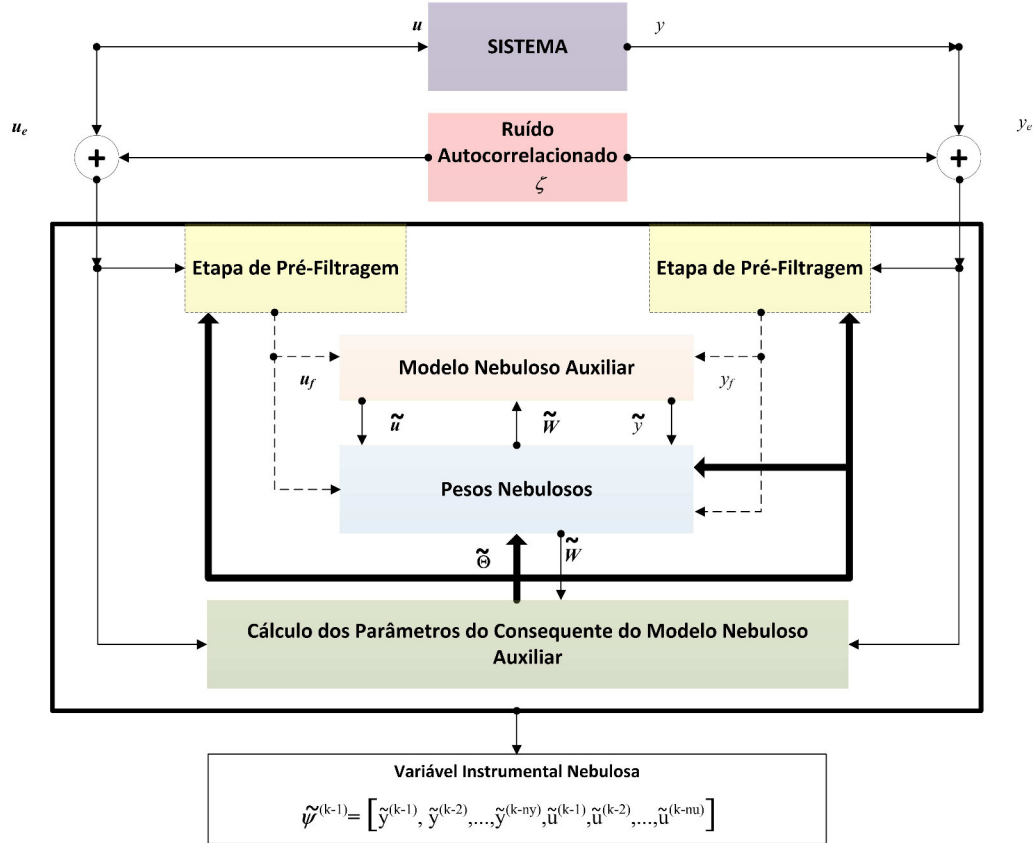
$$\boldsymbol{\psi}^{(k-1)} = \left[y^{(k-1)} \quad \dots \quad y^{(k-n_y)} \quad u^{(k-1)} \quad \dots \quad u^{(k-n_u)} \right]^T \forall k \quad (3.42)$$

e

$$\boldsymbol{\psi}_e^{(k-1)} = \left[y_e^{(k-1)} \quad \dots \quad y_e^{(k-n_y)} \quad u_e^{(k-1)} \quad \dots \quad u_e^{(k-n_u)} \right]^T \forall k. \quad (3.43)$$

Além disso, $\boldsymbol{\xi}^{(k)}$ representa a ruído autocorrelacionado no instante k .

Figura 23: Diagrama de bloco da variável instrumental nebulosa inserida no contexto evolutivo com seus principais componentes: etapa de pré-filtragem, modelo nebuloso auxiliar, pesos nebulosos estimação paramétrica do consequente do modelo auxiliar nebuloso.



Fonte: Elaborado pelo autor.

3.3.1 Modelo Nebuloso Auxiliar

A variável instrumental nebulosa $\tilde{\mathbf{y}}^{(k)}$ correlacionada com a saída $\mathbf{y}_e^{(k)}$ e decorrelacionada com ruído autocorrelacionado do sistema $\boldsymbol{\xi}^{(k)}$ é obtida a partir da saída filtrada de um sistema dinâmico identificado como modelo nebuloso auxiliar *Takagi-Sugeno*. A fórmula de inferência do modelo nebuloso auxiliar *Takagi-Sugeno* inserido em uma estrutura do tipo ARX é dada por

$$\begin{aligned} \tilde{\mathbf{y}}^{(k)} = & \gamma_{(1)} \left(\boldsymbol{\psi}_f^{(k-1)} \right) \left[\tilde{a}_{y_{(1)}}^{(1)} y_f^{(k-1)} + \dots + \tilde{a}_{y_{(1)}}^{(n_y)} y_f^{(k-n_y)} + \tilde{b}_{y_{(1)}}^{(1)} u_f^{(k-1)} + \dots + \tilde{b}_{y_{(1)}}^{(n_u)} u_f^{(k-n_u)} \right] \\ & + \gamma_{(2)} \left(\boldsymbol{\psi}_f^{(k-1)} \right) \left[\tilde{a}_{y_{(2)}}^{(1)} y_f^{(k-1)} + \dots + \tilde{a}_{y_{(2)}}^{(n_y)} y_f^{(k-n_y)} + \tilde{b}_{y_{(2)}}^{(1)} u_f^{(k-1)} + \dots + \tilde{b}_{y_{(2)}}^{(n_u)} u_f^{(k-n_u)} \right] \\ & + \dots + \gamma_{(c)} \left(\boldsymbol{\psi}_f^{(k-1)} \right) \left[\tilde{a}_{y_{(c)}}^{(1)} y_f^{(k-1)} + \dots + \tilde{a}_{y_{(c)}}^{(n_y)} y_f^{(k-n_y)} + \tilde{b}_{y_{(c)}}^{(1)} u_f^{(k-1)} + \dots + \tilde{b}_{y_{(c)}}^{(n_u)} u_f^{(k-n_u)} \right]. \quad (3.44) \end{aligned}$$

A equação (3.44) é reescrita da seguinte forma,

$$\tilde{y}^{(k)} = \sum_{i=1}^c \left\{ \gamma_{(i)} \left(\boldsymbol{\psi}_f^{(k-1)} \right) \left[\left(\boldsymbol{\psi}_f^{(k-1)} \right)^T \tilde{\boldsymbol{\Theta}}_y^{(i)} \right] \right\} \quad (3.45)$$

onde

$$\tilde{\boldsymbol{\Theta}}_y^{(i)} = \left[\tilde{a}_{y_{(i)}}^{(1)} \quad \dots \quad \tilde{a}_{y_{(i)}}^{(n_y)} \quad \tilde{b}_{y_{(i)}}^{(1)} \quad \dots \quad \tilde{b}_{y_{(i)}}^{(n_u)} \right] \quad (3.46)$$

para

$$\boldsymbol{\psi}_f^{(k)} = \left[y_f^{(k-1)} \quad \dots \quad y_f^{(k-n_y)} \quad u_f^{(k-1)} \quad \dots \quad u_f^{(k-n_u)} \right], \quad (3.47)$$

com o grau de ativação da i ésima regra, $\gamma_{(i)} \left(\boldsymbol{\psi}_f^{(k-1)} \right)$, para $i = 1, 2, \dots, c$, obtido a partir da equação (2.39); $\tilde{\boldsymbol{\Theta}}_y^{(i)}$ são os parâmetros do consequente utilizados no modelo nebuloso auxiliar para determinar a variável $\tilde{y}^{(k)}$; $\tilde{y}^{(k)}$ é a saída filtrada do modelo nebuloso auxiliar no instante k ; $\boldsymbol{\psi}_f^{(k-1)} = \left[y_f^{(k-1)} \quad \dots \quad y_f^{(k-n_y)} \quad u_f^{(k-1)} \quad \dots \quad u_f^{(k-n_u)} \right]$ é o vetor de regressores do antecedente obtido a partir da etapa de pré-filtragem, com n_y e n_u as dimensionalidades das variáveis de entrada e saída, respectivamente; c é o total de regras nebulosas identificadas no instante $k = 1, 2, \dots, N$.

A equação (3.45), a qual é linear com os parâmetros do consequente (3.46), é reescrita na forma vetorial como

$$\tilde{\boldsymbol{y}}^{(k)} = \left[\tilde{y}^{(k-1)} \quad \tilde{y}^{(k-2)} \quad \dots \quad \tilde{y}^{(k-n_y)} \right]^T. \quad (3.48)$$

A partir de um conjunto de N pares de dados de entrada e saída $\left\{ \left(\boldsymbol{\psi}_f^{(k)}, y_f^k \right) \right\}$ disponível, a seguinte forma vetorial é expressa por,

$$\tilde{\boldsymbol{Y}} = \left[\boldsymbol{\Gamma}_1 \boldsymbol{\Psi}_f \quad \boldsymbol{\Gamma}_2 \boldsymbol{\Psi}_f \quad \dots \quad \boldsymbol{\Gamma}_c \boldsymbol{\Psi}_f \right] \tilde{\boldsymbol{\Theta}}_y \quad (3.49)$$

onde $\boldsymbol{\Psi}_f = \left[\boldsymbol{y}_f^{(k)} \quad \dots \quad \boldsymbol{y}_f^{(k-n_y)} \quad \boldsymbol{u}_f^{(k)} \quad \dots \quad \boldsymbol{u}_f^{(k-n_u)} \right] \in \mathbb{R}^{N \times (n_y + n_u)}$, $\boldsymbol{\Gamma}_i = \text{diag} \left(\gamma_{(i)} \left(\boldsymbol{\psi}_f^{(k)} \right) \right) \in \mathbb{R}^{N \times N}$, $\tilde{\boldsymbol{Y}} \in \mathbb{R}^{N \times 1}$ e $\tilde{\boldsymbol{\Theta}}_y \in \mathbb{R}^{c(n_y + n_u) \times 1}$, são a matriz de regressores, a matriz de grau de ativação normalizado, o vetor de saída filtrada do modelo nebuloso auxiliar e os parâmetros do consequente do modelo auxiliar, respectivamente.

A variável instrumental nebulosa $\tilde{\boldsymbol{u}}^{(k)}$ correlacionada com a entrada $\boldsymbol{u}_e^{(k)}$ e descorrelacionada com ruído autocorrelacionado do sistema $\boldsymbol{\xi}^{(k)}$ é obtida a partir da saída filtrada de um sistema dinâmico identificado como modelo nebuloso auxiliar *Takagi–Sugeno*. A fórmula de inferência do modelo nebuloso auxiliar *Takagi–Sugeno* inserido em uma estrutura do tipo ARX é dada por

$$\begin{aligned} \tilde{u}^{(k)} &= \gamma_{(1)} \left(\boldsymbol{u}_f^{(k-1)} \right) \left[\tilde{b}_{u_{(1)}}^{(1)} u_f^{(k-1)} + \dots + \tilde{b}_{u_{(1)}}^{(n_u)} u_f^{(k-n_u)} \right] \\ &+ \gamma_{(2)} \left(\boldsymbol{u}_f^{(k-1)} \right) \left[\tilde{b}_{u_{(2)}}^{(1)} u_f^{(k-1)} + \dots + \tilde{b}_{u_{(2)}}^{(n_u)} u_f^{(k-n_u)} \right] \\ &+ \dots + \gamma_{(c)} \left(\boldsymbol{u}_f^{(k-1)} \right) \left[\tilde{b}_{u_{(c)}}^{(1)} u_f^{(k-1)} + \dots + \tilde{b}_{u_{(c)}}^{(n_u)} u_f^{(k-n_u)} \right]. \end{aligned} \quad (3.50)$$

A equaco (3.50)  reescrita da seguinte forma,

$$\tilde{u}^{(k)} = \sum_{i=1}^c \left\{ \gamma_{(i)} \left(\mathbf{u}_f^{(k-1)} \right) \left[\left(\mathbf{u}_f^{(k-1)} \right)^T \tilde{\Theta}_u^{(i)} \right] \right\}, \quad (3.51)$$

onde

$$\tilde{\Theta}_u^{(i)} = \left[\tilde{b}_{u_{(i)}}^{(1)} \quad \dots \quad \tilde{b}_{u_{(i)}}^{(n_u)} \right], \quad (3.52)$$

de modo que

$$\mathbf{u}_f^{(k)} = \left[u_f^{(k-1)} \quad \dots \quad u_f^{(k-n_u)} \right], \quad (3.53)$$

com o grau de ativaco da i sima regra, $\gamma_{(i)} \left(\mathbf{u}_f^{(k-1)} \right)$, para $c = 1, 2, \dots, c$, obtido a partir da equaco (2.39); $\tilde{\Theta}_u^{(i)}$ so os parmetros do conseqente utilizados no modelo nebuloso auxiliar para determinar a varivel $\tilde{u}^{(k)}$; $\tilde{u}^{(k)}$ representa a entrada filtrada do modelo nebuloso auxiliar no instante k ; $\mathbf{u}_f^{(k-1)} = \left[u_f^{(k-1)} \quad u_f^{(k-2)} \quad \dots \quad u_f^{(k-n_u)} \right]$  o vetor de entrada obtido a partir da etapa de pr-filtragem, com n_u a dimensionalidade da varivel de entrada, respectivamente; c  o total de regras nebulosas identificadas no instante $k = 1, 2, \dots, N$.

A equaco (3.51), linear nos parmetros conseqentes (3.52),  reescrita na forma vetorial como

$$\tilde{\mathbf{u}}^{(k)} = \left[\tilde{u}^{(k-1)} \quad \tilde{u}^{(k-2)} \quad \dots \quad \tilde{u}^{(k-n_u)} \right]^T \quad (3.54)$$

A partir de um conjunto de N amostras da entrada $\left\{ \mathbf{u}_f^{(k)} \right\}$ disponvel, a seguinte forma vetorial  expressa por,

$$\tilde{\mathbf{U}} = \left[\mathbf{\Gamma}_1 \mathbf{U}_f \quad \mathbf{\Gamma}_2 \mathbf{U}_f \quad \dots \quad \mathbf{\Gamma}_c \mathbf{U}_f \right] \tilde{\Theta}_u, \quad (3.55)$$

onde $\mathbf{U}_f = \left[\mathbf{u}_f^{(k)} \quad \dots \quad \mathbf{u}_f^{(k-n_u)} \right] \in \mathbb{R}^{N \times (n_u)}$, $\mathbf{\Gamma}_i = \text{diag} \left(\gamma_{(i)} \left(\mathbf{u}_f^{(k)} \right) \right) \in \mathbb{R}^{N \times N}$, $\tilde{\mathbf{U}} \in \mathbb{R}^{N \times 1}$ e $\tilde{\Theta}_u \in \mathbb{R}^{c(n_u) \times 1}$, so a matriz de regressores, a matriz de grau de ativaco normalizado, o vetor de entrada filtrada do modelo nebuloso auxiliar e os parmetros do conseqente do modelo auxiliar, respectivamente.

3.3.2 Modelo de Pr-Filtragem

A funo $\mathbf{F}(z^{-1})$ da varivel instrumental estendida (SODERSTROM; STOICA, 1989; SODERSTROM; STOICA, 1983) no modelo de varivel instrumental utilizado nesta tese ser representado um filtro de mdia mvel recursivo de quatro estgios utilizado na etapa de pr-filtragem deste trabalho conforme (SMITH, 2002). O uso do modelo de varivel instrumental nebulosa  capaz de garantir uma polarizaco assinttica da estimaco dos parmetros do conseqente; porm, para problemas onde h erro na varivel externa do modelo dinmico, verifica-se a necessidade de uma pr-filtragem desta, como uma forma de otimizaco do modelo nebuloso auxiliar e garantir uma polarizaco assinttica da estimaco paramtrica dos parmetros do conseqente.

As equações utilizadas no filtro de média móvel de quatro estágios para os sinais $y_f^{(k)}$ e $u_f^{(k)}$ são expressas por

$$y_f^{(k)} = (1 - \omega_y)^4 y^{(k)} + 4\omega_y y_f^{(k-1)} - 6\omega_y^2 y_f^{(k-2)} + 4\omega_y^3 y_f^{(k-3)} - \omega_y^4 y_f^{(k-4)} \quad (3.56)$$

e

$$u_f^{(k)} = (1 - \omega_u)^4 u^{(k)} + 4\omega_u u_f^{(k-1)} - 6\omega_u^2 u_f^{(k-2)} + 4\omega_u^3 u_f^{(k-3)} - \omega_u^4 u_f^{(k-4)}, \quad (3.57)$$

onde $\omega_u \in (0, 1)$ e $\omega_y \in (0, 1)$. De acordo com (SMITH, 2002), a beleza deste método recursivo nestes filtros está na capacidade de criar uma ampla variedade de filtragem, alterando-se um único parâmetro, ω_u ou ω_y . Na configuração de pré-filtragem o especialista tem a liberdade de desativar o processo de filtragem fazendo $\omega_y = 0$ ou $\omega_u = 0$, considerando que a escolha dos instrumentos que irão compor a variável instrumental nebulosa é derivado da experiência do especialista diante do problema onde o modelo está inserido.

3.3.3 Matriz de Pesos Nebulosos

A matriz de pesos nebulosos da estimação paramétrica do modelo auxiliar é obtida a partir da matriz de partição do agrupamento nebuloso realizado nas amostras. Ela pode ser realizada de forma não-recursiva (batelada) na estimação inicial, e de forma recursiva na estimação evolutiva.

A formulação da matriz de partição nebulosa obtida a partir do conjunto de amostras é expressa por

$$\mathbf{U}_\mu \in \mathbb{R}^{c \times N} \left| \mu_i^{(k)} \in [0, 1], \forall i, k; \sum_{i=1}^c \mu_i^{(k)} = 1, \forall k; 0 < \sum_{k=1}^N \mu_i^{(k)} < N, \forall i. \right. \quad (3.58)$$

Seja o conjunto de N pares de dados de entrada e saída $\{(\boldsymbol{\psi}_f^{(k)}, y_f^k)\}$, o grau de pertinência de cada amostra é definido como

$$\mu_i^{(k)} = \mu_i^{(k)} \left(\{(\boldsymbol{\psi}_f^{(k)}, y_f^k)\} \right) \forall k \quad (3.59)$$

e a matriz de pesos nebulosos para a amostra $\{(\boldsymbol{\psi}_f^{(k)}, y_f^k)\}$ é definida como se segue

$$\widetilde{\mathbf{W}}_\mu^{(i)} = \text{diag}(\mu_i^{(k)}) \in \mathbb{R}^{N \times N}, \quad (3.60)$$

onde

$$\widetilde{\mathbf{W}}_\mu^{(i)} = \begin{pmatrix} \tilde{w}_i^{(1)} = \mu_i^{(1)} & 0 & \dots & 0 \\ 0 & \tilde{w}_i^{(2)} = \mu_i^{(2)} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \tilde{w}_i^{(N)} = \mu_i^{(N)} \end{pmatrix}. \quad (3.61)$$

Para um conjunto de N amostras da entrada $\{\mathbf{u}_f^{(k)}\}$, o grau de pertinência de cada amostra de entrada é definido como

$$\mu_i^{(k)} = \mu_i^{(k)} \left(\{\mathbf{u}_f^{(k)}\} \right) \forall k \quad (3.62)$$

e a matriz de pesos nebulosos para a amostra $\{\mathbf{u}_f^{(k)}\}$ é definida como se segue

$$\widehat{W}_\mu^{(i)} = \text{diag}(\mu_i^{(k)}) \in \mathbb{R}^{N \times N} \quad (3.63)$$

onde

$$\widehat{W}_\mu^{(i)} = \begin{pmatrix} \hat{w}_i^{(1)} = \mu_i^{(1)} & 0 & \dots & 0 \\ 0 & \hat{w}_i^{(2)} = \mu_i^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{w}_i^{(N)} = \mu_i^{(N)} \end{pmatrix} \quad (3.64)$$

3.3.4 Parâmetros do Consequente do Modelo Nebuloso Auxiliar

Os vetores de parâmetros do consequente, $\tilde{\Theta}_u$ e $\tilde{\Theta}_y$, do modelo nebuloso auxiliar podem ser obtidos de forma não-recursiva quando em estimação inicial, e de forma recursiva na estimação evolutiva.

3.3.4.1 Esquema Não-Recursivo (Batelada)

Adota-se a abordagem local neste procedimento e os parâmetros do consequente auxiliares são estimados para cada regra i , independentemente, minimizando um conjunto de critérios locais ponderados para $i = 1, 2, \dots, c$:

$$\tilde{\Theta}_y^{(i)} = \left(\Psi_f^T \widehat{W}_\mu^{(i)} \Psi_f \right)^{-1} \Psi_f^T \widehat{W}_\mu^{(i)} y_f, \quad (3.65)$$

$$\tilde{\Theta}_u^{(i)} = \left(U_f^T \widehat{W}_\mu^{(i)} U_f \right)^{-1} U_f^T \widehat{W}_\mu^{(i)} u_f^{(k)}, \quad (3.66)$$

onde

$$\Psi_f = \begin{bmatrix} y_f^{(1)} & \dots & y_f^{(n_y)} & u_f^{(1)} & \dots & u_f^{(n_u)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_f^{(N-n_y)} & \dots & y_f^{(N-1)} & u_f^{(N-n_u)} & \dots & u_f^{(N-1)} \end{bmatrix} \quad (3.67)$$

para

$$U_f = \begin{bmatrix} u_f^{(1)} & \dots & u_f^{(n_u)} \\ \vdots & \vdots & \vdots \\ u_f^{(N-n_u)} & \dots & u_f^{(N-1)} \end{bmatrix}. \quad (3.68)$$

No esquema não-recursivo para obtenção dos parâmetros do consequente dos modelos auxiliares, utiliza-se a forma ponderada. A matriz de pesos nebulosos obtida a partir do conjunto de dados é utilizada na matriz de pesos, o que possibilita um refinamento do cálculo destes parâmetros, ao contrário dos métodos não ponderados.

3.3.4.2 Esquema Recursivo

Neste procedimento, os parâmetros do esquema recursivo são atualizados a cada nova iteração k . O ajuste de regras por parte do modelo nebuloso evolutivo permite as

alteraces nos parmetros do conseqente  medida que o sistema nebuloso modifica o nmero de regras nebulosas e garante um refinamento mais preciso na determinaco destes.

A atualizaco dos parmetros do conseqente $\tilde{\Theta}_y^{(k)(i)}$ para $i = 1, 2, \dots, c$, e $k = 1, 2, \dots, N$,  expressa como

$$\tilde{\Theta}_y^{(k)(i)} = \tilde{\Theta}_y^{(k-1)(i)} + \mathbf{K}_{\tilde{\Theta}_y}^{(k)} \left(y_f^{(k)} - \boldsymbol{\psi}_f^{(k)T} \tilde{\Theta}_y^{(k-1)(i)} \right), \quad (3.69)$$

onde o ganho de adaptaco do vetor estimado, referido como ganho de *Kalman*, 

$$\mathbf{K}_{\tilde{\Theta}_y}^{(k)} = \frac{\tilde{w}_i^{(k)} \mathbf{P}_{\tilde{\Theta}_y}^{(k-1)} \boldsymbol{\psi}_f^{(k)}}{\left(\delta_{\tilde{\Theta}_y}^{(k)} + \tilde{w}_i^{(k)} \boldsymbol{\psi}_f^{(k)T} \mathbf{P}_{\tilde{\Theta}_y}^{(k-1)} \boldsymbol{\psi}_f^{(k)} \right)}. \quad (3.70)$$

A matriz de covarincia  expressa por

$$\mathbf{P}_{\tilde{\Theta}_y}^{(k)} = \frac{1}{\delta_{\tilde{\Theta}_y}^{(k)}} \left[\mathbf{P}_{\tilde{\Theta}_y}^{(k-1)} - \mathbf{K}_{\tilde{\Theta}_y}^{(k)} \boldsymbol{\psi}_f^{(k)T} \mathbf{P}_{\tilde{\Theta}_y}^{(k-1)} \right] \quad (3.71)$$

onde $\tilde{w}_i^{(k)}$  a atualizaco dos pesos nebulosos a cada nova iterao k para a i sima regra do modelo evolutivo, $i = 1, 2, \dots, c$. O fator de esquecimento atualizado a cada nova iterao do algoritmo recursivo  expresso por

$$\delta_{\tilde{\Theta}_y}^{(k)} = \lambda_0 \delta_{\tilde{\Theta}_y}^{(k-1)} + (1 - \lambda_0) (\delta_0) - \sigma_{\tilde{\Theta}_y}^{(k-1)} \quad (3.72)$$

com $i = 1, 2, \dots, c$, e $k = 1, 2, \dots, N$. Os parmetros adicionais δ_0 , λ_0 e $\delta_{\tilde{\Theta}_u}^{(0)}$ so escolhidos pelo especialista. Os valores tpicos so: $\delta_0 = 0,95$, $\lambda_0 = 0,99$ e $\delta_{\tilde{\Theta}_y}^{(0)} \in [0, 1]$ (YOUNG, 1984; YOUNG, 2011). A varincia mvel do erro de estimaco a cada nova iterao  expressa por

$$\sigma_{\tilde{\Theta}_y}^{(k)} = \sigma_{\tilde{\Theta}_y}^{(k-1)} - \frac{1}{k} \left[\sigma_{\tilde{\Theta}_y}^{(k-1)} - \left(y_f^{(k)} - \tilde{y}^{(k)} - \widetilde{me}_{\tilde{\Theta}_y}^{(k)} \right)^2 \right] \quad (3.73)$$

e

$$\widetilde{me}_{\tilde{\Theta}_y}^{(k)} = \frac{1}{k} \sum_{j=1}^k \left(y_f^{(j)} - \tilde{y}^{(j)} \right) \quad (3.74)$$

onde, $y^{(k)}$ e $\tilde{y}^{(k)}$, so as sadas real e estimada do modelo no instante k , respectivamente.

A atualizaco dos parmetros do conseqente $\tilde{\Theta}_u^{(k)(i)}$ para $i = 1, 2, \dots, c$, e $k = 1, 2, \dots, N$,  expressa como

$$\tilde{\Theta}_u^{(k)(i)} = \tilde{\Theta}_u^{(k-1)(i)} + \mathbf{K}_{\tilde{\Theta}_u}^{(k)} \left(u_f^{(k)} - \mathbf{u}_f^{(k)T} \tilde{\Theta}_u^{(k-1)(i)} \right) \quad (3.75)$$

onde o ganho de adaptaco do vetor estimado, referido como ganho de *Kalman*, 

$$\mathbf{K}_{\tilde{\Theta}_u}^{(k)} = \frac{\hat{w}_i^{(k)} \mathbf{P}_{\tilde{\Theta}_u}^{(k-1)} \mathbf{u}_f^{(k)}}{\left(\delta_{\tilde{\Theta}_u}^{(k)} + \hat{w}_i^{(k)} \mathbf{u}_f^{(k)T} \mathbf{P}_{\tilde{\Theta}_u}^{(k-1)} \mathbf{u}_f^{(k)} \right)} \quad (3.76)$$

A matriz de covariância é expressa por

$$\mathbf{P}_{\tilde{\Theta}_u}^{(k)} = \frac{1}{\delta_{\tilde{\Theta}_u}^{(k)}} \left[\mathbf{P}_{\tilde{\Theta}_u}^{(k-1)} - \mathbf{K}_{\tilde{\Theta}_u}^{(k)} \mathbf{u}_f^{(k)T} \mathbf{P}_{\tilde{\Theta}_u}^{(k-1)} \right]. \quad (3.77)$$

onde $\hat{w}_i^{(k)}$ é a atualização dos pesos nebulosos a cada nova iteração, para i ésima regra do modelo evolutivo para $i = 1, 2, \dots, c$.

O fator de esquecimento atualizado a cada nova iteração do algoritmo recursivo é dado por

$$\delta_{\tilde{\Theta}_u}^{(k)} = \lambda_0 \delta_{\tilde{\Theta}_u}^{(k-1)} + (1 - \lambda_0)(\delta_0) - \sigma_{\tilde{\Theta}_u}^{(k-1)} \quad (3.78)$$

com $i = 1, 2, \dots, c$ e $k = 1, 2, \dots, N$. Os parâmetros adicionais δ_0 , λ_0 e $\delta_{\tilde{\Theta}_u}^{(0)}$ são escolhidos pelo especialista, cujos valores típicos são: $\delta_0 = 0,95$, $\lambda_0 = 0,99$ e $\delta_{\tilde{\Theta}_u}^{(0)} \in [0, 1]$ (YOUNG, 2011).

A variância móvel do erro de estimação a cada nova iteração é expresso por

$$\sigma_{\tilde{\Theta}_u}^{(k)} = \sigma_{\tilde{\Theta}_u}^{(k-1)} - \frac{1}{k} \left[\sigma_{\tilde{\Theta}_u}^{(k-1)} - \left(u_f^{(k)} - \tilde{u}^{(k)} - \widetilde{me}_{\tilde{\Theta}_u}^{(k)} \right)^2 \right] \quad (3.79)$$

sendo

$$\widetilde{me}_{\tilde{\Theta}_u}^{(k)} = \frac{1}{k} \sum_{j=1}^k \left(u_f^{(j)} - \tilde{u}^{(j)} \right). \quad (3.80)$$

onde, $u^{(k)}$ e $\tilde{u}^{(k)}$, são as saídas real e estimada do modelo no instante k , respectivamente.

As matrizes dos parâmetros do conseqüente adotadas no modelo auxiliar são expressas por,

$$\tilde{\Theta}_y = \left[\tilde{\Theta}_y^{(1)} \quad \tilde{\Theta}_y^{(2)} \quad \dots \quad \tilde{\Theta}_y^{(c)} \right]^T \quad (3.81)$$

e

$$\tilde{\Theta}_u = \left[\tilde{\Theta}_u^{(1)} \quad \tilde{\Theta}_u^{(2)} \quad \dots \quad \tilde{\Theta}_u^{(c)} \right]^T. \quad (3.82)$$

3.3.5 Composição da Variável Instrumental Nebulosa

A variável instrumental nebulosa inserida em um contexto de máxima verossimilhança para estimação paramétrica dos modelos locais, é expressa da seguinte forma,

$$\tilde{\psi}^{(k-1)} = \left[\tilde{y}^{(k-1)} \quad \tilde{y}^{(k-2)} \quad \dots \quad \tilde{y}^{(k-n_y)} \quad \tilde{u}^{(k-1)} \quad \tilde{u}^{(k-2)} \quad \dots \quad \tilde{u}^{(k-n_u)} \right]^T, \quad (3.83)$$

onde as variáveis $\tilde{\mathbf{u}}^{(k)}$ e $\tilde{\mathbf{y}}^{(k)}$ são obtidas a partir das equações (3.50) e (3.44).

Utilizando a identificação paramétrica não-recursiva a partir das equações (3.66) e (3.65), para as variáveis $\tilde{\mathbf{u}}^{(k)}$ e $\tilde{\mathbf{y}}^{(k)}$, respectivamente; para refinamento com o esquema de identificação paramétrica recursiva inserida no contexto da máxima verossimilhança, utiliza-se as equações (3.75) e (3.69), respectivamente.

Outra abordagem simplificada da variável instrumental nebulosa da equaço (3.83), é exposto da seguinte forma:

$$\tilde{\boldsymbol{\psi}}^{(k-1)} = \left[\tilde{y}^{(k-1)} \tilde{y}^{(k-2)} \dots \tilde{y}^{(k-n_y)} u_f^{(k-1)} u_f^{(k-2)} \dots u_f^{(k-n_u)} \right]^T, \quad (3.84)$$

onde as variáveis $\mathbf{u}_f^{(k)}$ e $\tilde{\mathbf{y}}^{(k)}$ são obtidas a partir das equaçes (3.57) e (3.44), respectivamente.

Utilizando a identificaço paramétrica não-recursiva a partir da equaço (3.65), para a variável $\tilde{\mathbf{y}}^{(k)}$; para refinamento com o esquema de identificaço paramétrica recursiva inserida no contexto da máxima verossimilhança, utiliza-se a equaço (3.69), respectivamente.

A simplificaço permite ao especialista escolher uma variável instrumental nebulosa em que as variáveis de saída são obtidas a partir do modelo nebuloso auxiliar e as variáveis de entradas são obtidas a partir do modelo de pré-filtragem e pode ser adequada a situaçes onde não há erro nas variáveis de entrada do modelo. Portanto não há comprometimento do desempenho final do algoritmo de identificaço paramétrica.

A variável instrumental nebulosa escrita na forma matricial é dada por

$$\tilde{\boldsymbol{\Psi}} = \left[\tilde{\boldsymbol{\psi}}^{(1)} \tilde{\boldsymbol{\psi}}^{(2)} \tilde{\boldsymbol{\psi}}^{(3)} \dots \tilde{\boldsymbol{\psi}}^{(N)} \right]^T. \quad (3.85)$$

3.4 Análise de Convergência da Variável Instrumental Nebulosa

Em (SERRA; BOTTURA, 2005; SERRA; BOTTURA, 2007; SERRA; BOTTURA, 2009) são apresentados Lemas e Teoremas a respeito da variável instrumental nebulosa e nesta tese será utilizado estes Lemas e Teoremas na variável instrumental nebulosa inserida no contexto evolutivo para fins de análise de convergência. Assim, garante a existência de uma soluço no método dos mínimos quadrados não-recursivo e/ou recursivo na presença de dados corrompidos por ruído autocorrelacionado. Considerando (SERRA; BOTTURA, 2009), formula-se as equaçes normais, no instante de tempo k , como

$$\sum_{j=1}^k \left\{ \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \dots \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right] \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \dots \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \right\} \hat{\boldsymbol{\Theta}}^{(k)} = \dots \\ \dots \sum_{j=1}^k \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \dots \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right] y^{(j)}, \quad (3.86)$$

onde $\gamma_{(i)}^{(j)}$ é o grau de ativaço normalizado, obtido na equaço (2.39), associado a $\boldsymbol{\psi}_e^{(j)}$. Considere $\boldsymbol{\psi}_e^{(j)}$ um vetor com erro na variável, $\boldsymbol{\psi}_e^{(j)} = \boldsymbol{\psi}^{(j)} + \boldsymbol{\xi}^{(j)}$; $\boldsymbol{\xi}^{(j)}$ é a perturbaço do sistema, para $i = 1, 2, \dots, c$; $j = 1, 2, \dots, k$ e $k = 1, 2, \dots, N$.

Seja a equaço do modelo nebuloso inserido na inferência *Takagi-Sugeno* exposto por

$$y^{(k)} = \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \dots \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \boldsymbol{\Theta} + \varepsilon^{(j)}, \quad (3.87)$$

onde $\varepsilon^{(j)}$ é o erro residual no instante j , para $j = 1, 2, \dots, k$, e $k = 1, 2, \dots, N$. Multiplicando-se ambos os lados da equação (3.86) por $\frac{1}{k}$ e substituindo-se na equação (3.87) obtém-se

$$\left\{ \frac{1}{k} \sum_{j=1}^k \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right] \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \right\} \hat{\boldsymbol{\Theta}}^{(k)} = \dots$$

$$\frac{1}{k} \sum_{j=1}^k \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right] \left\{ \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \boldsymbol{\Theta} + \varepsilon^{(j)} \right\}. \quad (3.88)$$

A estimação paramétrica com base na variável instrumental nebulosa neste contexto evolutivo, substituindo-se na equação (3.88), é dada por

$$\left\{ \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \right] \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \right\} \hat{\boldsymbol{\Theta}}^{(k)} = \dots$$

$$\frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \right] \left\{ \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \boldsymbol{\Theta} + \varepsilon^{(j)} \right\}, \quad (3.89)$$

onde $\tilde{\boldsymbol{\psi}}^{(j)}$ um vetor nebuloso de mesma ordem de $\boldsymbol{\psi}^{(j)}$, associado à dinâmica do sistema e descorrelacionado com o ruído de entrada $\boldsymbol{\xi}^{(j)}$; e $\gamma_{f(i)}^{(j)}$ é o grau de ativação normalizado, como obtido a partir da equação (2.39), associado a $\boldsymbol{\psi}_f^{(j)}$ para $i = 1, 2, \dots, c$; $j = 1, 2, \dots, k$, e $k = 1, 2, \dots, N$.

Para análise de convergência dos estimadores de mínimos quadrados com base na variável instrumental nebulosa, segue-se como proposto nos seguintes lemas e teorema.

3.4.1 Análise de Convergência - Lema 1

Lema 1 Considere $\tilde{\boldsymbol{\psi}}^{(j)}$ um vetor nebuloso de mesma ordem de $\boldsymbol{\psi}^{(j)}$, associado à dinâmica do sistema e descorrelacionado com o ruído de entrada $\boldsymbol{\xi}^{(j)}$; e $\gamma_{f(i)}^{(j)}$ é o grau de ativação normalizado como obtido a partir da equação (2.39), associado a $\boldsymbol{\psi}_f^{(j)}$ para $i = 1, 2, \dots, c$; $j = 1, 2, \dots, k$ e $k = 1, 2, \dots, N$.

Então, no limite, tem-se

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \right] \left(\boldsymbol{\xi}^{(j)} \right)^T = \mathbf{0}. \quad (3.90)$$

Prova. Desenvolvendo o lado esquerdo da equação (3.90), obtém-se,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \right]. \quad (3.91)$$

Como $\gamma_{f(i)}^{(j)}, i = 1, 2, \dots, c; j = 1, 2, \dots, k; k = 1, 2, \dots, N$ é um escalar, e, por definição, a escolha das variáveis são independentes do ruído autocorrelacionado na entrada, o produto vetorial entre $\tilde{\boldsymbol{\psi}}^{(j)}$ e $\boldsymbol{\xi}^{(j)}$ resultará em zero.

Consequentemente

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} (\boldsymbol{\xi}^{(j)})^T \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} (\boldsymbol{\xi}^{(j)})^T \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} (\boldsymbol{\xi}^{(j)})^T \right] = \mathbf{0} \quad \square \quad (3.92)$$

3.4.2 Análise de Convergência - Lema 2

Lema 2 Considere $\tilde{\boldsymbol{\psi}}^{(j)}$, um vetor nebuloso de mesma ordem de $\boldsymbol{\psi}^{(j)}$, associado à dinâmica do sistema e decorrelacionado do erro residual do sistema $\varepsilon^{(j)}$; e $\gamma_{f(i)}^{(j)}$ é o grau de ativação normalizado, obtido pela equação (2.39), associado a $\boldsymbol{\psi}_f^{(j)}$ para $i = 1, 2, \dots, c; j = 1, 2, \dots, k$ e $k = 1, 2, \dots, N$.

Então, no limite, tem-se

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \right] \varepsilon^{(j)} = \mathbf{0}. \quad (3.93)$$

Prova. Desenvolvendo o lado esquerdo da equação (3.93), obtém-se

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \varepsilon^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \varepsilon^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \varepsilon^{(j)} \right]. \quad (3.94)$$

Considerando a escolha das variáveis independentes do ruído, o produto vetorial entre $\tilde{\boldsymbol{\psi}}^{(j)}$ e $\varepsilon^{(j)}$ será zero no limite.

Consequentemente,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \varepsilon^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \varepsilon^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \varepsilon^{(j)} \right] = \mathbf{0} \quad \square \quad (3.95)$$

3.4.3 Análise de Convergência - Lema 3

Lema 3 Considere $\tilde{\boldsymbol{\psi}}^{(j)}$ um vetor nebuloso de mesma ordem $\boldsymbol{\psi}^{(j)}$, associado a dinâmica do sistema e decorrelacionado com o ruído de entrada e saída $\boldsymbol{\xi}^{(j)}$. Considere $\boldsymbol{\psi}_e^{(j)}$ um vetor com erro na variável, $\boldsymbol{\psi}_e^{(j)} = \boldsymbol{\psi}^{(j)} + \boldsymbol{\xi}^{(j)}$; os graus de ativação normalizados $\gamma_{f(i)}^{(j)}$ e $\gamma_{(i)}^{(j)}$, obtidos pela equação (2.39), associados a $\boldsymbol{\psi}_f^{(j)}$ e $\boldsymbol{\psi}_e^{(j)}$, respectivamente, para $i = 1, 2, \dots, c; j = 1, 2, \dots, k$, e $k = 1, 2, \dots, N$.

Então, de acordo com a equação (3.89), no limite, tem-se

$$\left\{ \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \right] \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \right\} =$$

$$\dots \times \left[\gamma_{(1)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \gamma_{(2)}^{(j)} \boldsymbol{\psi}_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \boldsymbol{\psi}_e^{(j)} \right]^T \} = \mathbf{C}_{\tilde{\boldsymbol{\psi}}\boldsymbol{\psi}_e} \neq \mathbf{0}. \quad (3.96)$$

Prova. Desenvolvendo-se o lado esquerdo da equação (3.96), obtém-se

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \gamma_{(1)}^{(j)} \left(\tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T + \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \right) \quad \gamma_{f(2)}^{(j)} \gamma_{(1)}^{(j)} \left(\tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T + \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \right) \quad \dots \right. \\ \left. \dots \quad \gamma_{f(c)}^{(j)} \gamma_{(c)}^{(j)} \left(\tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T + \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \right) \right]. \end{aligned}$$

A partir do Lema 1, a expressão é simplificada como

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \gamma_{(1)}^{(j)} \left(\tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T + \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \right) \quad \gamma_{f(2)}^{(j)} \gamma_{(2)}^{(j)} \left(\tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T + \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \right) \quad \dots \right. \\ \left. \dots \quad \gamma_{f(c)}^{(j)} \gamma_{(c)}^{(j)} \left(\tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T + \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\xi}^{(j)} \right)^T \right) \right]. \end{aligned}$$

Em outras palavras, tem-se

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \gamma_{(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \quad \gamma_{f(1)}^{(j)} \gamma_{(2)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \quad \dots \right. \\ \left. \dots \quad \gamma_{f(c)}^{(j)} \gamma_{(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \right] \end{aligned} \quad (3.97)$$

e esta matriz de covariância nebulosa tem a seguinte propriedade

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \gamma_{(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}_{rc}^{(j)} \right)^T \quad \gamma_{f(2)}^{(j)} \gamma_{(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \quad \dots \right. \\ \left. \dots, \gamma_{f(c)}^{(j)} \gamma_{(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \right] \neq \mathbf{0} \end{aligned} \quad (3.98)$$

e

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \gamma_{(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \quad \gamma_{f(2)}^{(j)} \gamma_{(1)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \quad \dots, \right. \\ \left. \gamma_{f(c)}^{(j)} \gamma_{(c)}^{(j)} \tilde{\boldsymbol{\psi}}^{(j)} \left(\boldsymbol{\psi}^{(j)} \right)^T \right] \mathbf{C}_{\tilde{\boldsymbol{\psi}}\boldsymbol{\psi}_e} \neq \mathbf{0} \quad \square \end{aligned} \quad (3.99)$$

3.4.4 Análise de Convergência - Teorema 1

Teorema 1 Considere $\tilde{\boldsymbol{\psi}}^{(j)}$, um vetor nebuloso de mesma ordem $\boldsymbol{\psi}^{(j)}$, associado a dinâmica do sistema e decorrelacionado com o ruído de entrada e saída $\boldsymbol{\xi}^{(j)}$. Considere $\boldsymbol{\psi}_e^{(j)}$ um vetor com erro na variável, $\boldsymbol{\psi}_e^{(j)} = \boldsymbol{\psi}^{(j)} + \boldsymbol{\xi}^{(j)}$; os graus de ativação normalizados $\gamma_{f(i)}^{(j)}$ e $\gamma_{(i)}^{(j)}$, obtidos na equação (2.39), associados a $\boldsymbol{\psi}_f^{(j)}$ e $\boldsymbol{\psi}_e^{(j)}$, respectivamente, para $i = 1, 2, \dots, c$; $j = 1, 2, \dots, k$, e $k = 1, 2, \dots, N$.

Então, no limite, tem-se

$$p.\lim_{k \rightarrow \infty} \left(\hat{\Theta}^{(k)} - \Theta \right) = \mathbf{0}. \quad (3.100)$$

Prova. De acordo com a equação (3.89), esta pode ser expressa da seguinte forma

$$\begin{aligned} & \left\{ \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\psi}^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\psi}^{(j)} \right] \left[\gamma_{(1)}^{(j)} \psi_e^{(j)} \quad \dots \quad \gamma_{(c)}^{(j)} \psi_e^{(j)} \right]^T \right\} \left(\hat{\Theta}^{(k)} - \Theta \right) = \\ & \dots \frac{1}{k} \sum_{j=1}^k \left[\gamma_{f(1)}^{(j)} \tilde{\psi}^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \tilde{\psi}^{(j)} \right] \varepsilon^{(j)}. \end{aligned} \quad (3.101)$$

Aplicando-se a probabilidade no limite com $k \rightarrow \infty$, e aplicando-se o teorema de Slutsky (WILLIAMSON, 2006), tem-se,

$$\begin{aligned} p.\lim_{k \rightarrow \infty} \left(\hat{\Theta}^{(k)} - \Theta \right) &= \left\{ p.\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left\{ \left[\gamma_{f(1)}^{(j)} \tilde{\psi}^{(j)} \quad \gamma_{f(2)}^{(j)} \tilde{\psi}^{(j)} \quad \dots \right. \right. \right. \\ & \left. \left. \left. \dots, \gamma_{f(c)}^{(j)} \tilde{\psi}^{(j)} \right] \left[\gamma_{f(1)}^{(j)} \psi_f^{(j)} \quad \dots \quad \gamma_{f(c)}^{(j)} \psi_f^{(j)} \right]^T \right\} \right\}^{(-1)} \times \\ & \left\{ p.\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \left\{ \left[\gamma_{f(1)}^{(j)} \tilde{\psi}^{(j)}, \dots, \gamma_{f(c)}^{(j)} \tilde{\psi}^{(j)} \right] \varepsilon^{(j)} \right\} \right\}. \end{aligned} \quad (3.102)$$

De acordo com os Lemas 1 e 3, tem-se

$$p.\lim_{k \rightarrow \infty} \left(\hat{\Theta}^{(k)} - \Theta \right) = \left\{ p.\lim_{k \rightarrow \infty} \mathbf{C}_{\tilde{\Psi}\psi_f} \right\}^{(-1)} \mathbf{0}, \quad (3.103)$$

onde a covariância nebulosa $\mathbf{C}_{\tilde{\Psi}_{rc}\psi_{f,rc}}$ é não-singular e, como consequência, existe a matriz inversa. Assim, o valor do limite do parâmetro de erro, em probabilidade, é dado por,

$$p.\lim_{k \rightarrow \infty} \left(\hat{\Theta}^{(k)} - \Theta \right) = \mathbf{0}. \quad (3.104)$$

e os estimadores são assintoticamente não polarizados, como requerido. \square

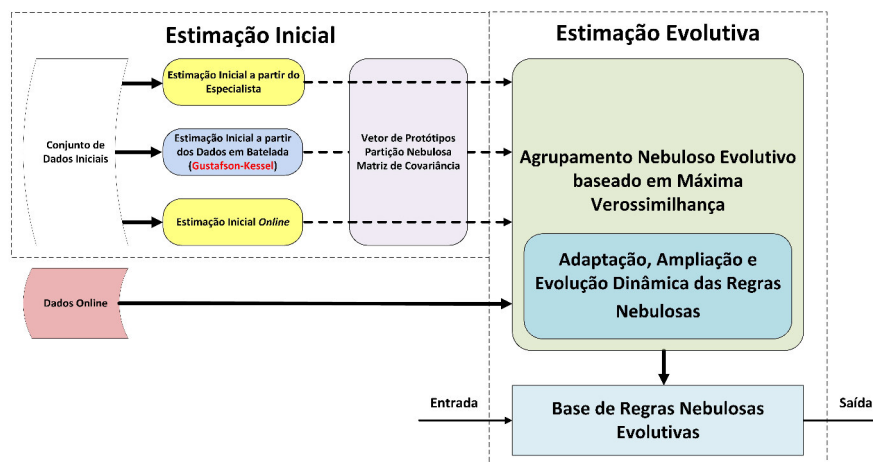
Este capítulo introduziu o uso da variável instrumental nebulosa inserida no contexto evolutivo, destacando-se os seus principais componentes, além da análise de convergência da variável instrumental nebulosa na presença de dados corrompidos por um ruído autocorrelacionado.

4 AGRUPAMENTO NEBULOSO EVOLUTIVO BASEADO EM MÁXIMA VEROSSIMILHANÇA

Neste capítulo, será apresentada a formulação matemática e a metodologia da modelagem baseada em agrupamento nebuloso evolutivo de máxima verossimilhança proposta nesta tese. A modelagem é realizada através de dois algoritmos subsequentes, denominados de: algoritmo de agrupamento nebuloso em batelada (estimação inicial), para estimação paramétrica inicial do modelo nebuloso em batelada, a partir de um conjunto de dados *offline*. O algoritmo de agrupamento nebuloso recursivo (estimação evolutiva), para a estimação paramétrica recursiva do modelo nebuloso evolutivo, a partir de um conjunto de dados *online*, conforme ilustrado na Figura 24 o diagrama de bloco geral desta abordagem.

Nesta metodologia, que é uma nova abordagem em relação as metodologias propostas por (KASABOV, 2002; GEORGIEVA; FILEV, 2009) amplamente divulgadas na teoria de sistemas nebulosos evolutivos, são apresentadas as principais contribuições que permitiram o desenvolvimento desta sistematização com aplicação em problemas reais de engenharia na área de modelagem de sistemas dinâmicos operando em ambiente não-estacionário.

Figura 24: Agrupamento nebuloso evolutivo de máxima verossimilhança compreendido pela etapa de estimação inicial e estimação evolutiva.



Fonte: Elaborado pelo autor.

4.1 Estimação Inicial

A etapa de estimação inicial do algoritmo de agrupamento nebuloso evolutivo é compreendida por três submodelos: estimação inicial a partir do especialista, estimação inicial *online* e **estimação inicial partir de dados em batelada (*offline*)**, que é realizada a partir do algoritmo de agrupamento *Gustafson–Kessel* (BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007; ANGELOV; FILEV; KASABOV, 2010).

- a) **Estimação inicial a partir de dados em batelada:** A partir da estimação inicial os parâmetros do vetor de protótipos, matriz de covariância nebulosa, matriz de partição nebulosa e parâmetros do consequente tem-se parâmetros iniciais do algoritmo de agrupamento evolutivo. Nesta tese, será usada estimação inicial em batelada baseada no algoritmo de agrupamento de *Gustafson–Kessel* por ser um algoritmo de agrupamento que utiliza uma norma adaptativa (BABUSKA, 1998).
- b) **Estimação inicial *online*:** a primeira amostra lida será considerada o primeiro vetor de protótipos agrupamento, a matriz de covariância nebulosa inicial será dada por $\kappa \mathbf{I}$, onde \mathbf{I} é a matriz identidade e κ é uma constante suficiente grande que pode assumir valores na faixa $[10^4, 10^8]$. A estrutura do consequente é definida a partir do primeiro grupo identificado, utilizando o método dos mínimos quadrados recursivos, conforme (KASABOV, 2002; GEORGIEVA; FILEV, 2009).
- c) **Estimação inicial a partir do especialista:** neste submodelo o especialista define quais os tipos de funções de pertinências que serão utilizadas nas variáveis linguísticas do antecedente da estimação inicial, consequente, serão as mesmas funções de pertinências utilizadas no antecedente da estimação evolutiva. Além do mais, permite que o especialista utilize uma combinação híbrida dos submodelos anteriores, permitindo uma flexibilidade maior do agrupamento evolutivo, conforme a experiência adquirida em relação ao conjunto de dados.

4.1.1 Estimação Inicial a partir de Dados em Batelada

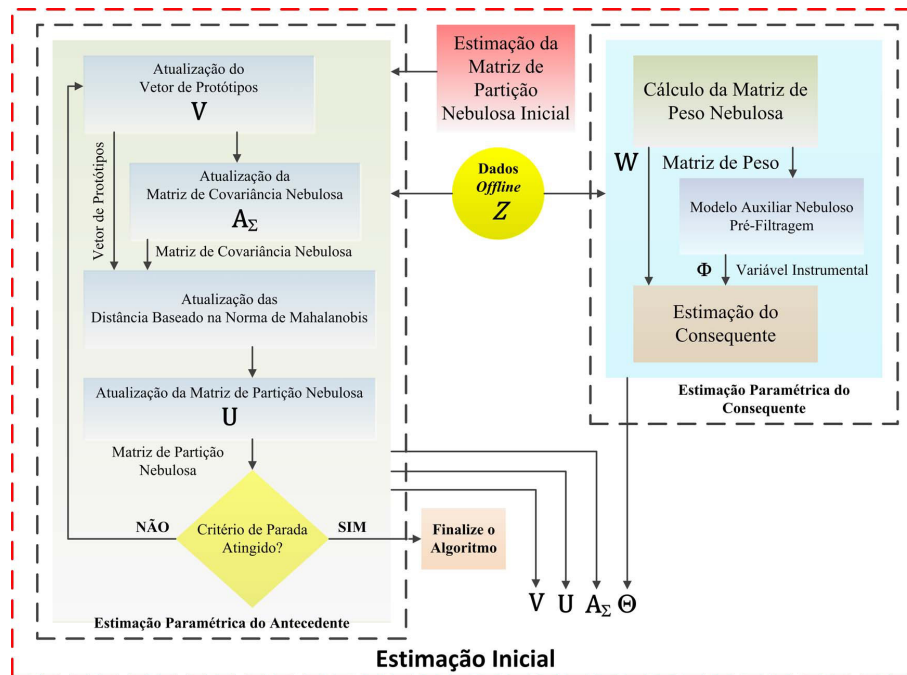
É realizado o agrupamento nebuloso em batelada no conjunto de dados *offline* \mathbf{Z} cujo tamanho não é fixo e é determinado a partir do conhecimento prévio do especialista com base nas observações dos dados experimentais e computacionais, uma vez que o tamanho deste conjunto de dados inicial pode influenciar os resultados finais. Isto é, grandes valores de observações aumentam a robustez, mas também aumentam o custo computacional, observações menores podem levar para um modelo impreciso na estimação inicial do modelo (LUGHOFER, 2011). Além disso, a largura do conjunto de dados *offline* da estimação paramétrica inicial deve ser baseada na análise estatística da razão do erro médio entre a saída real e saída estimada, de modo que o valor ideal do conjunto de dados

permitirá a minimização do erro da estimação paramétrica inicial (TSUMURA; OISHI, 1999; TSUMURA; MACIEJOWSKI, 2003; TSUMURA, 2009).

O algoritmo de agrupamento com matriz de covariância nebulosa é baseado na norma de *Mahalanobis* que é uma métrica de distância adaptativa, ideal para a inicialização do algoritmo de agrupamento nebuloso recursivo que é baseado na métrica de máxima verossimilhança (GATH; GEVA, 1989; BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007). O algoritmo é dividido em duas etapas subsequentes: estimação paramétrica do antecedente e estimação paramétrica do consequente.

Na estimação paramétrica do antecedente é realizado o agrupamento em batelada do conjunto de dados e os seguintes parâmetros são obtidos após a finalização deste: o vetor de protótipos \mathbf{V} , a matriz de partição nebulosa \mathbf{U} e a matriz de covariância nebulosa \mathbf{A}_Σ . Logo em seguida, a estimação paramétrica do consequente é realizada, a partir do método dos mínimos quadrados inserido no contexto de variável instrumental com a ponderação do conjunto de dados utilizando a matriz de pesos nebulosos \mathbf{W} . Após a finalização do algoritmo em batelada obtém-se os parâmetros do consequente Θ da estimação inicial, conforme ilustrado na Figura 25. Os parâmetros dos grupos obtidos na etapa da estimação inicial serão utilizados como parâmetros iniciais da etapa de estimação evolutiva.

Figura 25: Estimação inicial dos parâmetros do modelo nebuloso baseado no algoritmo de *Gustafson-Kessel* compreendido pelas etapas de estimação paramétrica do antecedente e estimação paramétrica do consequente (BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007).



Fonte: Elaborado pelo autor.

4.2 Estimação Evolutiva

Nesta etapa, a estimação paramétrica evolutiva do antecedente e estimação paramétrica evolutiva do conseqüente é realizada a cada iteração k . O algoritmo de agrupamento nebuloso evolutivo baseia-se na norma de máxima verossimilhança como uma nova abordagem em relação às metodologias já existentes na literatura (KASABOV, 2002; GEORGIEVA; FILEV, 2009). O algoritmo é baseado na orientação de dados, isto é, à medida que uma nova amostra é lida do conjuntos de dados *online*, a estimação paramétrica evolutiva do conseqüente é adaptada a nova estrutura modificada através da estimação paramétrica evolutiva do antecedente.

4.2.1 Estimação Paramétrica Evolutiva do Antecedente

O algoritmo é inicializado com dois grupos e os respectivos parâmetros iniciais são obtidos a partir da estimação inicial, *passo 0* do algoritmo. Nos demais passos do algoritmo são verificados se uma nova amostra lida deve ser incorporada aos grupos existentes e logo em seguida os parâmetros dos grupos são corrigidos; caso não seja incorporada aos grupos existentes, será criado um novo grupo e avaliado a atualização de seus parâmetros. Os parâmetros dos grupos atualizados no instante k são: vetor de protótipos, matriz de covariância nebulosa, matriz partição nebulosa e raio do grupo.

O raio do grupo é uma medida que se baseia no fator exponencial da máxima verossimilhança e determina a maior distância existente entre amostras e o respectivo centro do grupo as quais elas pertencem. A criação de novos grupos está condicionada a seguinte restrição: se a distância de uma nova amostra lida é maior do que o raio do grupo mais próximo, grupo de maior similaridade, é verdadeira, o mecanismo de criação de novos grupos avaliará a credibilidade da criação deste, se a condição do critério for satisfeita, o novo grupo será criado. Caso contrário, a amostra é incorporada ao grupo de maior similaridade e os parâmetros são atualizados conforme a estrutura atual no instante k .

As etapas que compõem a estimação paramétrica evolutiva do antecedente são:

- a) **Estimação do coeficiente evolutivo–recursivo:** é calculado a probabilidade a priori de uma amostra pertencer a um grupo em função do fator exponencial da máxima verossimilhança no instante k , uma nova abordagem em relação ao procedimento proposto em (BABUSKA, 1998; ABONYI, 2003; ABONYI et al., 2005; ABONYI; FEIL, 2007);
- b) **Estimação das distâncias em máxima verossimilhança:** as pertinências pertencentes a matriz de partição nebulosa serão calculadas em função da norma de máxima verossimilhança (GATH; GEVA, 1989; BABUSKA, 1998; ABONYI et al., 2005);

- c) **Estimação do grupo mais próximo:** o grupo de maior similaridade em relação à amostra lida é determinado, logo após estima-se o raio deste grupo;
- d) **Estimação do raio do grupo mais próximo:** a maior distância existente entre um centro do grupo e as demais amostras pertencentes, calculados a partir do fator exponencial da máxima verossimilhança;
- e) **Atualização dos parâmetros dos grupos:** os parâmetros dos grupos no instante k são atualizados, após determinar o grupo de maior similaridade da amostra lida no instante k , uma nova abordagem em relação ao procedimento proposto em (KASABOV, 2002; GEORGIEVA; FILEV, 2009);
- f) **Fator de credibilidade:** fator que evita crescimento exponencial de grupos à medida que novas amostras são lidas;
- g) **Atualização e junção dos grupos:** critério adotado para otimizar o menor número de grupos existentes, de forma a garantir o mínimo possível segundo um critério de compatibilidade (BABUSKA, 1998).

Para minimizar o problema do crescimento exponencial dos grupos (regras nebulosas), *maldição da dimensionalidade*, duas condições restritivas na estimação paramétrica evolutiva do antecedente serão avaliadas: o fator de credibilidade e o algoritmo de atualização e junção de grupos. Diferente de (KASABOV, 2002; GEORGIEVA; FILEV, 2009) que adota o fator de credibilidade segundo um critério estabelecido para evitar o problema do crescimento exponencial dos grupos; nesta tese, a partir de condições pré-estabelecidas, ambas as condições serão avaliadas no instante k a fim de garantir um menor número de regras possível, sem comprometer o desempenho do modelo final obtido na estimação paramétrica evolutiva do consequente.

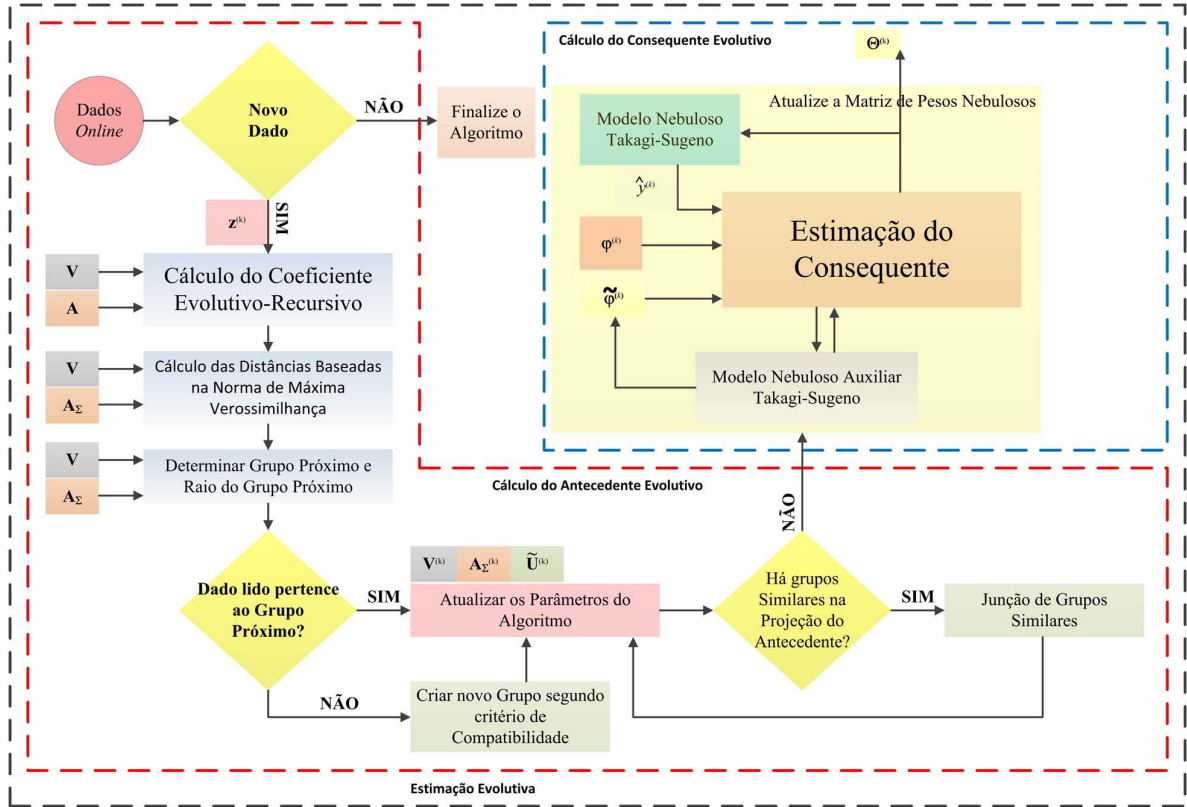
4.2.2 Estimação Paramétrica Evolutiva do Consequente

Nesta etapa, a estrutura de modelo nebuloso *Takagi-Sugeno* inserido no contexto evolutivo é modificada a cada iteração k , após a finalização da estimação paramétrica evolutiva do antecedente. Os parâmetros do consequente são inicializados a partir dos parâmetros obtidos do modelo inicial.

Os parâmetros do consequente a partir do instante k da etapa evolutiva são obtidos através do método dos mínimos quadrados ponderados, com uso da variável instrumental nebulosa inserida no contexto evolutivo. A ponderação deste método é realizado a partir dos pesos nebulosos obtidos da matriz de partição nebulosa no instante k . A formulação matemática da variável instrumental nebulosa inserida no contexto evolutivo foi apresentada no capítulo 3.

O diagrama de bloco a etapa evolutiva composta pela estimação paramétrica evolutiva do antecedente e estimação paramétrica evolutiva do consequente é ilustrado na Figura 26.

Figura 26: Estimação evolutiva: etapa de ajustes da estimação paramétrica evolutiva do antecedente e a estimação paramétrica evolutiva do consequente no instante k .



Fonte: Elaborado pelo autor.

4.3 Formulação do Algoritmo de Agrupamento Nebuloso em Bate-lada: estimação inicial

Seja c o número de grupos, com $2 \leq c \leq N_{total}$, que podem ser representados por um vetor de protótipos

$$\mathbf{V} = [\mathbf{v}_{(1)}, \mathbf{v}_{(2)}, \dots, \mathbf{v}_{(c)}], \mathbf{v}_i \in R^{p+1}, i = 1, 2, \dots, c, \quad (4.1)$$

com c centros e a dimensionalidade, p , do vetor de entrada de dados, a qual representa as variáveis medidas. N_{total} é o total de dados experimentais ou computacionais obtidos do sistema a ser identificado, dividido em N_{bt} dados para estimação em batelada e N_{rc} dados para estimação evolutiva.

A matriz de dados utilizada pelo algoritmo de agrupamento nebuloso em batelada é expressa por

$$\mathbf{Z}_{bt} = [\mathbf{X}_{bt}, \mathbf{y}_{bt}], \quad (4.2)$$

como segue

$$\mathbf{Z}_{bt} = [\mathbf{z}_{bt}^{(1)}, \mathbf{z}_{bt}^{(2)}, \dots, \mathbf{z}_{bt}^{(N_{bt})}]^T, \quad (4.3)$$

com $\mathbf{z}_{bt}^{(j)} \in \mathbb{R}^{p+1}$, onde

$$\mathbf{X}_{bt} = [\mathbf{x}_{bt}^{(1)}, \mathbf{x}_{bt}^{(2)}, \dots, \mathbf{x}_{bt}^{(N_{bt})}], \quad (4.4)$$

com a matriz de regressores representada por $\mathbf{x}_{bt}^{(j)} \in \mathbb{R}^p$, e

$$\mathbf{y}_{bt} = [y_{bt}^{(1)}, y_{bt}^{(2)}, \dots, y_{bt}^{(N_{bt})}]^T, \quad (4.5)$$

com o vetor de dados de saída, $y_{bt}^{(k)} \in \mathbb{R}$, para $\forall k = 1, 2, \dots, N_{bt}$.

4.3.1 Estimação Paramétrica em Batelada do Antecedente

O algoritmo de agrupamento nebuloso em batelada, inicialmente, é implementado a partir de um conjunto de dados com N_{bt} elementos para estimação inicial do modelo nebuloso, obtido via algoritmo *Gustafson-Kessel* através da minimização da função-objetivo (FGK) (BABUSKA, 1998), como segue:

$$(\tilde{\mathbf{U}}_{bt}, \mathbf{V}_{bt}, \mathbf{A}_{\Sigma_{bt}}) = FGK(\mathbf{Z}_{N_{bt}}, c, \tilde{\mathbf{U}}_0, \varepsilon), \quad (4.6)$$

com

$$\tilde{\mathbf{U}}_0 \in \mathbb{R}^{c \times N_{bt}} \left| \begin{array}{l} \mu_{0(i)}^{(k)} \in [0, 1], \forall i, k; \\ \sum_{i=1}^c \mu_{0(i)}^{(k)} = 1, \forall k; \\ 0 < \sum_{k=1}^{N_{bt}} \mu_{0(i)}^{(k)} < N_{bt}, \forall i, \end{array} \right. \quad (4.7)$$

onde

$\tilde{\mathbf{U}}_0$ é matriz de partição nebulosa inicial obtida de forma aleatória satisfazendo as condições impostas em (4.7);

ε é a tolerância do erro de convergência da matriz de partição nebulosa;

$\tilde{\mathbf{U}}_{bt} \in \mathbb{R}^{c \times N_{bt}}$ corresponde a matriz de partição nebulosa obtida através do algoritmo em batelada;

$\mathbf{V}_{bt} \in \mathbb{R}^{c \times p+1}$ corresponde ao vetor de protótipos dos grupos obtido através do algoritmo em batelada;

$\mathbf{A}_{\Sigma_{bt}} \in \mathbb{R}^{c \cdot p \times p}$ corresponde à matriz de covariância nebulosa obtida através do algoritmo em batelada.

O algoritmo proposto em batelada corresponde à solução da função-objetivo *FGK* que emprega a métrica de *Mahalanobis*. Na sequência, será discutida a formulação matemática do algoritmo proposto e seus principais passos, (BABUSKA, 1998; BABUSKA B. SCHUTTER; GUERRA, 2010; JEEVA; PONNUSAMY; SANGAPILLAI, 2013).

O vetor de protótipos do algoritmo em batelada, na iteração (l) , para $1 \leq i \leq c$ e $1 \leq k \leq N_{bt}$, é dada por:

$$\mathbf{v}_{bt(i)}^{(l)} = \frac{\sum_{k=1}^{N_{bt}} \left[\left(\mu_{bt(i)}^{(k)(l-1)} \right)^m \mathbf{z}_{bt}^{(k)} \right]}{\sum_{k=1}^{N_{bt}} \left(\mu_{bt(i)}^{(k)(l-1)} \right)^m}, \forall i, k. \quad (4.8)$$

A matriz de covariância nebulosa $A_{\Sigma_{bt(i)}}$ no algoritmo em batelada, na iteração $(l-1)$, para $1 \leq i \leq c$ e $1 \leq k \leq N_{bt}$, é dada por:

$$A_{\Sigma_{bt(i)}} = \frac{\sum_{k=1}^{N_{bt}} \left(\mu_{bt(i)}^{(k)(l-1)} \right)^m \left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(i)}^{(l)} \right) \left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(i)}^{(l)} \right)^T}{\sum_{k=1}^{N_{bt}} \left(\mu_{bt(i)}^{(k)(l-1)} \right)^m}, \quad (4.9)$$

com fator de incerteza $m > 1$.

As distâncias dos dados em relação aos centros dos grupos nebulosos, segundo o critério da métrica de *Mahalanobis*, para $1 \leq i \leq c$, e $1 \leq k \leq N_{bt}$, é dado por:

$$D_{ikA_{\Sigma_{bt}}}^2 = \left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(i)}^{(l)} \right)^T \left[\det \left(A_{\Sigma_{bt(i)}} \right)^{1/p} \left(A_{\Sigma_{bt(i)}} \right)^{-1} \right] \left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(i)}^{(l)} \right), \forall i, k. \quad (4.10)$$

Na iteração l , a atualização dos elementos da matriz de partição nebulosa na i ésima regra e instante k , $\mu_{bt(i)}^{(k)(l)}$, para a condição $D_{ikA_{\Sigma_{bt}}} > 0$ com $1 \leq i \leq c$, $1 \leq j \leq c$ e $1 \leq k \leq N_{bt}$, é dada por:

$$\mu_{bt(i)}^{(k)(l)} = \frac{1}{\sum_{j=1}^c \left(\frac{D_{ikA_{\Sigma_{bt}}}}{D_{jkA_{\Sigma_{bt}}}} \right)^{2/(m-1)}}. \quad (4.11)$$

Escrevendo a equação (4.11) nos termos da equação (4.10), obtém-se

$$\mu_{bt(i)}^{(k)(l)} = \frac{1}{\sum_{j=1}^c \left(\frac{\left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(i)}^{(l)} \right)^T \left[\det \left(A_{\Sigma_{bt(i)}} \right)^{1/p} \left(A_{\Sigma_{bt(i)}} \right)^{-1} \right] \left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(i)}^{(l)} \right)}{\left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(j)}^{(l)} \right)^T \left[\det \left(A_{\Sigma_{bt(j)}} \right)^{1/p} \left(A_{\Sigma_{bt(j)}} \right)^{-1} \right] \left(\mathbf{z}_{bt}^{(k)} - \mathbf{v}_{bt(j)}^{(l)} \right)} \right)^{1/(m-1)}}. \quad (4.12)$$

Considerando a equação (4.11) para as condições $D_{ikA_{\Sigma_{bt}}}$ igual a zero e $j \neq i$ com $1 \leq i \leq c$, $1 \leq j \leq c$ e $1 \leq k \leq N_{bt}$, a atualização dos elementos da matriz de partição nebulosa na i ésima regra e instante k , $\mu_{bt(i)}^{(k)(l)}$, é dada por:

$$\mu_{bt(i)}^{(k)(l)} = 1, \quad (4.13)$$

$$\mu_{bt(j)}^{(k)(l)} = 0, \quad (4.14)$$

com, $m > 1$, e $\mu_{bt(i)}^{(k)(l)} \in [0, 1]$ para $\sum_{i=1}^c \mu_{bt(i)}^{(k)(l)} = 1$.

4.3.2 Passos do Algoritmo do Antecedente para Gustafson–Kessel

O algoritmo é representado pelos passos descritos a seguir, dado o conjunto de dados $\mathbf{Z}_{bt} = [\mathbf{X}_{bt}, \mathbf{y}_{bt}]$ com N_{bt} elementos.

-
- 1: Definir o parâmetro: valor da incerteza $m > 1$.
 - 2: Definir o parâmetro: tolerância do erro de convergência da matriz de partição nebulosa $\varepsilon > 0$.
 - 3: Definir o parâmetro: o número inicial de protótipos $c > 2$.
 - 4: Inicializar a matriz de partição nebulosa inicial $\tilde{\mathbf{U}}_0$ de acordo com as restrições (4.7).
 - 5: $l \leftarrow 0$
 - 6: **repetir**
 - 7: $l \leftarrow l + 1$
 - 8: Determinar o vetor de protótipos dos grupos nebulosos através da equação (4.8).
 - 9: Determinar a matriz de covariância nebulosa através da equação (4.9).
 - 10: Determinar as distâncias de \mathbf{z}_{bt} em relação ao vetor de protótipos $\mathbf{v}_{bt(i)}$ através da equação (4.10).
 - 11: Atualizar a matriz de partição nebulosa através das equações (4.11).
 - 12: **até que** $\left\| \tilde{\mathbf{U}}_{bt}^{(l)} - \tilde{\mathbf{U}}_{bt}^{(l-1)} \right\| < \varepsilon$.
 - 13: Parâmetros obtidos: \mathbf{V}_{bt} , $\mathbf{A}_{\Sigma_{bt}}$ e $\tilde{\mathbf{U}}_{bt}$.
-

Após a finalização dos passos do algoritmo do antecedente para *Gustafson–Kessel*, os parâmetros obtidos: \mathbf{V}_{bt} , $\mathbf{A}_{\Sigma_{bt}}$ e $\tilde{\mathbf{U}}_{bt}$ são utilizados para determinar os parâmetros do consequente na estimação paramétrica em batelada do consequente.

4.3.3 Estimação Paramétrica em Batelada do Consequente

No sentido de superar a influência do ruído autocorrelacionado, possivelmente presente nos dados experimentais, na estimação paramétrica do consequente da etapa inicial, propõe-se o uso do método da variável instrumental nebulosa não-recursiva inserida em um contexto de máxima verossimilhança para obtenção dos parâmetros do consequente.

Inicialmente, seja o vetor de regressores de dados do sistema dinâmico no antecedente, expressos a seguir

$$\mathbf{x}_{bt}^{(k-1)} = \left[y_{bt}^{(k-1)} \quad \dots \quad y_{bt}^{(k-n_y)} \quad u_{bt}^{(k-1)} \quad \dots \quad u_{bt}^{(k-n_u)} \right]^T, \quad (4.15)$$

onde $y_{bt}^{(k)}$, $u_{bt}^{(k)}$, n_y e n_u são: saída do sistema dinâmico, entrada externa do sistema dinâmico, dimensionalidade do vetor $y_{bt}^{(k)}$ e dimensionalidade do vetor $u_{bt}^{(k)}$, respectivamente. A matriz de regressores é expressa por

$$\mathbf{X}_{bt} = \left[\mathbf{x}_{bt}^{(1)} \quad \mathbf{x}_{bt}^{(2)} \quad \dots \quad \mathbf{x}_{bt}^{(N_{bt})} \right]^T. \quad (4.16)$$

Seja a variável de entrada auxiliar escrita na forma vetorial obtida a partir da equação (3.50) do modelo nebuloso auxiliar, expressa por

$$\tilde{\mathbf{u}}_{bt}^{(k)} = \left[\tilde{u}_{bt}^{(k-1)} \quad \tilde{u}_{bt}^{(k-2)} \quad \dots \quad \tilde{u}_{bt}^{(k-n_u)} \right]^T. \quad (4.17)$$

Seja a variável de saída auxiliar escrita na forma vetorial obtida a partir da equação (3.44) do modelo nebuloso auxiliar, expressa por

$$\tilde{\mathbf{y}}_{bt}^{(k)} = \left[\tilde{y}_{bt}^{(k-1)} \quad \tilde{y}_{bt}^{(k-2)} \quad \dots \quad \tilde{y}_{bt}^{(k-n_y)} \right]^T. \quad (4.18)$$

A variável instrumental nebulosa, equação (3.83), nos termos das equações (4.15), (4.17) e (4.18) é reescrita da seguinte forma

$$\tilde{\mathbf{x}}_{bt}^{(k-1)} = \left[\tilde{y}_{bt}^{(k-1)} \quad \tilde{y}_{bt}^{(k-2)} \quad \dots \quad \tilde{y}_{bt}^{(k-n_y)} \quad \tilde{u}_{bt}^{(k-1)} \quad \tilde{u}_{bt}^{(k-2)} \quad \dots \quad \tilde{u}_{bt}^{(k-n_u)} \right]^T. \quad (4.19)$$

A variável instrumental nebulosa (4.19) escrita na forma vetorial é dada por

$$\tilde{\mathbf{X}}_{bt} = \left[\tilde{\mathbf{x}}_{bt}^{(1)} \quad \tilde{\mathbf{x}}_{bt}^{(2)} \quad \tilde{\mathbf{x}}_{bt}^{(3)} \quad \dots \quad \tilde{\mathbf{x}}_{bt}^{(N)} \right]^T. \quad (4.20)$$

A matriz de pesos nebulosos da estimação paramétrica em batelada do consequente é obtida a partir da matriz de partição do agrupamento inicial, realizado nas N_{bt} observações do conjuntos de dados, definida como segue:

$$\tilde{\mathbf{U}}_{bt} \in \mathbb{R}^{c \times N_{bt}} \left| \mu_{bt(i)}^{(k)} \in [0, 1], \forall i, k; \sum_{i=1}^c \mu_{bt(i)}^{(k)} = 1, \forall k; 0 < \sum_{k=1}^{N_{bt}} \mu_{bt(i)}^{(k)} < N_{bt}, \forall i, \right. \quad (4.21)$$

onde o grau de pertinência da amostra $\left\{ \left(\mathbf{x}_{bt}^{(k)}, y_{bt}^{(k)} \right) \right\}$ é expresso por

$$\mu_{(i)}^{(k)} = \mu_{(i)}^{(k)} \left(\left\{ \left(\mathbf{x}_{bt}^{(k)}, y_{bt}^{(k)} \right) \right\} \right) \forall k \quad (4.22)$$

e a matriz de pesos nebulosos para a amostra $\left\{ \left(\mathbf{x}_{bt}^{(k)}, y_{bt}^{(k)} \right) \right\}$ é definida como segue

$$\widetilde{\mathbf{W}}_{bt}^{(i)} = \text{diag} \left(\mu_{(i)}^{(k)} \right) \in \mathbb{R}^{N \times N}, \quad (4.23)$$

onde

$$\widetilde{\mathbf{W}}_{bt}^{(i)} = \begin{pmatrix} \tilde{w}_{bt}^{(1)} = \mu_{(i)}^{(1)} & 0 & \dots & 0 \\ 0 & \tilde{w}_{bt}^{(2)} = \mu_{(i)}^{(2)} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \tilde{w}_{bt}^{(N)} = \mu_{(i)}^{(N)} \end{pmatrix}, \quad (4.24)$$

para $i = 1, 2, \dots, c; k = 1, 2, \dots, N_{bt}$.

O vetor de parâmetros do consequente obtidos localmente é dado por

$$\Theta_{bt}^{(i)} = \left(\tilde{\mathbf{X}}_{bt}^T \widetilde{\mathbf{W}}_{bt(i)} \mathbf{X}_{bt} \right)^{-1} \tilde{\mathbf{X}}_{bt}^T \widetilde{\mathbf{W}}_{bt(i)} y_{bt} \quad i = 1, 2, \dots, c. \quad (4.25)$$

O refinamento do vetor de parâmetros do conseqüente obtidos a partir da equação (4.25) é dado por

$$\tilde{\Theta}_{bt}^{(k)(i)} = \tilde{\Theta}_{bt}^{(k-1)(i)} + \mathbf{K}_{bt}^{(k)} \left(y_{bt}^{(k)} - \mathbf{x}_{bt}^{(k)T} \tilde{\Theta}_{bt}^{(k-1)(i)} \right), \quad (4.26)$$

onde o ganho de adaptação do vetor estimado, referido como ganho de *Kalman*, é

$$\mathbf{K}_{bt}^{(k)} = \frac{\tilde{w}_{bt_i}^{(k)} \mathbf{P}_{bt}^{(k-1)} \tilde{\mathbf{x}}_{bt}^{(k)}}{\left(\delta_{bt}^{(k)} + \tilde{w}_{bt_i}^{(k)} \mathbf{x}_{bt}^{(k)T} \mathbf{P}_{bt}^{(k-1)} \tilde{\mathbf{x}}_{bt}^{(k)} \right)}. \quad (4.27)$$

A matriz de covariância é expressa por

$$\mathbf{P}_{bt}^{(k)} = \frac{1}{\delta_{bt}^{(k)}} \left[\mathbf{P}_{bt}^{(k-1)} - \mathbf{K}_{bt}^{(k)} \tilde{\Theta}_{bt}^{(k)} \mathbf{x}_{bt}^{(k)T} \mathbf{P}_{bt}^{(k-1)} \right], \quad (4.28)$$

onde $\tilde{w}_{bt(i)}^{(k)} \Big|_{i=1,2,\dots,c}$ é a atualização dos pesos nebulosos a cada nova iteração, para a i ésima regra.

O fator de esquecimento atualizado a cada nova iteração do algoritmo recursivo é dado por

$$\delta_{bt}^{(k)} = \lambda_0 \delta_{bt}^{(k-1)} + (1 - \lambda_0) (\delta_0) - \left(\sigma_{bt}^{(k-1)} \right)^2, \quad (4.29)$$

com $i = 1, 2, \dots, c$ e $k = 1, 2, \dots, N_{bt}$.

A variância móvel do erro de estimação a cada nova iteração é expresso por

$$\left(\sigma_{bt}^{(k)} \right)^2 = \left(\sigma_{bt}^{(k-1)} \right)^2 - \frac{1}{k} \left[\left(\sigma_{bt}^{(k-1)} \right)^2 - (\Delta_y)^2 \right], \quad (4.30)$$

onde

$$\Delta_y = y_{bt}^{(k)} - \hat{y}_{bt}^{(k)} - \bar{y}_{bt}^{(k)}, \quad (4.31)$$

de modo que

$$\bar{y}_{bt}^{(k)} = \frac{1}{k} \sum_{j=1}^k \left(y_{bt}^{(j)} - \hat{y}_{bt}^{(j)} \right), \quad (4.32)$$

onde $y_{bt}^{(k)}$ e $\hat{y}_{bt}^{(k)}$ são as saídas real e estimada do modelo no instante k , respectivamente.

4.3.4 Passos do Algoritmo do Conseqüente Gustafson–Kessel

O algoritmo é representado pelos passos descritos a seguir:

-
- 1: Determinar a variável instrumental nebulosa da estimação paramétrica em batelada através das equações (4.15)–(4.20).
 - 2: Determinar a matriz nebulosa de pesos a partir da matriz de partição nebulosa (**Estimação Paramétrica em Batelada**) através das equações (4.21)–(4.24).
 - 3: Calcular o vetor de parâmetros em batelada através da equação (4.25).
 - 4: Refinar os parâmetros do conseqüente em batelada através das equações (4.26)–(4.32).
 - 5: Parâmetros obtidos: Θ_{bt} .
-

4.4 Algoritmo de Agrupamento Nebuloso Recursivo: estimação evolutiva

O algoritmo evolutivo proposto corresponde à atualização do vetor de protótipos dos grupos $\mathbf{V}_{rc}^{(k)}$, da matriz de covariância $\mathbf{A}_{\Sigma_{rc}}^{(k)}$ e da matriz de partição nebulosa $\tilde{\mathbf{U}}_{rc}^{(k)}$, bem como dos parâmetros do conseqüente $\Theta_{rc}^{(i)(k)}$ a cada nova amostra lida, $\mathbf{z}_{rc}^{(k)}$, na iteração k . A similaridade entre uma nova amostra lida em relação aos centros dos grupos é mensurado a partir da métrica de máxima verossimilhança.

4.4.1 Estimação Paramétrica Evolutiva do Antecedente

Os valores iniciais do vetor de protótipos $\mathbf{V}_{rc}^{(0)}$, da matriz de covariância $\mathbf{A}_{\Sigma_{rc}}^{(0)}$, da matriz de partição nebulosa $\tilde{\mathbf{U}}_{rc}^{(0)}$ e parâmetros do conseqüente $\Theta_{rc}^{(i)(0)}$ são dados pelos respectivos valores obtidos na equação 4.3, como segue:

$$\mathbf{V}_{rc}^{(0)} = \mathbf{V}_{bt}, \quad (4.33)$$

$$\mathbf{A}_{\Sigma_{rc}}^{(0)} = \mathbf{A}_{\Sigma_{bt}}, \quad (4.34)$$

$$\mathbf{U}_{rc}^{(0)} = \tilde{\mathbf{U}}_{bt}, \quad (4.35)$$

$$\Theta_{rc}^{(i)(0)} = \Theta_{bt}^{(i)}, \quad (4.36)$$

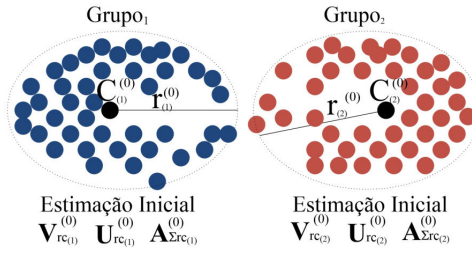
com $i = 1, 2, \dots, c$.

Uma descrição da formulação explícita dos critérios de evolução das regras nebulosas, que são incorporados na metodologia proposta, são mostrados nas Figuras 27–31. Os critérios de evolução da regra podem ser explicados da seguinte forma:

- **Passo 0:** na estimação inicial, o vetor de protótipos, a matriz de covariância e a matriz de partição são estimados a partir do algoritmo visto em 4.3.1, no instante (0), como ilustrado na Figura 27.
- **Passo 1:** após a obtenção de uma nova amostra, as distâncias $D_{ikA_{\Sigma}}$ entre esta amostra e os centros existentes $\mathbf{C}_{(i)}^{(k)}$, para $i = 1, 2, \dots, c$, no instante k , são computadas:
 - **Se** uma nova amostra não pertence a qualquer grupo existente, um novo grupo é criado e, portanto, uma nova regra nebulosa é criada, como ilustrado nas Figuras 28 e 30.
 - **Se** uma nova amostra pertence a um dos grupos existentes, as regras nebulosas são atualizadas, como ilustrado nas Figuras 29 e 31.

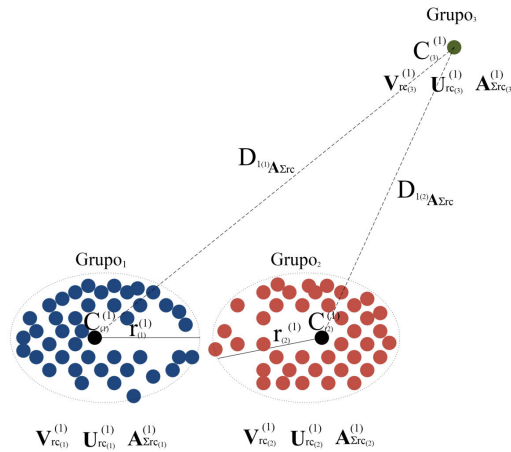
- **Passo 2:** enquanto há nova amostra, os passos 0 e 1 são repetidos.

Figura 27: Um breve processo de agrupamento nebuloso evolutivo, representado com amostras em um espaço 2-D. Após a estimativa inicial do agrupamento em batelada, os centros $C_{(1)}^{(0)}$ e $C_{(2)}^{(0)}$ definidos para os grupos 1 e grupos 2, respectivamente; os raios $r_{(1)}^{(0)}$ e $r_{(2)}^{(0)}$ definidos para os grupos 1 e grupo 2, respectivamente. Obtidos a partir da etapa em batelada denominado de instante inicial (0).



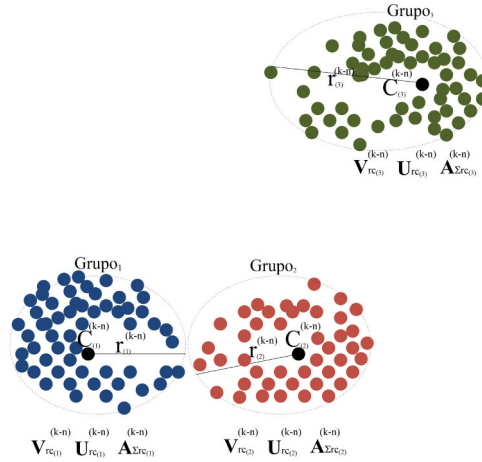
Fonte: Elaborado pelo autor.

Figura 28: Após a leitura da primeira amostra, um novo grupo é criado, $C_{(3)}^{(1)}$, e os demais grupos são atualizados: $C_{(1)}^{(0)} \rightarrow C_{(1)}^{(1)}$ e $C_{(2)}^{(0)} \rightarrow C_{(2)}^{(1)}$. Os parâmetros dos grupos são atualizados: $V_{(i)}^{(0)} \rightarrow V_{(i)}^{(1)}$, $U^{(0)} \rightarrow U^{(1)}$ e $A_{\Sigma(i)}^{(0)} \rightarrow A_{\Sigma(i)}^{(1)}$, para $i = 1, 2, 3$.



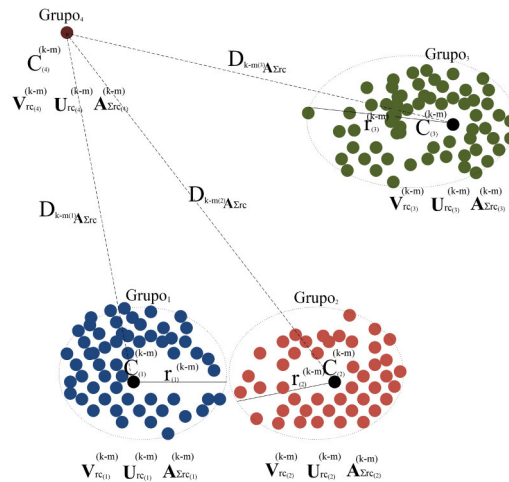
Fonte: Elaborado pelo autor.

Figura 29: Após a leitura da $(k-n)$ ésima amostra, os parâmetros do grupos são atualizados: $\mathbf{V}_{(i)}^{(k-n-1)} \rightarrow \mathbf{V}_{(i)}^{(k-n)}$, $\mathbf{U}_{(i)}^{(k-n-1)} \rightarrow \mathbf{U}_{(i)}^{(k-n)}$ e $\mathbf{A}_{\Sigma(i)}^{(k-n-1)} \rightarrow \mathbf{A}_{\Sigma(i)}^{(k-n)}$, para $i = 1, 2, 3$.



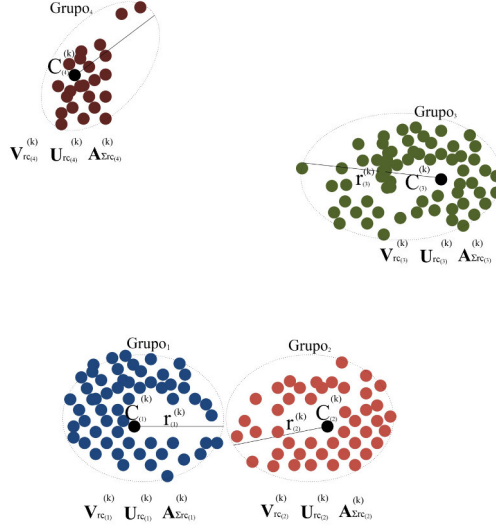
Fonte: Elaborado pelo autor.

Figura 30: Após a leitura da $(k-m)$ ésima amostra, um novo grupo é criado. Os parâmetros dos grupos são atualizados: $\mathbf{V}_{(i)}^{(k-m-1)} \rightarrow \mathbf{V}_{(i)}^{(k-m)}$, $\mathbf{U}_{(i)}^{(k-m-1)} \rightarrow \mathbf{U}_{(i)}^{(k-m)}$ e $\mathbf{A}_{\Sigma(i)}^{(k-m-1)} \rightarrow \mathbf{A}_{\Sigma(i)}^{(k-m)}$, para $i = 1, 2, 3, 4$.



Fonte: Elaborado pelo autor.

Figura 31: Após a leitura da (k) ésima amostra, os parâmetros dos grupos são atualizados:
 $\mathbf{V}_{(i)}^{(k-1)} \rightarrow \mathbf{V}_{(i)}^{(k)}$, $\mathbf{U}_{(i)}^{(k-1)} \rightarrow \mathbf{U}_{(i)}^{(k)}$ e $\mathbf{A}_{\Sigma_{(i)}}^{(k-1)} \rightarrow \mathbf{A}_{\Sigma_{(i)}}^{(k)}$, para $i = 1, 2, 3, 4$.



Fonte: Elaborado pelo autor.

4.4.1.1 Estimação da Probabilidade a Priori

De acordo com (GATH; GEVA, 1989; ABONYI, 2003; BABUSKA, 1998; ABONYI; FEIL, 2007; BABUSKA B. SCHUTTER; GUERRA, 2010), a métrica da máxima verossimilhança de uma dada amostra $\mathbf{z}_{rc}^{(k)}$ em relação ao i ésimo centro do grupo no instante k $\|\mathbf{d}_{zv(i)}^{(k)}\|_{ML}$ para $i = 1, 2, \dots, c$, é expressa por

$$\|\mathbf{d}_{zv(i)}^{(k)}\|_{ML} = \frac{(2\pi)^{p/2} \sqrt{A_{\Sigma_{rc}}^{(k-1)}}}{P_{(i)}} \exp \left[\frac{1}{2} \left(\mathbf{d}_{zv(i)}^{(k)} \right) \left(A_{\Sigma_{rc}}^{(k-1)} \right)^{-1} \left(\mathbf{d}_{zv(i)}^{(k)} \right)^T \right], \quad \forall k, i \quad (4.37)$$

e

$$\mathbf{d}_{zv(i)}^{(k)} = \mathbf{z}_{rc}^{(k)} - \mathbf{v}_{rc(i)}^{(k-1)}. \quad (4.38)$$

A probabilidade *a priori* calculada no instante k para $i = 1, 2, \dots, c$, é expressa por

$$P_{(i)} = \frac{1}{k} \sum_{j=1}^k \mu_{rc(i)}^{(j)}, \quad (4.39)$$

O cálculo da probabilidade *a priori* definida pela equação (4.39) é em função da quantidade de amostras pertencentes a cada i ésimo grupo, ou seja, quanto mais amostras pertencentes existirem em um determinado grupo, maior será probabilidade de uma nova amostra lida pertencer ao grupo com maior probabilidade. Nesta tese, adota-se a distância

de uma amostra lida em relação aos demais grupos para cálculo da probabilidade *a priori*, denominado de **Coefficiente Evolutivo–Recurivo**.

O procedimento denominado de **Coefficiente Evolutivo–Recurivo**, $C_{r_0}(i)$, é uma versão otimizada da probabilidade *a priori* de se selecionar o i ésimo grupo no instante k . O coeficiente avalia a probabilidade de selecionar um grupo a partir das distâncias entre a amostra lida, $\mathbf{z}_{rc}^{(k)}$, e os grupos existentes; de tal modo que, a maior probabilidade indicará que há uma maior proximidade ao centro do grupo. Na iteração k , a atualização do coeficiente evolutivo–recurivo para i ésima regra, $C_{r_0}(i)$, para a condição $D_{(i)k} > 0$, é dado por

$$C_{r_0}(i) = \frac{1}{\sum_{j=1}^c \left(\frac{D_{(i)k}}{D_{(j)k}} \right)^{2/(m-1)}}, \quad (4.40)$$

onde

$$D_{(i)k}^2 = \left(\mathbf{d}_{zv(i)}^{(k)} \right) \left(\mathbf{A}_{\Sigma_{rc(i)}}^{(k-1)} \right)^{-1} \left(\mathbf{d}_{zv(i)}^{(k)} \right)^T \quad (4.41)$$

e

$$D_{(j)k}^2 = \left(\mathbf{d}_{zv(j)}^{(k)} \right) \left(\mathbf{A}_{\Sigma_{rc(j)}}^{(k-1)} \right)^{-1} \left(\mathbf{d}_{zv(j)}^{(k)} \right)^T, \quad (4.42)$$

com

$$\mathbf{d}_{zv(j)}^{(k)} = \mathbf{z}_{rc}^{(k)} - \mathbf{v}_{rc(j)}^{(k-1)}. \quad (4.43)$$

Considerando a equação (4.40) para as condições de $D_{(i)k}$ igual a zero e $j \neq i$ com $i = 1, 2, \dots, c$ e $j = 1, 2, \dots, c$, tem-se

$$C_{r_0}(i) = 1, \quad (4.44)$$

$$C_{r_0}(j) = 0, \quad (4.45)$$

onde, as equações (4.41) e (4.42) representam as distâncias da amostra lida em relação aos centros $\mathbf{v}_{rc(i)}^{(k-1)}$ e $\mathbf{v}_{rc(j)}^{(k-1)}$, no instante k , respectivamente.

Com $C_{r_0}(i) \in [0, 1]$ com $\sum_{i=1}^c C_{r_0}(i) = 1$, satisfazendo os axiomas de probabilidade, com $i = 1, 2, \dots, c$, no instante k (A.PAPOULIS; PILAI, 2002; GHAMRAMANI, 2005; ALBUQUERQUE J. M. P. FORTES, 2008).

Observe que o critério da métrica de distância, equação (4.39), foi modificado pela eliminação do coeficiente exponencial utilizado na norma do algoritmo de agrupamento Gath–Geva (ABONYI et al., 2005), de modo que a parcela,

$$\frac{(2\pi)^{p/2} \sqrt{A_{\Sigma_{rc}}^{(k-1)}}}{C_{r_0}(i)} \quad (4.46)$$

é mais adequada quando usado para comparação entre grupos. Além disso, este critério também pode ser utilizado como grau de pertinência quando se avalia se um ponto pertence ao grupo selecionado pelo coeficiente (SOLEIMANI-B; C.; BABAK, 2010).

4.4.1.2 Estimação das Distâncias

As distâncias dos dados em relação aos i -ésimos grupos selecionados são calculadas a partir da similaridade existente entre a amostra lida, $\mathbf{z}_{rc}^{(k)}$, no instante k , e os respectivos centros dos agrupamentos, $\mathbf{v}_{rc_i}^{(k-1)}$, calculada a partir da norma de máxima verossimilhança, utilizada para estimar a matriz partição nebulosa, é expressa por

$$D_{ikA_{\Sigma rc}}^2 = \frac{(2\pi)^{p/2} \sqrt{A_{\Sigma rc}^{(k-1)}}}{C_{r0}(i)} \exp \left[\frac{1}{2} \left(\mathbf{d}_{zv(i)}^{(k)} \right) \left(A_{\Sigma rc}^{(k-1)} \right)^{-1} \left(\mathbf{d}_{zv(i)}^{(k)} \right)^T \right], \quad \forall k, \quad (4.47)$$

$$D_{jkA_{\Sigma rc}}^2 = \frac{(2\pi)^{p/2} \sqrt{A_{\Sigma rc}^{(k-1)}}}{C_{r0}(j)} \exp \left[\frac{1}{2} \left(\mathbf{d}_{zv(j)}^{(k)} \right) \left(A_{\Sigma rc}^{(k-1)} \right)^{-1} \left(\mathbf{d}_{zv(j)}^{(k)} \right)^T \right], \quad \forall k, \quad (4.48)$$

com $i, j = 1, 2, \dots, c$. O fator exponencial da máxima verossimilhança é utilizado para estimar o grupo mais próximo, ou seja, o grupo com maior similaridade em relação a amostra lida, é expressa por

$$\left\| \mathbf{d}_{zv(i)}^{(k)} \right\|_{A_{\Sigma}} = \left[\left(\mathbf{d}_{zv(i)}^{(k)} \right) \left(A_{\Sigma rc}^{(k-1)} \right)^{-1} \left(\mathbf{d}_{zv(i)}^{(k)} \right)^T \right], \quad \forall k, i. \quad (4.49)$$

4.4.1.3 Estimação do Grupo Mais Próximo

O **grupo mais próximo** é o grupo que apresenta a melhor similaridade existente entre a amostra lida e os grupos existentes e a partir deste **grupo mais próximo** é tomada a decisão da criação de um novo grupo ou se manterá a estrutura atual. Após calcular as distâncias em relação aos i -ésimos centros dos grupos, determina-se o grupo mais próximo e o raio deste grupo é calculado. Isto permite ao algoritmo decidir sobre a criação de um novo grupo, considerando que, se o grupo mais próximo tem o raio inferior à distância da amostra em relação ao centro mais próximo, então a amostra lida não pertence aos grupos existentes e um novo grupo deve ser criado. Identifica-se o grupo mais próximo g_0 , é expresso por

$$g_0 = \arg \min_{i=1,2,\dots,c} \left(\sqrt{\left\| \mathbf{d}_{zv(i)}^{(k)} \right\|_{A_{\Sigma}}} \right). \quad (4.50)$$

4.4.1.4 Estimação do Raio do Grupo Mais Próximo

Ao se determinar o vetor de protótipo do **grupo mais próximo**, \mathbf{v}_{g_0} , calcula-se o raio deste, no instante k , que é dado por

$$r_{g_0} = \operatorname{argmax}_{\forall \mathbf{z}_{rc}^{(j)} \in g_0 \text{ e } \mu_{g_0(j)} > \mu_h \quad j=1,2,\dots,k} \left\| \mathbf{z}_{rc}^{(j)} - \mathbf{v}_{g_0} \right\|_{A_{\Sigma g_0}}. \quad (4.51)$$

Neste trabalho, $\|\cdot\|_{A_{\Sigma g(i)}}$ representa a norma que permite mensurar o raio de grupo com base nas condições de restrições dadas por:

$$\forall \mathbf{z}_{rc}^{(j)} \in g_0 \quad (4.52)$$

para

$$\mu_{g_0(j)} > \mu_h \quad (4.53)$$

com $j = 1, 2, \dots, k$.

A condição de restrição (4.52) diz respeito a todas as amostras que pertencem ao grupo mais próximo determinadas pela condição de restrição da equação (4.53) quando satisfeita, indicará qual delas será utilizada para mensurar o raio deste agrupamento, ou seja, todas as amostras em que seu grau de pertinência, $\mu_{g_0(j)}$, do grupo mais próximo é maior que um grau de pertinência limiar, μ_h , indica que a amostra deve pertencer ao grupo, e portanto, esta será utilizada para mensurar o raio do grupo mais próximo, o fator r_{g_0} .

O valor do raio do agrupamento mais próximo é calculado de acordo com o fator exponencial da máxima verossimilhança, que é expresso por

$$\left\| \mathbf{z}_{rc}^{(j)} - \mathbf{v}_{g_0} \right\|_{A_{\Sigma_{g_0}}} = \left[\left(\mathbf{z}_{rc}^{(j)} - \mathbf{v}_{g_0} \right) \left(A_{\Sigma_{g_0}} \right)^{-1} \left(\mathbf{z}_{rc}^{(j)} - \mathbf{v}_{g_0} \right)^T \right] \quad (4.54)$$

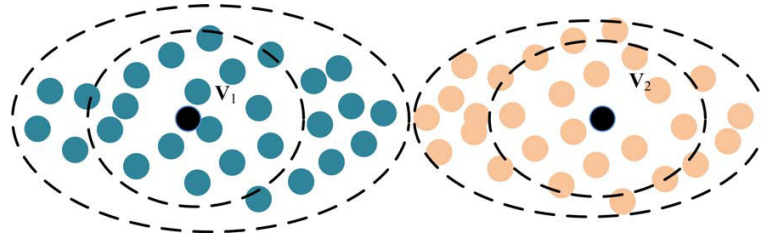
com $j = 1, 2, \dots, k$.

Diferente da metodologia proposta por (ANGELOV; FILEV; KASABOV, 2010) para atualização do raio mais próximo, nesta tese atualiza-se todos os raios dos grupos existentes e então determina-se o raio do grupo mais próximo, considerando que todos os grupos mantêm informações a respeito da dinâmica do conjunto de dados e devem contribuir na formulação dos estimadores paramétricos na etapa de estimação paramétrica evolutiva do consequente.

O grau de pertinência limiar é $\mu_h \in [0, 1]$, o qual é definido por um especialista, sendo que valores típicos são da ordem de 0,5 (ANGELOV; FILEV; KASABOV, 2010). Utilizar valores baixos do grau de pertinência limiar, $u_h < 0,5$, implica que uma amostra poderá pertencer a mais de um grupo no cálculo do raio; para valores altos grau de pertinência limiar, $u_h > 0,5$, implica que amostras pertencentes ao grupo deixarão de fazer parte do cálculo do raio deste. Na Figura 32 é ilustrado um conjunto de dois grupos com seus respectivos vetores de protótipos e a região de similaridade determinada pela linha tracejada. Observe-se que as linhas tracejadas mais próximas dos centros dos grupos sinalizam maior grau de pertinência e as distâncias mais afastadas sinalizam um menor grau de pertinência.

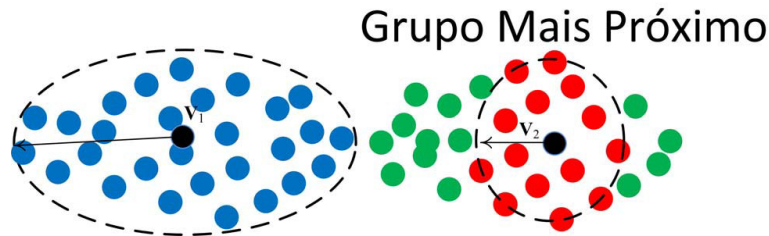
A partir da Figura 32 será ilustrado o efeito do cálculo do raio de um agrupamento em função da escolha do grau de pertinência limiar u_h , conforme se observa nas Figuras 33–35. Ao escolher um grau de pertinência limiar superior ao valor de 0,5; observe-se que algumas amostras pertencentes aos grupos deixarão de fazer parte na estimativa do raio do grupo (amostras sinalizadas com a cor verde), ou seja, o raio é calculado com poucas amostras similares, conforme ilustrado na Figura 33.

Figura 32: Conjunto de dois grupos com seus respectivos vetores de protótipos e a região de similaridade determinada pela linha tracejada.



Fonte: Elaborado pelo autor.

Figura 33: Raio do agrupamento determinado para um grau de pertinência limiar alto $u_h \gg 0,5$.



Fonte: Elaborado pelo autor.

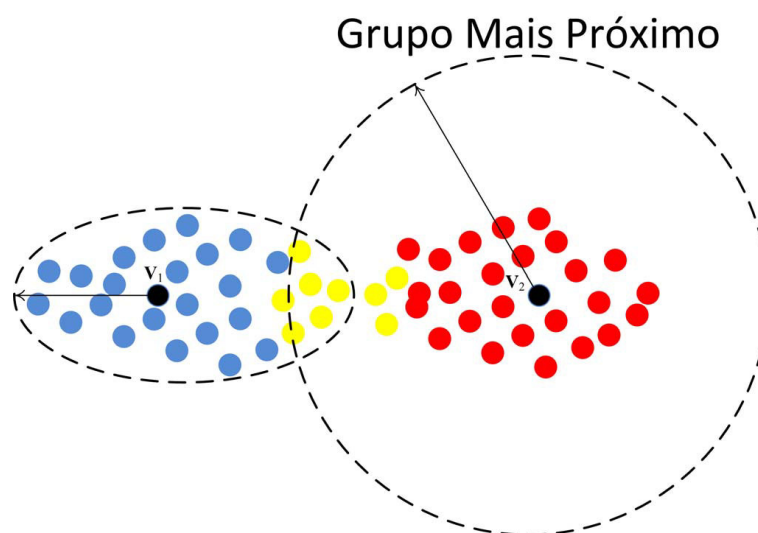
Quando o especialista escolhe um grau de pertinência limiar inferior ao valor de 0,5; observe-se que as amostras com maior similaridade para seu grupo e com menor grau de similaridade para os demais grupos farão parte na estimativa do raio do grupo (amostras sinalizadas com a cor amarela), ou seja, o raio é calculado a partir de amostras com menor grau de similaridade, conforme ilustrado na Figura 34.

Valores típicos de $0,4 \leq u_h \leq 0,6$ obtidos a partir de simulações experimentais demonstraram uma forma ideal para se determinar o raio do agrupamento conforme ilustrado na Figura 35, garante uma aproximação ideal da maior distância existente entre as amostras e seus respectivos grupos pertencentes, o que pode ser comprovado também na literatura, conforme visto em (GEORGIEVA; FILEV, 2009).

4.4.1.5 Atualização dos Parâmetros dos Grupos

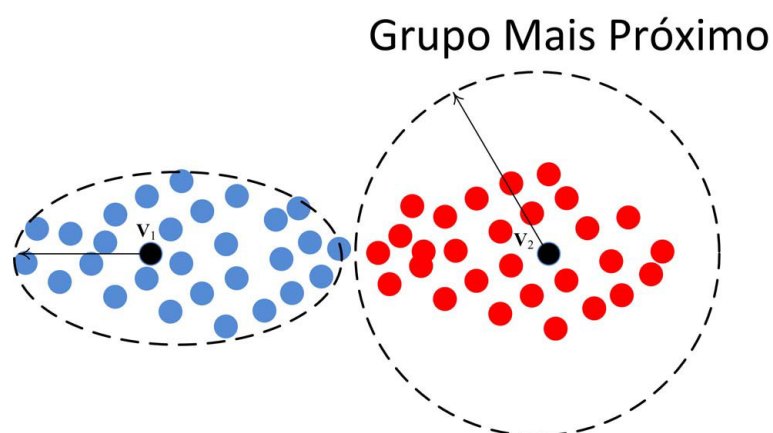
A atualização do vetor de protótipos, $\mathbf{v}_{rc(i)}^{(k)}$; da matriz de covariância nebulosa, $\mathbf{A}_{\Sigma_{rc}}^{(k)}$; e da matriz de partição nebulosa, $\tilde{\mathbf{U}}_{rc}^{(k)}$, conforme proposto por (ANGELOV; FILEV; KASABOV, 2010; GEORGIEVA; FILEV, 2009) baseia-se na regra de *Kohonen* (KOHONEN, 1998), apenas no grupo mais próximo identificado. Nesta tese, atualizam-se todos os parâmetros dos grupos existentes utilizando-se as regras de *Kohonen* em conjunto com

Figura 34: Raio do agrupamento determinado para um grau de pertinência limiar baixo $u_h \ll 0,5$.



Fonte: Elaborado pelo autor.

Figura 35: Raio do agrupamento determinado para um grau de pertinência limiar na faixa de $0,4 \leq u_h \leq 0,6$.



Fonte: Elaborado pelo autor.

um fator de aprendizado participativo, permitindo uma maior consistência na matriz de covariância nebulosa dos dados de cada grupo e um melhor desempenho do modelo final obtido na estimação paramétrica evolutiva do consequente. (PEDRYCZ; GOMIDE, 2007; LIMA et al., 2010; ANGELOV; FILEV; KASABOV, 2010).

A atualização do parâmetros é dada a seguir:

$$\mathbf{v}_{rc(i)}^{(k)} = \mathbf{v}_{rc(i)}^{(k-1)} + \alpha \left(\mathbf{z}^{(k)} - \mathbf{v}_{rc(i)}^{(k-1)} \right), \quad (4.55)$$

$$\mathbf{A}_{\Sigma_{rc}}^{(k)} = (1 - \alpha) \mathbf{A}_{\Sigma_{rc}}^{(k-1)} + \alpha \left(\mathbf{z}^{(k)} - \mathbf{v}_{rc(i)}^{(k-1)} \right)^T \left(\mathbf{z}^{(k)} - \mathbf{v}_{rc(i)}^{(k-1)} \right), \quad (4.56)$$

$$\tilde{\mathbf{U}}_{rc}^{(k)} = \tilde{h} \left(\mathbf{z}_{rc}^{(k)}, c, \tilde{\mathbf{U}}_{rc}^{(k-1)} \right), \quad (4.57)$$

com $i = 1, 2, \dots, c$.

A função $\tilde{h} \left(\mathbf{z}_{rc}^{(k)}, c, \tilde{\mathbf{U}}_{rc}^{(k-1)} \right)$ é a atualização da matriz de partição nebulosa dos c grupos atualizados no instante k . Para a condição de $D_{ikA_{\Sigma_{rc}}} > 0$ o grau de pertinência é expresso por

$$\mu_{rc(i)}^{(k)} = \frac{1}{\sum_{j=1}^c \left(\frac{D_{ikA_{\Sigma_{rc}}}}{D_{jkA_{\Sigma_{rc}}}} \right)^{2/(m-1)}}, \quad (4.58)$$

considerando a equação (4.58) para as condições $D_{ikA_{\Sigma_{rc}}}$ igual a zero e $j \neq i$ com $i = 1, 2, \dots, c$ e $j = 1, 2, \dots, c$, tem-se

$$\mu_{rc(i)}^{(k)} = 1, \quad (4.59)$$

$$\mu_{rc(j)}^{(k)} = 0, \quad (4.60)$$

com, $m > 1$, e $\mu_{rc(i)}^{(k)(l)} \in [0, 1]$ para $\sum_{i=1}^c \mu_{rc(i)}^{(k)(l)} = 1$. A distâncias $D_{ikA_{\Sigma_{rc}}}$ e $D_{jkA_{\Sigma_{rc}}}$ são as distâncias segundo a norma da máxima verossimilhança da amostra lida no instante k em relação aos centros dos grupos i e j , respectivamente.

Define-se particionamento nebuloso, M_{fc} , do conjunto de dados \mathbf{Z}_{rc} , até o instante k como segue:

$$M_{fc} = \left\{ \tilde{\mathbf{U}}_{\mu} \in \mathbb{R}^{c \times k} \mid \mu_{rc(i)}^{(j)} \in [0, 1], \forall i, j; \sum_{i=1}^c \mu_{rc(i)}^{(j)} = 1; \forall j; 0 < \sum_{j=1}^k \mu_{rc(i)}^{(j)} < k, \forall i \right\}. \quad (4.61)$$

4.4.1.6 Fator de Aprendizado Participativo

O processo do fator de aprendizado participativo permite que os dados modulem as taxas de aprendizado, de modo que dados mais próximos aos centros demonstram que há uma maior compatibilidade entre eles. Conseqüentemente o processo de aprendizagem pode ser acelerado. À medida que dados espúrios (deslocados, do inglês *outliers*) aparecem, o fator de

aprendizado é reduzido para que possa dar mais robustez ao aprendizado, ou seja, o processo de aprendizagem é desacelerado. Observa-se que pequenas variações nas amostras lidas não serão tratadas como novos agrupamentos, em virtude da modulação do fator de aprendizado participativo agir no sentido de evitar desvios na orientação do fluxo das amostras lidas. As aplicações cujas amostras diferentes também são consideradas mais importantes do que as demais, o fator de aprendizado participativo não interferirá no processo, em virtude das variações no fluxo de dados serem tratadas como novos agrupamentos e validados de acordo com o fator de credibilidade, ver seção 4.4.1.9 (ANGELOV; FILEV; KASABOV, 2010; LIMA et al., 2010).

As taxas de aprendizado são determinadas da seguinte forma (MACIEL; GOMIDE; BALLINI, 2012):

$$\alpha = \lambda \rho_{(k)}^{1-a_{(k)}}, \quad (4.62)$$

onde $\lambda \in [0, 1]$ é a taxa de aprendizado; o termo $\rho_{(k)}$ é o fator de compatibilidade da observação $\mathbf{z}^{(k)}$, (ANGELOV; FILEV; KASABOV, 2010; CAMINHAS; LEMOS; GOMIDE, 2013). É calculado como segue:

$$\rho_{(k)} = \exp \left[-\frac{1}{2} \left(\mathbf{z}^{(k)} - \mathbf{v}_{rc(i)}^{(k-1)} \right) \left(A_{\Sigma_{rc}}^{(k-1)} \right)^{-1} \left(\mathbf{z}^{(k)} - \mathbf{v}_{rc(i)}^{(k-1)} \right)^T \right] \quad (4.63)$$

e

$$a_{(k)} = (1 - \beta)a_{(k-1)} + \beta(1 - \rho_{(k)}), \quad (4.64)$$

onde $\beta \in [0, 1]$ é uma taxa de aprendizado menor que λ .

4.4.1.7 Atualização e Junção de Grupos

O algoritmo proposto precisa garantir uma maior transparência na criação de regras, todavia, para se evitar a maldição da dimensionalidade na criação de novas regras, faz-se necessário uma avaliação dos grupos já criados e avaliar se os mesmos podem ser agregados a partir de critérios de proximidade, em todas as iterações, de modo que, o menor número de regras seja garantido sem comprometer o desempenho final do algoritmo na etapa de estimação paramétrica evolutiva do consequente.

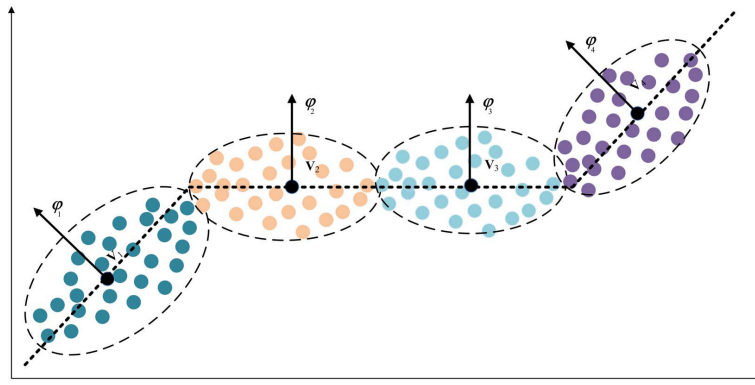
Os elementos chaves para a junção de grupos são os critérios que medem o grau de compatibilidade entre grupos. O grau de compatibilidade é determinado com base nas propriedades geométricas dos grupos, avaliando-se os autovalores e os autovetores da matriz de covariância nebulosa de cada grupo criado no instante k (BABUSKA, 1998; BABUSKA B. SCHUTTER; GUERRA, 2010). Os critérios adotados para atualizar e juntar os grupos são descritos a seguir:

$$s_{(ij)}^{(1)} = \left\| \boldsymbol{\varphi}_{(ip)}^T \boldsymbol{\varphi}_{(jp)} \right\| \geq k_1 \quad \forall i, j = 1, 2, \dots, c, \quad (4.65)$$

$$s_{(ij)}^{(2)} = \left\| \mathbf{v}_{v_i}^* - \mathbf{v}_{v_j}^* \right\| \leq k_2 \quad \forall i, j = 1, 2, \dots, c, \quad (4.66)$$

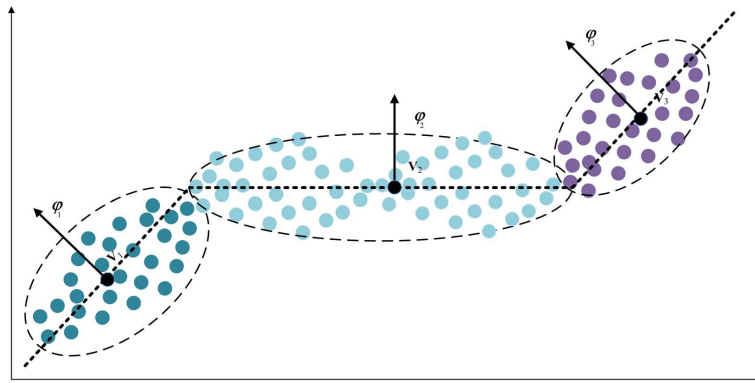
onde $\varphi_{(ip)}^T$ e $\varphi_{(jp)}^T$ são os autovetores unitários da matriz de covariância nebulosa dos grupos i e j a serem avaliados; as constantes k_1 e k_2 são definidas pelo especialista com valores de k_1 próximo a 1 e valores de k_2 próximos a 0 (BABUSKA, 1998); e $\mathbf{v}_{v_i}^*$ e $\mathbf{v}_{v_j}^*$ são os centros dos agrupamentos normalizados, obtidos a partir da equação (2.41). A primeira condição imposta pela equação (4.65) afirma que os grupos devem ser juntados se forem paralelos. A segunda condição imposta pela equação (4.66) avalia a distância existente entre os centros dos protótipos dos grupos candidatos a serem juntados. Conforme ilustrado na Figura 36 que retrata a ideia principal de junção de grupos compatíveis, observam-se que os grupos 2 e 3 apresentam alto grau de compatibilidade que foi estabelecido pelas condições impostas nas equações (4.65) e (4.66). O resultado final do algoritmo de junção de grupos compatíveis podem ser ilustrados na Figura 37 e todos os parâmetros envolvidos atualizados após a nova estrutura resultante.

Figura 36: Grupos compatíveis para atuação do mecanismo de junção de grupos no espaço \mathbb{R}^2 .



Fonte: Elaborado pelo autor.

Figura 37: Resultado final do agrupamento de junção entre grupos compatíveis no espaço \mathbb{R}^2 .



Fonte: Elaborado pelo autor.

O procedimento do algoritmo para junção de grupos compatíveis são definidos a partir das matrizes de agregação dos critérios citados nas equações (4.65) e (4.66), definidas como segue

$$\mathbf{S}^{(1)} = \begin{bmatrix} s_{(11)}^{(1)} & \cdots & s_{(1c)}^{(1)} \\ \vdots & \vdots & \vdots \\ s_{(c1)}^{(1)} & \cdots & s_{(cc)}^{(1)} \end{bmatrix} \quad (4.67)$$

e

$$\mathbf{S}^{(2)} = \begin{bmatrix} s_{(11)}^{(2)} & \cdots & s_{(1c)}^{(2)} \\ \vdots & \vdots & \vdots \\ s_{(c1)}^{(2)} & \cdots & s_{(cc)}^{(2)} \end{bmatrix}. \quad (4.68)$$

As matrizes dos critérios de agregação refinadas utilizadas para gerar os grupos candidatos que compatíveis à atualização e junção, são dadas por:

$$\tilde{\mathbf{S}}^{(1)} = \begin{bmatrix} \tilde{s}_{(11)}^{(1)} & \cdots & \tilde{s}_{(1c)}^{(1)} \\ \vdots & \vdots & \vdots \\ \tilde{s}_{(c1)}^{(1)} & \cdots & \tilde{s}_{(cc)}^{(1)} \end{bmatrix} \quad (4.69)$$

e

$$\tilde{\mathbf{S}}^{(2)} = \begin{bmatrix} \tilde{s}_{(11)}^{(2)} & \cdots & \tilde{s}_{(1c)}^{(2)} \\ \vdots & \vdots & \vdots \\ \tilde{s}_{(c1)}^{(2)} & \cdots & \tilde{s}_{(cc)}^{(2)} \end{bmatrix}, \quad (4.70)$$

onde

$$\tilde{s}_{(ij)}^{(1)} = \exp^{-\left(s_{(ij)}^{(1)}\right)^2} \quad \forall i, j = 1, 2, \dots, c \quad (4.71)$$

e

$$\tilde{s}_{(ij)}^{(2)} = \exp^{-\left(s_{(ij)}^{(2)}\right)^2} \quad \forall i, j = 1, 2, \dots, c. \quad (4.72)$$

Define-se a matriz geométrica de agregação, como segue:

$$s_{(ij)}^{(0)} = \sqrt{\tilde{s}_{(ij)}^{(1)} \tilde{s}_{(ij)}^{(2)}} \quad \forall i, j = 1, 2, \dots, c. \quad (4.73)$$

A matriz final que definirá os grupos a serem agregados é definida como:

$$s_{(ij)}^{(\gamma_c)} = \begin{cases} 1 & \text{SE } s_{(ij)}^{(0)} \geq \gamma_c \\ 0 & \text{CASO CONTRÁRIO} \end{cases}, \quad (4.74)$$

onde $\gamma_c \in (0, 1)$ corresponde a um fator limiar de agregação para duas ou mais regras similares. Os valores de γ_c próximos de zero permitem a agregação de grupos nebulosos (regras nebulosas) cujas distâncias sejam maiores que seus raios; em contrapartida, os valores de γ_c próximos de um permitem a agregação de grupos nebulosos (regras nebulosas) cujas distâncias sejam menores que seus raios. Não há método específico para atribuir valores padrões para γ_c durante a fase de estimação inicial do algoritmo; embora, que

a partir de análises computacionais, verificou-se que valores de $\gamma_{(c)}$ no intervalo de 0,85 a 0,95 demonstraram ser fatores limiares ideais para agregação de regras nebulosas similares (BABUSKA, 1998).

O vetor de protótipos atualizados após a agregação dos grupos similares é o vetor médio dos vetores protótipos dos grupos similares entre si, definidos pelo critério de agregação de similaridade $s_{(ij)}^{(\gamma_{(c)})}$ para $i = 1, 2, \dots, c$ e $j = 1, 2, \dots, c$ com $i \neq j$, no instante k , que é expresso por

$$\mathbf{v}_{rc(i)}^{(k)} = \frac{1}{a_{g(i)}} \left[\mathbf{v}_{rc(i)}^{(k)} + \sum_{j>i}^c \left(\mathbf{v}_{rc(i)}^{(k)} s_{(ij)}^{(\gamma_{(c)})} \right) \right] \quad i = 1, 2, \dots, c, \forall k, \quad (4.75)$$

onde $a_{g(i)} < c$ é o total de grupos agregados ao i ésimo grupo, observando-se que após a agregação final, o algoritmo manterá apenas dois grupos restantes, o mínimo aceitável para a implementação das regras nebulosas da estimação paramétrica evolutiva do consequente.

Após a adaptação da estrutura com os novos grupos agregados, o algoritmo atualizará a matriz de covariância nebulosa e matriz de partição nebulosa dos grupos remanescente. De acordo com (BABUSKA; VEEN; KAYMAK, 2002; BABUSKA B. SCHUTTER; GUERRA, 2010), a nova matriz de covariância nebulosa de inicialização dos grupos é expressa por

$$\mathbf{A}_{\Sigma_{rc(i)}}^{(nova)} = (1 - \gamma_a) \mathbf{A}_{\Sigma_{rc(i)}}^{(k)} + \gamma_a \left[\det \left(\mathbf{A}_{\Sigma_{rc(0)}}^{(k)} \right) \right]^{1/p} \mathbf{I}, \quad (4.76)$$

onde $\gamma_u \in [0, 1]$ é o parâmetro de sintonia de ajuste da matriz de covariância nebulosa de inicialização $\mathbf{A}_{\Sigma_{rc(i)}}^{(nova)}$. Dependendo do valor de γ_a , os grupos são forçados a ter uma forma geométrica idêntica e para valores iguais a 1 torna-se um fator limitante do algoritmo para identificar adequadamente grupos. $\mathbf{A}_{\Sigma_{rc(0)}}^{(k)}$ é a matriz de covariância obtida a partir do conjunto de dados no instante k e independe do total de grupos existentes. \mathbf{I} é matriz identidade de mesma ordem da matriz de covariância nebulosa $\mathbf{A}_{\Sigma_{rc(i)}}^{(k)}$ (BABUSKA; VEEN; KAYMAK, 2002; BABUSKA B. SCHUTTER; GUERRA, 2010). Obtém-se os autovalores $\lambda_{a(ij)}$ e autovetores $\phi_{a(ij)}$ a partir da matriz de covariância nebulosa de inicialização $\mathbf{A}_{\Sigma_{rc(i)}}^{(nova)}$ e determine seus valores máximos dos autovalores, expressos por

$$\lambda_{i(max)} = \max_j \left(\lambda_{a(ij)} \right). \quad (4.77)$$

Após determinar os valores máximos dos autovalores, modifique-os da seguinte forma:

$$\lambda_{a(ij)} = \frac{\lambda_{i(max)}}{\beta_a} \quad \forall j \text{ para todo } \frac{\lambda_{i(max)}}{\lambda_{a(ij)}} > \beta_a. \quad (4.78)$$

Reconstruir a matriz de covariância nebulosa dos i ésimos grupos, para $i = 1, 2, \dots, c$, a partir de:

$$\mathbf{A}_{\Sigma_{rc(i)}}^{(k)} = \left[\phi_{a(i1)} \dots \phi_{a(ip)} \right] \text{diag} \left(\lambda_{a(i1)} \dots \lambda_{a(ip)} \right) \left[\phi_{a(i1)} \dots \phi_{a(ip)} \right]^{-1}. \quad (4.79)$$

Os passos seguintes permitem ao algoritmo refinar os parâmetros dos grupos após a realização do algoritmo de agregação dos grupos similares. O cálculo da probabilidade *a priori* a partir do coeficiente evolutivo–recursivo, para $i = 1, 2, \dots, c$, no instante n , para $n = 1, 2, \dots, k$, é expressa por

$$C_{r0}(i) = \frac{1}{\sum_{j=1}^c \left(\frac{D_{(i)n}}{D_{(j)n}} \right)^{2/(m-1)}}, \quad (4.80)$$

onde

$$D_{(i)n}^2 = \left(\mathbf{d}_{zv(i)}^{(n)} \right) \left(\mathbf{A}_{\Sigma_{rc(i)}}^{(k)} \right)^{-1} \left(\mathbf{d}_{zv(i)}^{(n)} \right)^T \quad (4.81)$$

e

$$D_{(j)n}^2 = \left(\mathbf{d}_{zv(j)}^{(n)} \right) \left(\mathbf{A}_{\Sigma_{rc(j)}}^{(k)} \right)^{-1} \left(\mathbf{d}_{zv(j)}^{(n)} \right)^T, \quad (4.82)$$

com

$$\mathbf{d}_{zv(i)}^{(n)} = \mathbf{z}_{rc}^{(n)} - \mathbf{v}_{rc(i)}^{(k)} \quad (4.83)$$

e

$$\mathbf{d}_{zv(j)}^{(n)} = \mathbf{z}_{rc}^{(n)} - \mathbf{v}_{rc(j)}^{(k)}. \quad (4.84)$$

Os elementos da matriz de partição nebulosa são calculados para $1 \leq i \leq c$, $1 \leq n \leq k$, é expresso por

$$\mu_{rc(i)}^{(n)} = \frac{1}{\sum_{j=1}^c \left(\frac{\left[\frac{\det \left(\mathbf{A}_{\Sigma_{rc(i)}}^{(k)} \right)^{1/2}}{C_{r0(i)}^{(l)}} \exp \frac{1}{2} \left(\mathbf{d}_{zv(i)}^{(n)} \right)^T \left(\mathbf{A}_{\Sigma_{rc(i)}}^{(k)} \right)^{-1} \left(\mathbf{d}_{zv(i)}^{(n)} \right) \right]}{\left[\frac{\det \left(\mathbf{A}_{\Sigma_{rc(j)}}^{(k)} \right)^{1/2}}{C_{r0(j)}^{(l)}} \exp \frac{1}{2} \left(\mathbf{d}_{zv(j)}^{(n)} \right)^T \left(\mathbf{A}_{\Sigma_{rc(j)}}^{(k)} \right)^{-1} \left(\mathbf{d}_{zv(j)}^{(n)} \right) \right]} \right)}. \quad (4.85)$$

Define-se particionamento nebuloso, M_{fc} , do conjunto de dados \mathbf{Z}_{rc} , até o instante k , que é expressa por

$$M_{fc} = \left\{ \tilde{\mathbf{U}}_{\mu} \in \mathbb{R}^{c \times k} \mid \mu_{rc(i)}^{(j)} \in [0, 1], \forall i, j; \sum_{i=1}^c \mu_{rc(i)}^{(j)} = 1; \forall j; 0 < \sum_{j=1}^k \mu_{rc(i)}^{(j)} < k, \forall i \right\}. \quad (4.86)$$

O vetor de protótipos atualizado para $i = 1, 2, \dots, c$, é expresso por

$$\mathbf{v}_{rc(i)}^{(k)} = \frac{\sum_{n=1}^k \left[\left(\mu_{rc(i)}^{(n)} \right)^{w_g} \mathbf{z}_{rc}^{(n)} \right]}{\sum_{n=1}^k \left(\mu_{rc(i)}^{(n)} \right)^{w_g}}. \quad (4.87)$$

O cálculo da matriz de covariância nebulosa refinada após os ajustes do centros dos grupos e da matriz de partição nebulosa, é expresso por

$$\mathbf{A}_{\Sigma_{rc(i)}}^{(k)} = \frac{\sum_{n=1}^k \left[\left(\mu_{rc(i)}^{(n)} \right)^{w_g} \left(\mathbf{d}_{zv(i)}^{(n)} \right)^T \left(\mathbf{d}_{zv(i)}^{(n)} \right) \right]}{\sum_{n=1}^k \left(\mu_{rc(i)}^{(n)} \right)^{w_g}}. \quad (4.88)$$

Após determinar a matriz de covariância nebulosa a partir da equação (4.88), utiliza-se o refinamento final desta a partir das equações (4.76) a (4.79).

4.4.1.8 Passos do Algoritmo de Junção dos Grupos

O algoritmo para atualização e junção de grupos é representado pelos passos descritos a seguir:

-
- 1: **enquanto** Grupos Identificados > 2 **É VERDADEIRO faça**
 - 2: Calcular critérios de identificação através das equações (4.65)–(4.66).
 - 3: **se** Há Grupos Similares **É VERDADEIRO então**
 - 4: Construir as matrizes de agregação dos critérios através das equações (4.67)–(4.68).

 - 5: Refinar as matrizes de agregação dos critérios através das equações (4.69)–(4.72).
 - 6: Construir a matriz geométrica de agregação através da equação (4.73).
 - 7: Construir a matriz de sinalização de grupos similares através da equação (4.74).
 - 8: Agregar vetor de protótipos dos grupos similares através da equação (4.75).
 - 9: Refinar a matriz de covariância nebulosa através das equações (4.76)–(4.79).
 - 10: Atualizar a probabilidade *a priori* através do coeficiente evolutivo–recursivo através das equações (4.80)–(4.84).
 - 11: Atualizar a matriz de partição nebulosa através das equações (4.85)–(4.86).
 - 12: Atualizar o vetor de protótipos através da equação (4.87).
 - 13: Atualizar a matriz de covariância nebulosa através da equação (4.88).
 - 14: Refinar a matriz de covariância nebulosa através das equações (4.76)–(4.79).
 - 15: Os parâmetros dos grupos: vetor de protótipos, matriz de covariância nebulosa e matriz de partição nebulosa são atualizados.
 - 16: **fim se**
 - 17: **fim enquanto**
-

4.4.1.9 Fator de Credibilidade

A fim de garantir a credibilidade do grupo recém-criado, um parâmetro $P_n(g_0)$ é introduzido. O fator de credibilidade $P_n(g_0)$ é definido como o número de amostras pertencentes ao grupo mais próximo da amostra lida no instante k . Utiliza-se este parâmetro para avaliar se o número de amostras pertencentes ao grupo mais próximo está dentro de um limite máximo aceitável. Este limite garante que há um número considerável de amostras por grupo e garante a robustez com respeito as amostras espúrias, como também valida a matriz de covariância do conjunto de amostras (ANGELOV; FILEV; KASABOV, 2010).

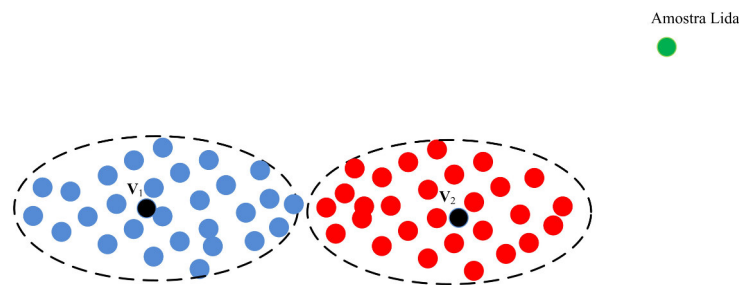
O limite mínimo determinado em função da dimensionalidade do vetor de dados é

expressa por

$$P_{tol} = \frac{p(p+1)}{2}. \tag{4.89}$$

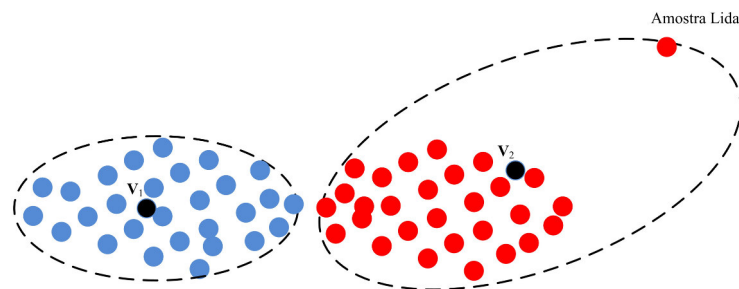
Um novo grupo será criado e validado se a condição $P_n(g_0) > P_{tol}$ for satisfeita; caso contrário a criação de um novo grupo será rejeitado mantendo-se a estrutura antiga. O procedimento é ilustrado nas Figuras 38–40.

Figura 38: Uma amostra é lida e verifica-se se a condição do fator de credibilidade é satisfeita.



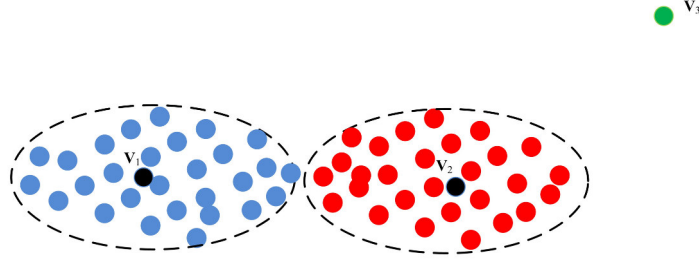
Fonte: Elaborado pelo autor.

Figura 39: A condição do fator de credibilidade não é satisfeita e a amostra lida é incorporada ao grupo mais próximo.



Fonte: Elaborado pelo autor.

Figura 40: A condição do fator de credibilidade é satisfeita e um novo grupo é criado a partir da amostra lida.



Fonte: Elaborado pelo autor.

4.4.2 Estimação Paramétrica Evolutiva do Consequente

Inicialmente, seja o vetor de regressores de dados do sistema dinâmico no antecedente, expressos a seguir

$$\mathbf{x}_{rc}^{(k-1)} = \left[y_{rc}^{(k-1)} \quad \dots \quad y_{rc}^{(k-n_y)} \quad u_{rc}^{(k-1)} \quad \dots \quad u_{rc}^{(k-n_u)} \right]^T. \quad (4.90)$$

A matriz de regressores é dada por

$$\mathbf{X}_{rc} = \left[\mathbf{x}_{rc}^{(1)} \quad \mathbf{x}_{rc}^{(2)} \quad \dots \quad \mathbf{x}_{rc}^{(N_{rc})} \right]^T. \quad (4.91)$$

Seja a variável de entrada auxiliar escrita na forma vetorial obtida a partir da equação (3.50) do modelo modelo nebuloso auxiliar, expressa por

$$\tilde{\mathbf{u}}_{rc}^{(k)} = \left[\tilde{u}_{rc}^{(k-1)} \quad \tilde{u}_{rc}^{(k-2)} \quad \dots \quad \tilde{u}_{rc}^{(k-n_u)} \right]^T. \quad (4.92)$$

Seja a variável de saída auxiliar escrita na forma vetorial obtida a partir da equação (3.44) do modelo modelo nebuloso auxiliar, expressa por

$$\tilde{\mathbf{y}}_{rc}^{(k)} = \left[\tilde{y}_{rc}^{(k-1)} \quad \tilde{y}_{rc}^{(k-2)} \quad \dots \quad \tilde{y}_{rc}^{(k-n_y)} \right]^T. \quad (4.93)$$

A variável instrumental nebulosa, equação (3.83), nos termos das equações (4.90), (4.92) e (4.93) é reescrita da seguinte forma

$$\tilde{\mathbf{x}}_{rc}^{(k-1)} = \left[\tilde{y}_{rc}^{(k-1)} \quad \tilde{y}_{rc}^{(k-2)} \quad \dots \quad \tilde{y}_{rc}^{(k-n_y)} \quad \tilde{u}_{rc}^{(k-1)} \quad \tilde{u}_{rc}^{(k-2)} \quad \dots \quad \tilde{u}_{rc}^{(k-n_u)} \right]^T. \quad (4.94)$$

A variável instrumental nebulosa (4.94) escrita na forma vetorial é dada por

$$\tilde{\mathbf{X}}_{rc} = \left[\tilde{\mathbf{x}}_{rc}^{(1)} \quad \tilde{\mathbf{x}}_{rc}^{(2)} \quad \tilde{\mathbf{x}}_{rc}^{(3)} \quad \dots \quad \tilde{\mathbf{x}}_{rc}^{(N)} \right]^T. \quad (4.95)$$

A matriz de pesos nebulosos da estimação paramétrica recursiva do consequente é obtida a partir da matriz de partição do agrupamento evolutivo, realizado nas N_{rc}

observações do conjuntos de dados como segue:

$$\tilde{\mathbf{U}}_{rc} \in \mathbb{R}^{c \times N_{rc}} \left| \begin{array}{l} \mu_{rci}^{(k)} \in [0, 1], \forall i, k; \sum_{i=1}^c \mu_{rci}^{(k)} = 1, \forall k; 0 < \sum_{k=1}^{N_{rc}} \mu_{rci}^{(k)} < N_{rc}, \forall i, \end{array} \right. \quad (4.96)$$

onde o grau de pertinência da amostra $\{(\mathbf{x}_{rc}^{(k)}, y_{rc}^{(k)})\}$ é definido como,

$$\mu_i^{(k)} = \mu_i^{(k)} \left(\{(\mathbf{x}_{rc}^{(k)}, y_{rc}^{(k)})\} \right) \forall k \quad (4.97)$$

e a matriz de pesos nebulosos para a amostra $\{(\mathbf{x}_{rc}^{(k)}, y_{rc}^{(k)})\}$ é definida como segue

$$\tilde{\mathbf{W}}_{rc}^{(i)} = \text{diag}(\mu_i^{(k)}) \in \mathbb{R}^{N \times N}, \quad (4.98)$$

onde

$$\tilde{\mathbf{W}}_{rc}^{(i)} = \begin{pmatrix} \tilde{w}_{rc}^{(1)} = \mu_i^{(1)} & 0 & \dots & 0 \\ 0 & \tilde{w}_{rc}^{(2)} = \mu_i^{(2)} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \tilde{w}_{rc}^{(N_{rc})} = \mu_i^{(N_{rc})} \end{pmatrix}, \quad (4.99)$$

para $i = 1, 2, \dots, c; k = 1, 2, \dots, N_{rc}$, onde,

O refinamento do vetor de parâmetros do consequente é dado por

$$\tilde{\Theta}_{rc}^{(k)(i)} = \tilde{\Theta}_{rc}^{(k-1)(i)} + \mathbf{K}_{rc}^{(k)} \left(y_{rc}^{(k)} - \mathbf{x}_{rc}^{(k)T} \tilde{\Theta}_{rc}^{(k-1)(i)} \right), \quad (4.100)$$

onde o ganho de adaptação do vetor estimado é

$$\mathbf{K}_{rc}^{(k)} = \frac{\tilde{w}_{rc}^{(k)} \mathbf{P}_{rc}^{(k-1)} \tilde{\mathbf{x}}_{rc}^{(k)}}{\left(\delta_{rc}^{(k)} + \tilde{w}_{rc}^{(k)} \mathbf{x}_{rc}^{(k)T} \mathbf{P}_{rc}^{(k-1)} \tilde{\mathbf{x}}_{rc}^{(k)} \right)} \quad (4.101)$$

A matriz de covariância é expressa por

$$\mathbf{P}_{rc}^{(k)} = \frac{1}{\delta_{rc}^{(k)}} \left[\mathbf{P}_{rc}^{(k-1)} - \mathbf{K}_{rc}^{(k)} \mathbf{x}_{rc}^{(k)T} \mathbf{P}_{rc}^{(k-1)} \right], \quad (4.102)$$

onde $\tilde{w}_{rci}^{(k)} \Big|_{i=1,2,\dots,c}$ é a atualização dos pesos nebulosos a cada nova iteração para a i ésima regra.

O fator de esquecimento atualizado a cada nova iteração do algoritmo recursivo é dado por

$$\delta_{rc}^{(k)} = \lambda_0 \delta_{rc}^{(k-1)} + (1 - \lambda_0)(\delta_0) - \sigma_{rc}^{(k-1)} \quad (4.103)$$

com $i = 1, 2, \dots, c$, e $k = 1, 2, \dots, N_{rc}$.

A variância móvel do erro de estimação a cada nova iteração é expressa por

$$\sigma_{rc}^{(k)} = \sigma_{rc}^{(k-1)} - \frac{1}{k} \left[\sigma_{rc}^{(k-1)} - \left(y_{rc}^{(k)} - \hat{y}_{rc}^{(k)} - \widehat{m}e_{rc}^{(k)} \right)^2 \right], \quad (4.104)$$

tal que

$$\widehat{m}e_{rc}^{(k)} = \frac{1}{k} \sum_{j=1}^k \left(y_f^{(j)} - \hat{y}_{rc}^{(j)} \right), \quad (4.105)$$

onde $y^{(k)}$ e $\hat{y}_{rc}^{(k)}$ são as saídas real e estimada do modelo no instante k .

4.4.2.1 Modelo Nebuloso Takagi–Sugeno

A estrutura nebulosa Takagi–Sugeno utilizada no algoritmo evolutivo é dada pelo conjunto de regras nebulosas SE–ENTÃO, como segue:

$$\mathbf{R}_i : \mathbf{SE} \left(\eta_1^{(k)} \text{ é } \Gamma_1^i \right) \text{ E } \dots \text{ E } \left(\eta_n^{(k)} \text{ é } \Gamma_n^i \right) \mathbf{ENTÃO} y_i^{(k)} = \sum_{j=1}^n \zeta_{\eta_j} \eta_j^{(k-j)}. \quad (4.106)$$

onde $i = 1, \dots, N_{regras}$; N_{regras} é o número de regras SE–ENTÃO do modelo evolutivo nebuloso; $\boldsymbol{\eta}^{(k)} = [\eta_1^{(k)} \dots \eta_p^{(k)}]$ são as variáveis do antecedente no instante k ; p é a dimensionalidade do vetor do antecedente; $\Gamma_{1,\dots,p}^i$ são os conjuntos nebulosos do antecedente; ζ_{η_j} correspondem aos parâmetros do consequente $j = 1, 2, \dots, p$; e $y_i^{(k)}$ é a saída na i -ésima regra.

4.4.3 Passos do Algoritmo de Estimação Evolutiva–Recursiva

O algoritmo é representado pelos passos descritos a seguir:

-
- 1: Definir os parâmetros da etapa de pré-filtragem: ω_u e ω_y .
 - 2: Definir o parâmetro: grau de pertinência limiar μ_h .
 - 3: Definir os parâmetros: taxas de aprendizagem λ e β .
 - 4: Definir os parâmetros: critérios de agregação de regras γ_c, k_1, k_2 .
 - 5: Definir os parâmetros: taxas de aprendizagem do algoritmo recursivo de estimação paramétrica γ_0, λ_0 .
 - 6: $k \leftarrow 0$.
 - 7: **repetir**
 - 8: $k \leftarrow k + 1$.
 - 9: Ler amostra $\mathbf{z}_{rc}^{(k)}$.
 - 10: Normalizar a amostra lida de acordo com as equações (2.42), (2.43) e (2.44).
 - 11: Calcular o coeficiente recursivo–evolutivo através das equações (4.39)–(4.42).
 - 12: Calcular as distâncias da amostra lida, $\mathbf{z}_{rc}^{(k)}$, aos centros dos grupos através da equação (4.47).
 - 13: Determinar o centro do grupo mais próximo através da equação (4.50).
 - 14: Atualizar o raio do i -ésimo grupo para $i = 1, 2, \dots, c$, através das equações (4.51)–(4.54).
 - 15: Determinar o raio do grupo mais próximo r_{g_0} .
 - 16: **se** $\sqrt{D_{ikA_{\Sigma_{rc}}}^2} \leq r_{g_0}$ **então**
 - 17: Manter a estrutura de grupos e regras nebulosas.
 - 18: Atualizar o vetor de protótipos através da equação (4.55).
 - 19: Atualizar a matriz de covariância nebulosa através da equação (4.56).
 - 20: Atualizar a matriz de partição nebulosa através das equações (4.57)–(4.59).
 - 21: Atualizar o fator de aprendizagem participativo através das equações (4.62)–(4.64).

- 22: **senão**
- 23: Determinar o número de amostras pertencentes ao grupo mais próximo $P_n(g_0)$
- 24: Determinar o fator de credibilidade limiar P_{tol}
- 25: **se** $P_n(g_0) < P_{tol}$ **então**
- 26: Manter a estrutura anterior de grupo e regras.
- 27: Atualizar o vetor de protótipos através da equação (4.55).
- 28: Atualizar a matriz de covariância nebulosa através da equação (4.56).
- 29: Atualizar a matriz de partição nebulosa através das equações (4.57)–(4.59).
- 30: Atualizar o fator de aprendizagem participativo através das equações (4.62)–(4.64).
- 31: **senão se** $P_n(g_0) \geq P_{tol}$ **então**
- 32: Criar um novo grupo e uma nova regra nebulosa.
- 33: Atualizar o número de regras nebulosas.
- 34: $c \leftarrow c + 1$
- 35: Determinar o centro do grupo recém-criado $\mathbf{v}_{rc(c)}^{(k)} \leftarrow \mathbf{z}_{rc}^{(k)}$
- 36: Atualizar a matriz de covariância do novo grupo a partir da matriz de covariância do grupo mais próximo.
- 37: Atualizar a matriz de partição nebulosa através das equações (4.57)–(4.59).
- 38: Atualizar o fator de aprendizagem participativo através das equações (4.62)–(4.64).
- 39: **fim se**
- 40: **fim se**
- 41: **enquanto** Grupos similares > 2 **É VERDADEIRO faça**
- 42: Identificar grupos similares.
- 43: Atualizar grupos muitos próximos através das equações (4.65)–(4.73).
- 44: Agregar grupos e regras similares.
- 45: **fim enquanto**
- 46: **se** Grupos Atualizados **então**
- 47: Atualizar o vetor de protótipos através da equação (4.55).
- 48: Atualizar a matriz de covariância nebulosa através da equação (4.56).
- 49: Atualizar a matriz de partição nebulosa através das equações (4.57)–(4.59).
- 50: Atualizar o fator de aprendizagem participativo através das equações (4.62)–(4.64).
- 51: **fim se**
- 52: Atualizar a variável instrumental nebulosa recursiva através das equações (4.90)–(4.95).
- 53: Atualizar a matriz nebulosa de peso através das equações (4.96)–(4.99).
- 54: Refinar os parâmetros do consequente do modelo final através das equações (4.100)–(4.105).

55: **até que** Há amostra para ser processada

56: Finalizar.

Este capítulo introduziu a metodologia proposta nesta tese, um algoritmo nebuloso evolutivo de máxima verossimilhança, inserido no contexto de variável instrumental nebulosa para estimação paramétrica do modelo do conseqüente.

5 RESULTADOS COMPUTACIONAIS E EXPERIMENTAIS

Neste capítulo, são apresentados os resultados experimentais e computacionais com o objetivo de avaliar o desempenho da metodologia proposta no capítulo 4. As métricas de avaliação amplamente disponíveis na literatura adotadas neste capítulo são descritas na seção 5.1. A análise da variável instrumental nebulosa adotada nesta tese é apresentada na seção 5.3. Os modelos e algoritmos são avaliados em problemas de identificação de sistemas, sendo: a seção 5.5 descreve os resultados da avaliação da identificação de um sistema não-linear; a seção 5.6 descreve os resultados da modelagem caixa preta de um helicóptero com dois graus de liberdade.

5.1 Métricas de Avaliação

Para se avaliar os resultados numéricos dos problemas de predição e identificação, verificou-se a necessidade da análise do seu desempenho em relação à série de dados originais; além de se comparar com outras metodologias e técnicas de identificação de sistemas (BABUSKA B. SCHUTTER; GUERRA, 2010; SKRJANC; DOVZAN; GOMIDE, 2014). As métricas utilizadas são descritas nesta seção:

- i) **Índice de Desigualdade de Theil** (U_{THEIL}): ideal para se avaliar o ajustamento de uma série de dados estimada em relação à série de dados original. Identifica-se quanto mais próximo à zero, mais ajustada é a série de dados estimada em relação a sua série de dados original (MORETTIN, 2006; LEMOS; CAMINHAS; GOMIDE, 2011). O índice é dado por:

$$(U_{THEIL}) = \frac{\sqrt{\frac{1}{N_y} \sum_{k=1}^{N_y} (y^{(k)} - \hat{y}^{(k)})^2}}{\sqrt{\frac{1}{N_y} \sum_{k=1}^{N_y} (y^{(k)})^2} + \sqrt{\frac{1}{N_y} \sum_{k=1}^{N_y} (\hat{y}^{(k)})^2}} \quad (5.1)$$

- ii) **Soma dos Quadrados dos Erros (SSE)**: calcular a soma quadrática dos erros da série de dados estimada em relação à sua série original de dados (MORETTIN, 2006; LEMOS; CAMINHAS; GOMIDE, 2011). A SSE é dado por:

$$SSE = \sum_{k=1}^{N_y} (\hat{y}^{(k)} - y^{(k)})^2 \quad (5.2)$$

- iii) **Erro Quadrático Médio (MSE)**: calcular a média dos erros quadráticos (MORETTIN, 2006; LEMOS; CAMINHAS; GOMIDE, 2011). O MSE é dado por:

$$MSE = \frac{SSE}{N_y} \quad (5.3)$$

- iv) **Raiz do Erro Quadrático Médio (RMSE)**: avaliar a amplitude do erro médio quadrático da série de dados estimada em relação à sua série de dados original (MORETTIN, 2006; LEMOS; CAMINHAS; GOMIDE, 2011). A RMSE é dado por:

$$RMSE = \sqrt{MSE} \quad (5.4)$$

- v) **Índice de Erro Não-Dimensional (NDEI)**: avaliar o erro independente da escala de dados, defini-se como sendo a razão entre o RMSE e o desvio padrão (**std**) da série de dados original (LEMOS; CAMINHAS; GOMIDE, 2011). O índice NDEI é dado por:

$$NDEI = \frac{RMSE}{std(y^{(k)})} \quad (5.5)$$

- vi) **Variance Accounted For (VAF)**: avaliar variância (**var**) do erro em relação a variância da série de dados original (SERRA; BOTTURA, 2007). O índice VAF dado por:

$$VAF(\%) = 100 \times \left[1 - \frac{var(\mathbf{y} - \hat{\mathbf{y}})}{var(\mathbf{y})} \right] \quad (5.6)$$

onde N_y é o tamanho da série de dados original, $y^{(k)}$ é a saída de dados da série original e $\hat{y}^{(k)}$ é a saída da série estimada no instante k .

5.2 Coeficiente de Correlação

A análise de correlação é uma ferramenta estatística útil que é utilizada em séries de dados distintas, a fim de se obter o grau de correlação existentes entre elas. Através desta análise bivariada, é possível medir os pontos fortes de associação entre suas séries de dados distintas e a direção do relacionamento entre elas. A relação é medida a partir de um grau que varia de +1 a -1 e quanto mais próximo de ± 1 indica um grau perfeito de associação entre elas (XIAO et al., 2016).

A força da correlação pode ser descrita da seguinte maneira:

Tabela 1: Força da correlação entre duas séries de dados

Coefficiente ρ_c	Força
$0.00 \leq \rho_c < 0.20$	Muito Fraco
$0.20 \leq \rho_c < 0.40$	Fraco
$0.40 \leq \rho_c < 0.60$	Moderado
$0.60 \leq \rho_c < 0.80$	Forte
$0.80 \leq \rho_c < 1.00$	Muito Forte

Fonte: Elaborado pelo autor

Os coeficiente de correlação utilizados para avaliação deste grau de força são:

- a) **Coefficiente de Pearson:** Utilizado para a avaliar a correlação existente entre duas séries dados que estão relacionadas linearmente (XIAO et al., 2016). Sejam $\mathbf{x} = [x^{(1)} \ x^{(2)} \ \dots \ x^{(N)}]$ e $\mathbf{y} = [y^{(1)} \ y^{(2)} \ \dots \ y^{(N)}]$ duas séries de dados distintas, o índice de correlação é dado por:

$$\rho_c = \frac{N \sum_{k=1}^N x^{(k)} y^{(k)} - \sum_{k=1}^N x^{(k)} \sum_{k=1}^N y^{(k)}}{\sqrt{N \sum_{k=1}^N (x^{(k)})^2 - \left(\sum_{k=1}^N x^{(k)}\right)^2} \sqrt{N \sum_{k=1}^N (y^{(k)})^2 - \left(\sum_{k=1}^N y^{(k)}\right)^2}}. \quad (5.7)$$

- b) **Coefficiente de Kendall:** Teste não-paramétrico que mede a força de dependência entre duas séries de dados. Avalia-se o total de pares de dados **concordantes** e o total de pares **discordantes** (XIAO et al., 2016). O índice de correlação é dado por:

$$\rho_c = \frac{(\text{número de pares concordantes}) - (\text{número de pares discordantes})}{N(N-1)/2}, \quad (5.8)$$

$$(\text{número de pares concordantes}) \text{ se } x^{(i)} > x^{(j)} \text{ e } y^{(i)} > y^{(j)} \text{ para } i \neq j \quad (5.9)$$

e

$$(\text{número de pares discordantes}) \text{ se } x^{(i)} > x^{(j)} \text{ e } y^{(i)} < y^{(j)} \text{ para } i \neq j. \quad (5.10)$$

- c) **Coefficiente de Spearman:** A correlação de *Spearman* é um teste não-paramétrico que é usado para medir o grau de associação entre duas variáveis. O teste de correlação de classificação *Spearman* não assume quaisquer pressupostos sobre a distribuição dos dados e é a análise de correlação apropriada quando as variáveis são medidas em

uma escala pelo menos ordinária (XIAO et al., 2016). O índice de correlação é dado por:

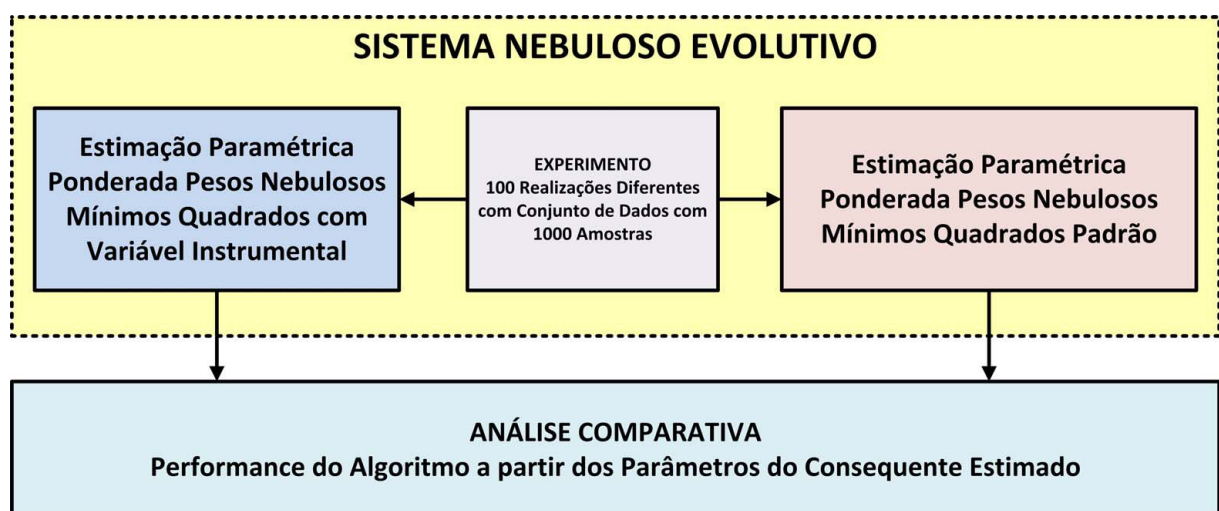
$$\rho_c = 1 - \frac{6 \sum_{k=1}^N \left(\text{rank}(x^{(k)}) - \text{rank}(y^{(k)}) \right)}{N(N^2 - 1)}, \quad (5.11)$$

onde $\text{rank}(x^{(k)})$ é a posição de $x^{(k)}$ na série \mathbf{x} e $\text{rank}(y^{(k)})$ é a posição de $y^{(k)}$ na série \mathbf{y} .

5.3 Análise da Variável Instrumental Nebulosa

Nesta seção, a análise de desempenho do uso da variável instrumental nebulosa não-recursiva (batelada) e variável instrumental nebulosa recursiva inserida no método dos mínimos quadrados é comparada com o desempenho do métodos dos mínimos quadrados tradicional. A metodologia consiste em várias simulações computacionais de um sistema dinâmico não linear e após as realizações computacionais é realizado uma comparação estatística dos parâmetros do consequente obtidos a partir do algoritmo evolutivo utilizando o método dos mínimos quadrados ponderados e do método da variável instrumental nebulosa ponderada, conforme ilustrado na Figura 41. A simulação é realizada em um sistema dinâmico de quarta ordem amplamente utilizado na literatura (ZHU, 2011; YOUNG, 2015), cujo objetivo é a validação e análise da metodologia adotada com variável instrumental nebulosa inserida no contexto evolutivo.

Figura 41: Diagrama de bloco da análise dos parâmetros do experimento baseado nos métodos dos mínimos quadrados e variável instrumental.



Fonte: Elaborado pelo autor.

O modelo adotado é expresso por

$$y^{(k)} = x_r^{(k)} + \epsilon^{(k)}, \quad (5.12a)$$

$$x_r^{(k)} = \frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})} u^{(k)}, \quad (5.12b)$$

$$\epsilon^{(k)} = \frac{\mathbf{C}(z^{-1})}{\mathbf{D}(z^{-1})} e^{(k)}, \quad (5.12c)$$

$$\mathbf{A}(z^{-1}) = 1 - 1.5z^{-1} + 0.7z^{-2} + 0.3z^{-3} - 0.2z^{-4}, \quad (5.12d)$$

$$\mathbf{B}(z^{-1}) = z^{-1} + 0.5z^{-2} - 2.0z^{-3} + z^{-4}, \quad (5.12e)$$

$$\mathbf{C}(z^{-1}) = 1 - 0.6z^{-1} + 0.4z^{-2}, \quad (5.12f)$$

$$\mathbf{D}(z^{-1}) = 1 - 0.95z^{-1} + 0.9506z^{-2}, \quad (5.12g)$$

$$e^{(k)} = \mathbf{N}(0, \sigma_e^2). \quad (5.12h)$$

Foram executadas 100 realizações computacionais de modo que o valor de $\sigma_e^2 \in [0,11 \dots 0,20]$ é obtido aleatoriamente para cada realização individual. A relação sinal/ruído para cada realização individual é de $[35 \dots 10]$ dB, conforme observadas nas Figuras 42–43. A simulação computacional é composta por um conjunto de dados de 1000 amostras. O sinal de entrada $u^{(k)}$ é um sinal pseudo-aleatório compreendido entre -3 a $+3$.

De acordo com o capítulo 3, a variável instrumental deve ser bem correlacionada com o vetor de regressores e totalmente descorrelacionada com a variável do erro inserido no sistema. A partir dos índices de correlação médio de cada realização computacional, pode-se observar que a variável instrumental utilizada demonstrou um alto grau de correlacionamento com o vetor de regressores do sistema, baseado nos índices de correlação de *Kendall* (valor médio igual a 0,89), *Pearson* (valor médio igual a 0,99) e *Spearman* (valor médio igual a 0,98). Também, percebe-se que a variável instrumental utilizada demonstrou um alto grau de descorrelacionamento com o erro inserido no sistema, baseado nos índices de correlação de *Kendall* (valor médio igual a 0,03), *Pearson* (valor médio igual a 0,04) e *Spearman* (valor médio igual a 0,04), conforme ilustrado nas Figuras 44–45.

O modelo de inferência nebuloso *Takagi–Sugeno* utilizado para identificação, é descrito a seguir:

$$R_{(i)} : \text{SE } x^{(k)} \text{ É } \Gamma_{x^{(k)}} \text{ ENTÃO } y_{(i)}^{(k)} = b_4^{(i)} u^{(k-4)} + b_3^{(i)} u^{(k-3)} + b_2^{(i)} u^{(k-2)} + b_1^{(i)} u^{(k-1)} + \dots$$

$$\dots + a_4^{(i)} y^{(k-4)} + a_3^{(i)} y^{(k-3)} + a_2^{(i)} y^{(k-2)} + a_1^{(i)} y^{(k-1)}. \quad (5.13)$$

A estrutura identificou dois grupos de regras nebulosas distintas, ou seja, $i = 1, 2$; total de amostras do conjunto de dados igual a 1000, $k = 1, 2, \dots, 1000$; $\Theta = [a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4]$ são os parâmetros do conseqüente a serem identificados.

O experimento para obtenção dos parâmetros estimados a partir dos métodos do mínimos quadrados de forma não–recursiva e recursiva e comparados com o método da variável instrumental nebulosa de forma não–recursiva e recursiva é realizado nos seguintes passos:

-
- 1: Definir o total de grupos a partir do algoritmo nebuloso evolutivo baseado em máxima verossimilhança.
 - 2: Realização $k = 1$
 - 3: **enquanto** $k \leq 100$ **É VERDADEIRO faça**
 - 4: Determine o valor σ_e^2 para a equação (5.12h).
 - 5: Realização $l = 1$
 - 6: **enquanto** $l \leq 100$ **É VERDADEIRO faça**
 - 7: Realize o agrupamento em batelada em máxima verossimilhança de acordo com os grupos já identificados.
 - 8: Determine a matriz de pesos nebulosos.
 - 9: Realizar o algoritmo de mínimos quadrados ponderado em batelada.
 - 10: Estimar os parâmetros a partir do algoritmo de mínimos quadrados ponderado com os pesos nebulosos em batelada.
 - 11: Realizar o algoritmo de variável instrumental nebulosa ponderado em batelada.
 - 12: Estimar os parâmetros a partir do algoritmo de variável instrumental nebulosa com os pesos nebulosos em batelada.
 - 13: Realizar o algoritmo de mínimos quadrados ponderado recursivo.
 - 14: Inicializar o algoritmo de mínimos quadrados recursivos a partir dos parâmetros estimados em batelada.
 - 15: Realizar o algoritmo de variável instrumental nebulosa ponderado recursivo.
 - 16: Inicializar o algoritmo de mínimos quadrados recursivos de variável instrumental nebulosa a partir dos parâmetros estimados em batelada.
 - 17: $l \leftarrow l + 1$
 - 18: **fim enquanto**
 - 19: Armazenar os valores dos parâmetros estimados de forma batelada e recursiva para a realização k .
 - 20: $k \leftarrow k + 1$
 - 21: **fim enquanto**
-

Os valores médios e o desvio-padrão dos parâmetros estimados para ambas metodologias apresentadas são ilustrados nas Tabelas 2–3, após a analisá-los percebe-se uma eficiência da variável instrumental nebulosa na forma batelada (BFIV) e recursiva (RFIV) em relação ao método dos mínimos quadrados na forma batelada (BLS) e recursiva (RLS) executados sob as mesmas condições.

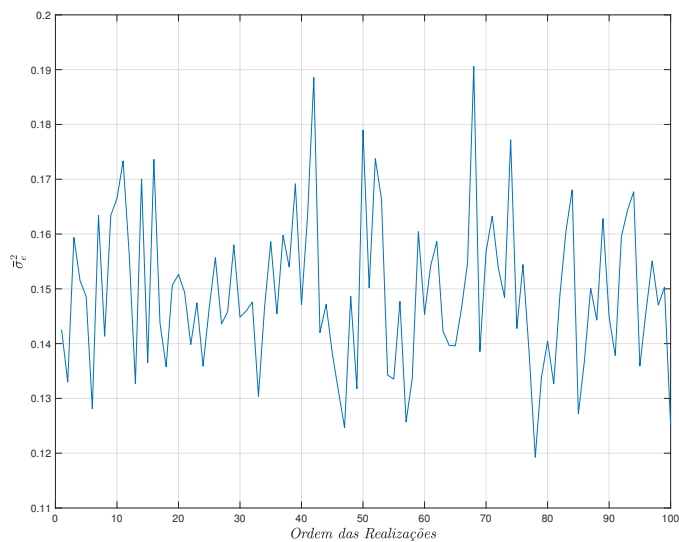
Os resultados dos parâmetros estimados de forma em batelada ao final de cada realização k são ilustrados nas Figuras 46–49, comparados com os respectivos valores nominais dos parâmetros e observa-se que ao final de cada realização a eficiência do método de variável instrumental nebulosa (BFIV) em relação ao método dos mínimos quadrados em batelada (BLS).

Para uma única realização do experimento computacional é ilustrado nas Figuras 50–53 o resultado dos algoritmo de mínimos quadrados (RLS) e variável instrumental nebulosa (RFIV) na forma recursiva sob as mesma condições de inicialização e pesos nebulosos. Percebe-se que o método recursivo com variável instrumental nebulosa na presença do ruído manteve mais estável do que o método dos mínimos quadrados diante do ruído autocorrelacionado.

Os resultados dos parâmetros estimados de forma recursiva ao final de cada realização k são ilustrados nas Figuras 54–57, comparados com os respectivos valores nominais dos parâmetros e observa-se que ao final de cada realização a eficiência do método de variável instrumental nebulosa (RFIV) em relação ao método dos mínimos quadrados recursivo (RLS).

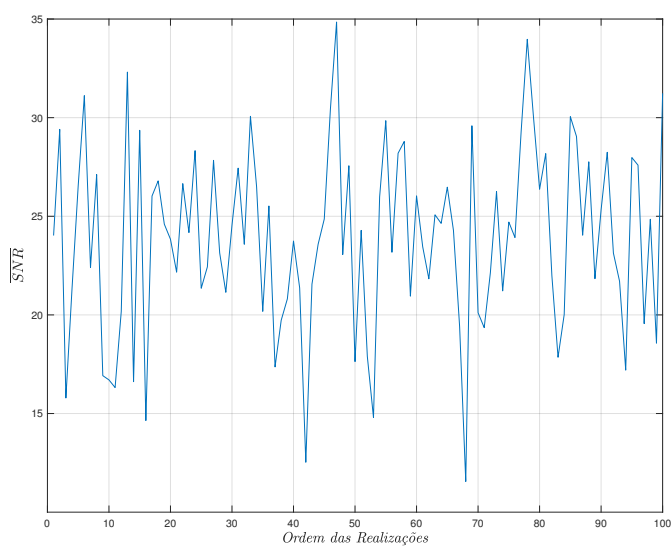
Os histogramas dos parâmetros estimados no método dos mínimos quadrados em batelada (BLS) e recursiva (RLS) e os histogramas dos parâmetros estimados no método de variável instrumental nebulosa em batelada (BFIV) e recursiva (RFIV) são ilustrados nas Figuras 58–65 de modo a visualizar as dispersões dos resultados computacionais em relação aos parâmetros nominais para cada uma das realizações.

Figura 42: Média da variância $\bar{\sigma}_e^2$ utilizada na equação (5.12h) para cada realização de simulação computacional.



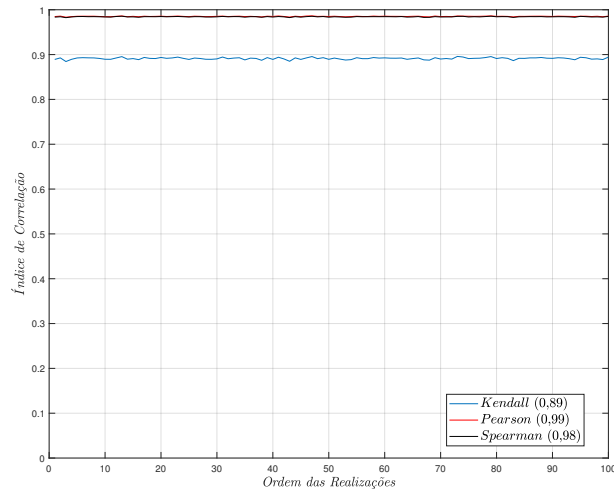
Fonte: Elaborado pelo autor.

Figura 43: Média da relação de sinal/ruído \overline{SNR} observada para cada realização de simulação computacional.



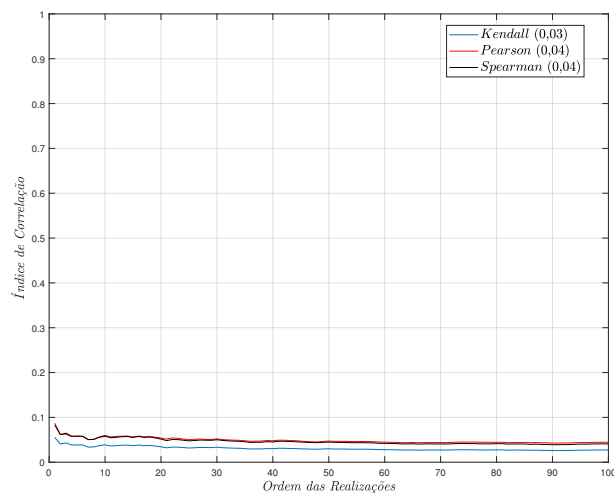
Fonte: Elaborado pelo autor.

Figura 44: Índice de correlação da variável instrumental nebulosa com a variável do vetor de regressores. Avaliado pelos índices de correlação médios de *Kendall* (0,89), *Pearson* (0,99) e *Spearman* (0,98); com os respectivos valores de índices médios obtidos a partir das 100 realizações.



Fonte: Elaborado pelo autor.

Figura 45: Índice de correlação da variável instrumental nebulosa com o erro inserido na variável de saída. Avaliado pelos índices de correlação médios de *Kendall* (0,03), *Pearson* (0,04) e *Spearman* (0,04); com os respectivos valores de índices médios obtidos a partir das 100 realizações.



Fonte: Elaborado pelo autor.

Tabela 2: Análise comparativa da média (**desvio-padrão**) dos parâmetros estimados $\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4$ para os métodos de mínimos quadrados não-recursivo (batelada) (**BLS**) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (**BFIV**) .

Parâmetros	Valores Nominais	BLS	BFIV
\hat{a}_1	1,5	1,3554 (0,0361)	1,5324 (0,0215)
\hat{a}_2	-0,7	-0,5737 (0,0330)	-0,7271 (0,0190)
\hat{a}_3	-0,3	-0,3101 (0,0066)	-0,2997 (0,0036)
\hat{a}_4	0,2	0,1395 (0,0152)	0,2164 (0,0093)
\hat{b}_1	1,0	1,0000 (0,0045)	1,0000 (0,0043)
\hat{b}_2	0,5	0,6445 (0,0382)	0,4676 (0,0229)
\hat{b}_3	-2,0	-1,8373 (0,0409)	-2,0376 (0,0256)
\hat{b}_4	1,0	0,8008 (0,0510)	1,0441 (0,0296)

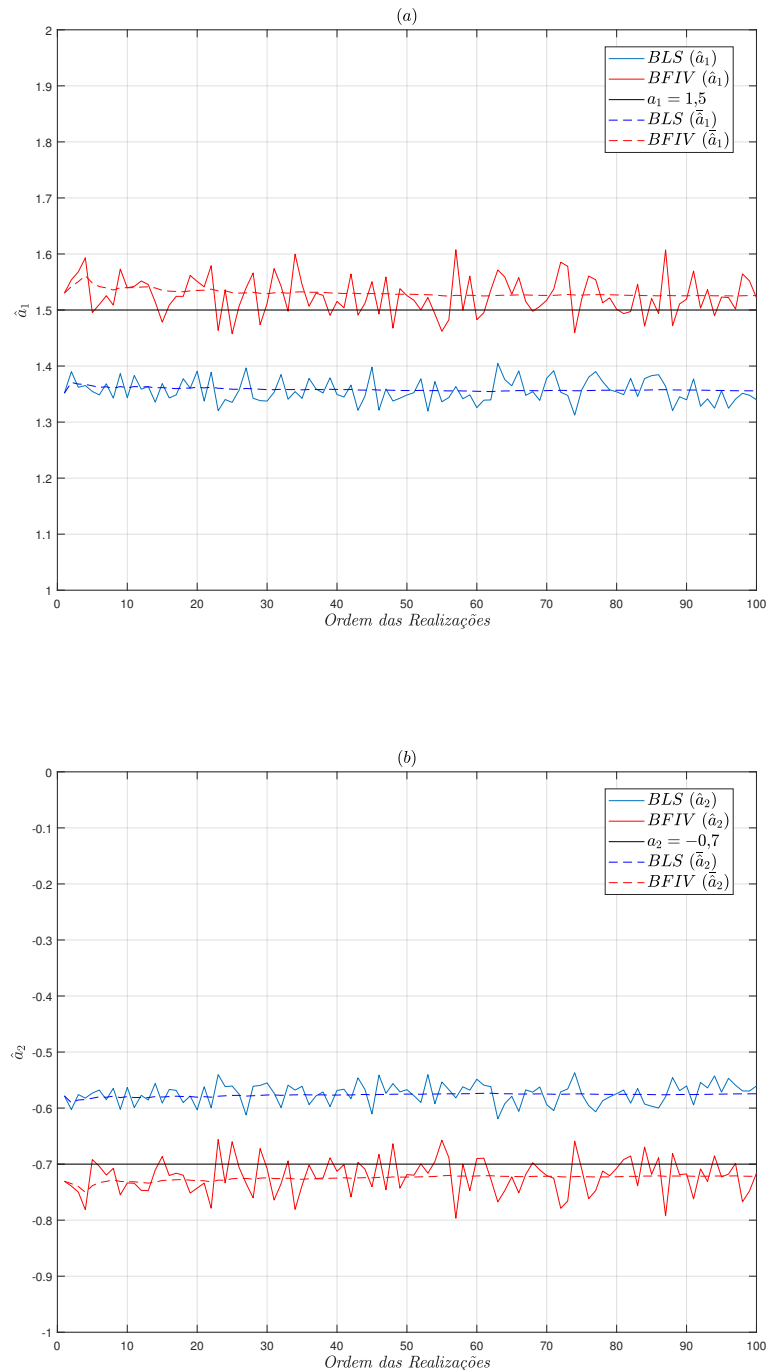
Fonte: Elaborado pelo autor

Tabela 3: Análise comparativa da média (**desvio-padrão**) dos parâmetros estimados $\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4$ para os métodos dos mínimos quadrados recursivo (**RLS**) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (**RFIV**).

Parâmetros	Valores Nominais	RLS	RFIV
\hat{a}_1	1,5	1,3544 (0,0329)	1,5288 (0,0209)
\hat{a}_2	-0,7	-0,5728 (0,0299)	-0,7239 (0,0184)
\hat{a}_3	-0,3	-0,3102 (0,0064)	-0,3001 (0,0036)
\hat{a}_4	0,2	0,1391 (0,0137)	0,2150 (0,0091)
\hat{b}_1	1,0	1,0000 (0,0045)	1,0000 (0,0043)
\hat{b}_2	0,5	0,6455 (0,0344)	0,4714 (0,0223)
\hat{b}_3	-2,0	-1,8362 (0,0371)	-2,0337 (0,0249)
\hat{b}_4	1,0	0,7994 (0,0462)	1,0391 (0,0288)

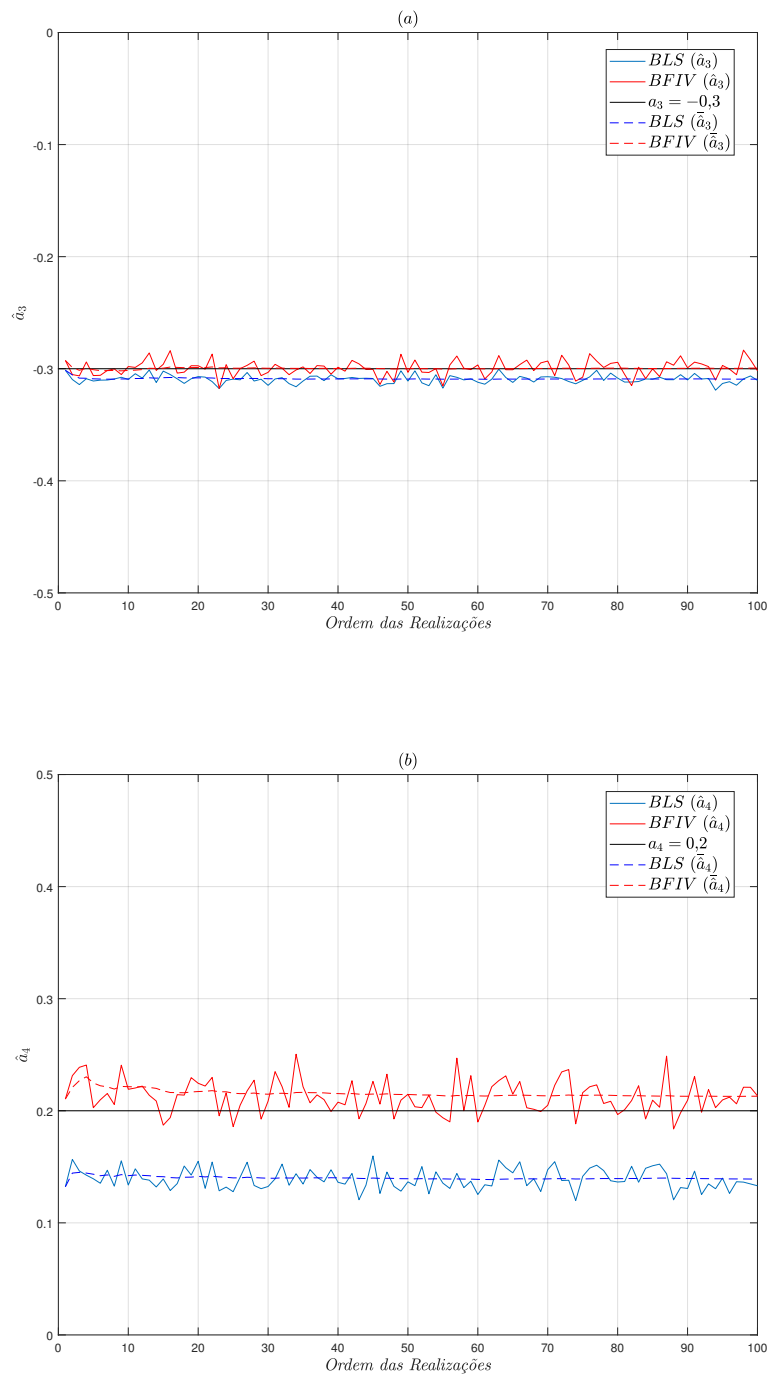
Fonte: Elaborado pelo autor

Figura 46: Resultado da estimação dos parâmetros \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos dos mínimos quadrados não-recursivo (batelada) (BLS) e do método dos mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV) em função do número de realizações.



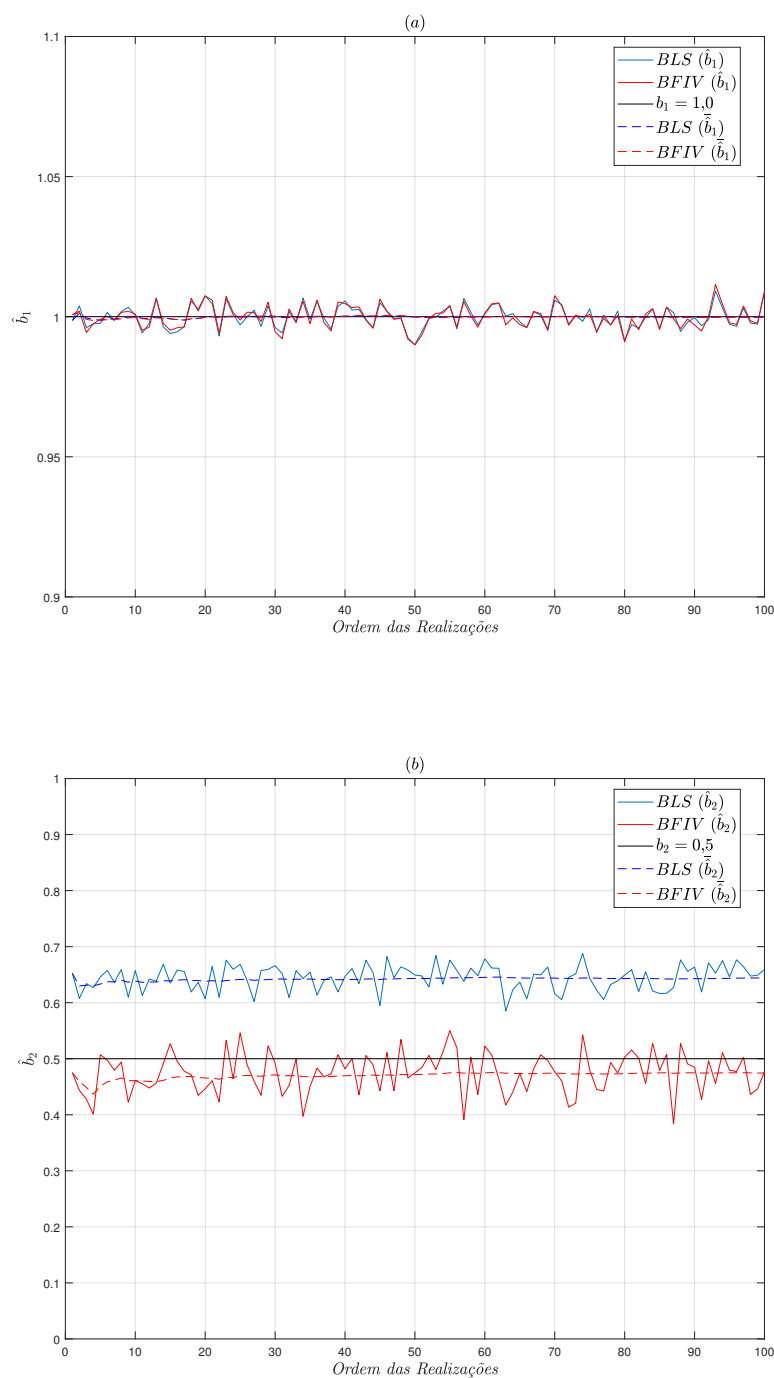
Fonte: Elaborado pelo autor.

Figura 47: Resultado da estimação dos parâmetros \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos dos mínimos quadrados não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV) em função do número de realizações.



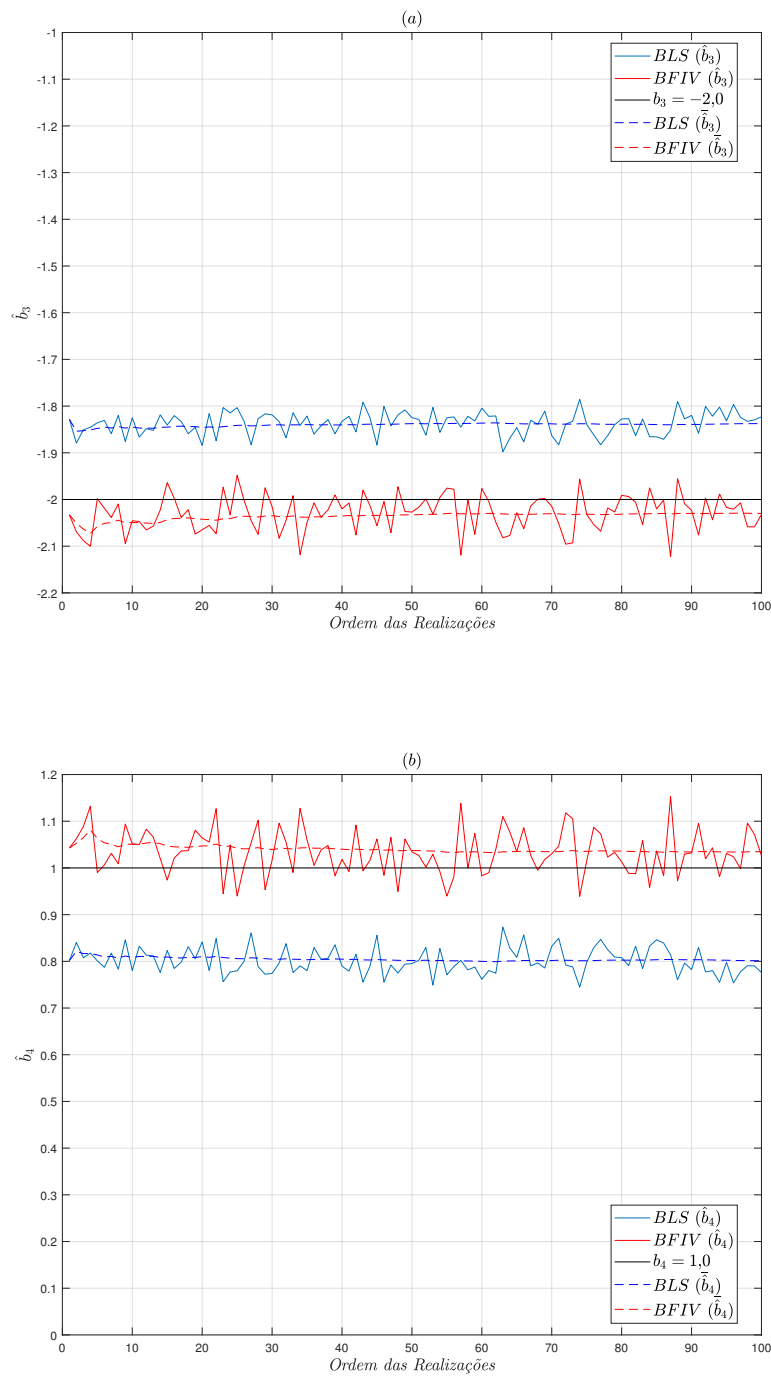
Fonte: Elaborado pelo autor.

Figura 48: Resultado da estimação dos parâmetros \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos dos mínimos quadrados não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV) em função do número de realizações.



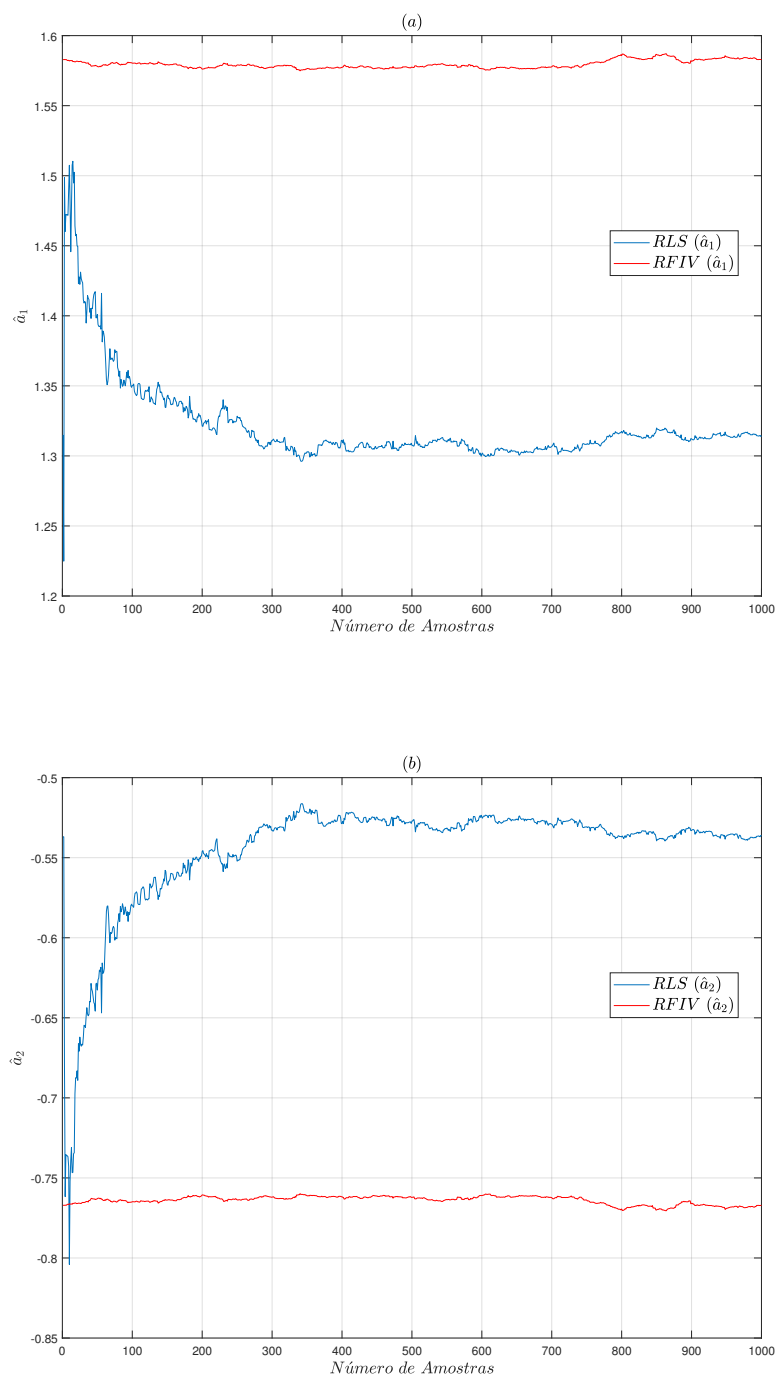
Fonte: Elaborado pelo autor.

Figura 49: Resultado da estimação dos parâmetros \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos dos mínimos quadrados não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV) em função do número de realizações.



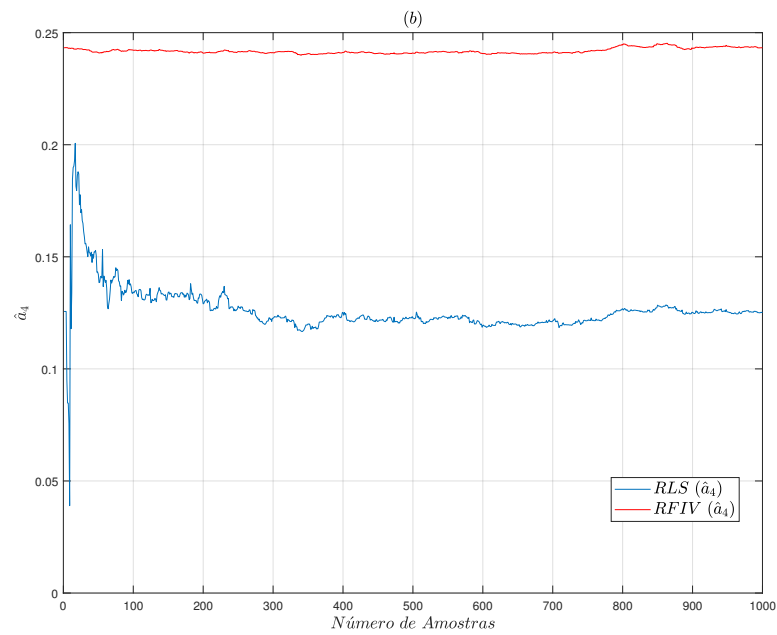
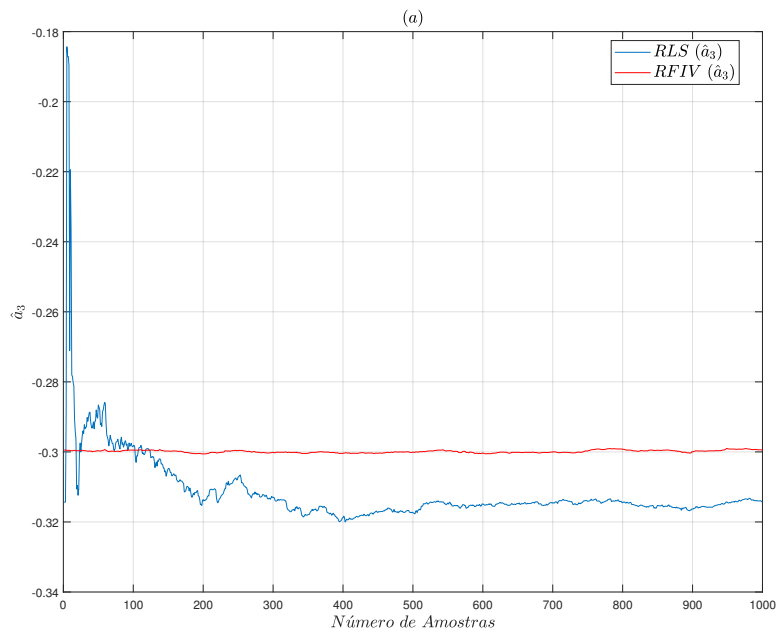
Fonte: Elaborado pelo autor.

Figura 50: Resultado da estimação dos parâmetros \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.



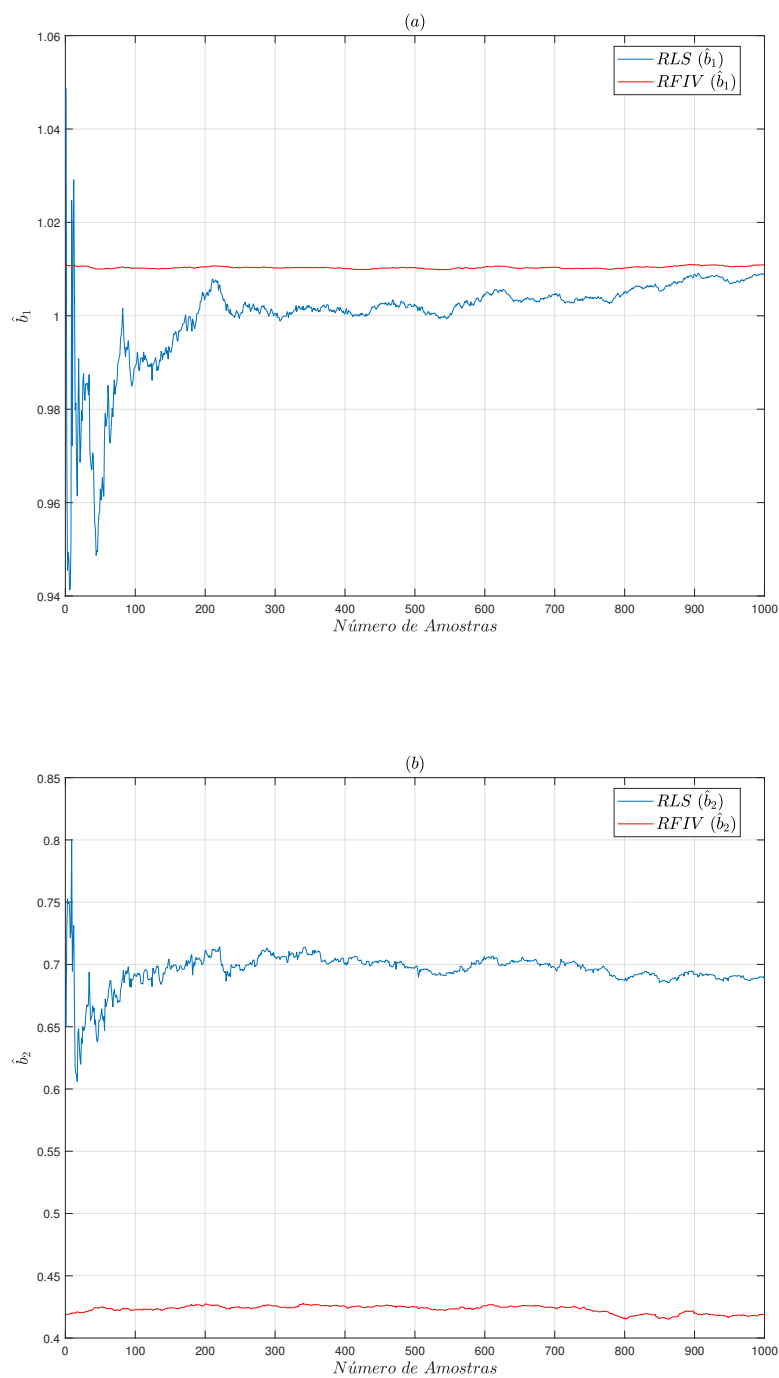
Fonte: Elaborado pelo autor.

Figura 51: Resultado da estimação dos parâmetros \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.



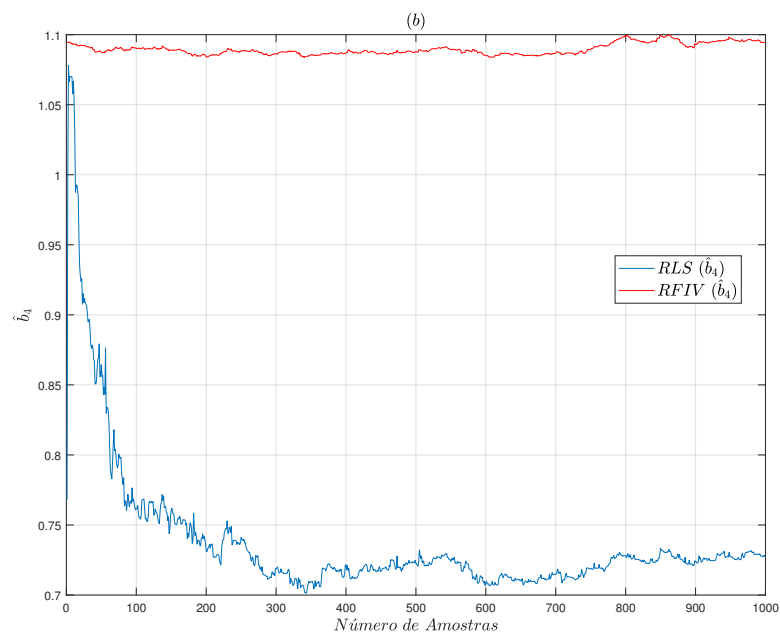
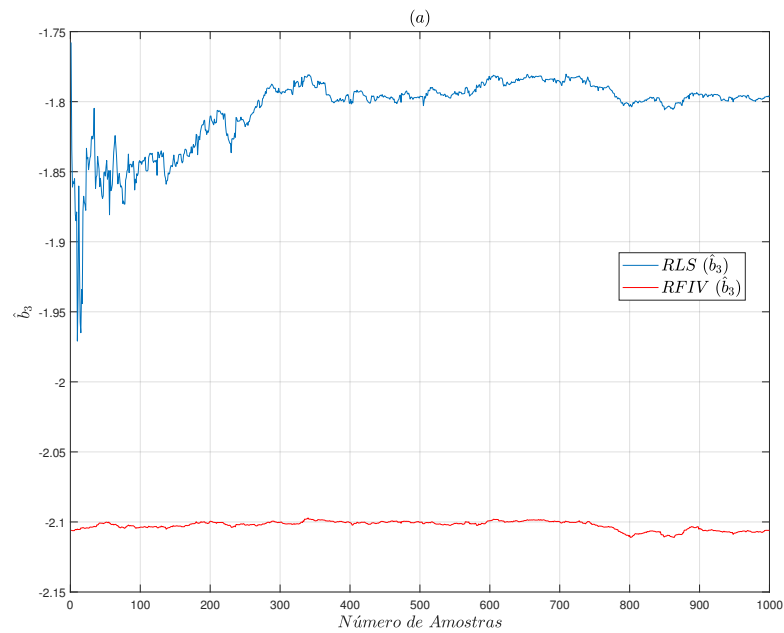
Fonte: Elaborado pelo autor.

Figura 52: Resultado da estimação dos parâmetros \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.



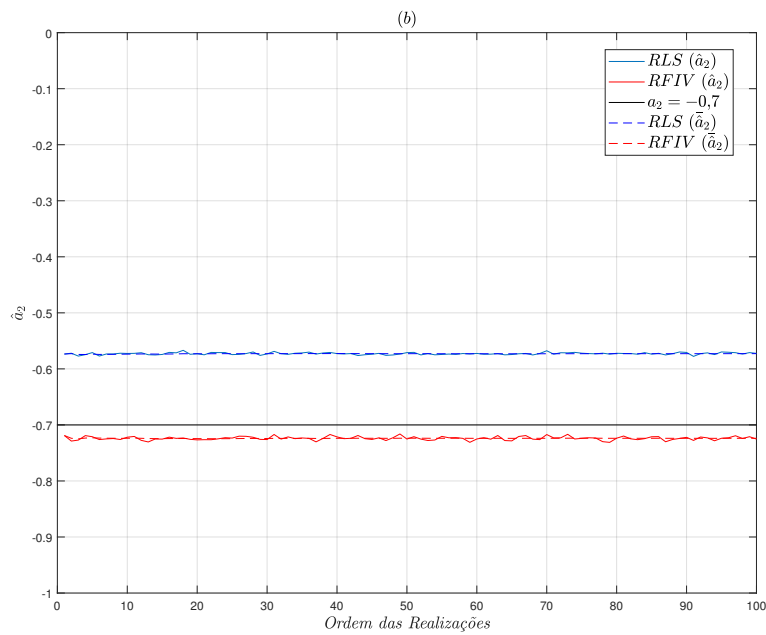
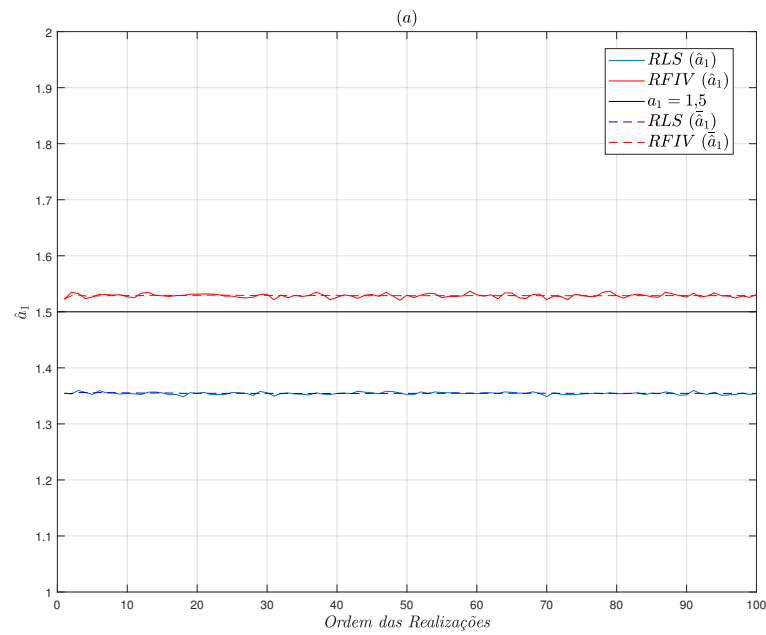
Fonte: Elaborado pelo autor.

Figura 53: Resultado da estimação dos parâmetros \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número do número de amostras para a única realização.



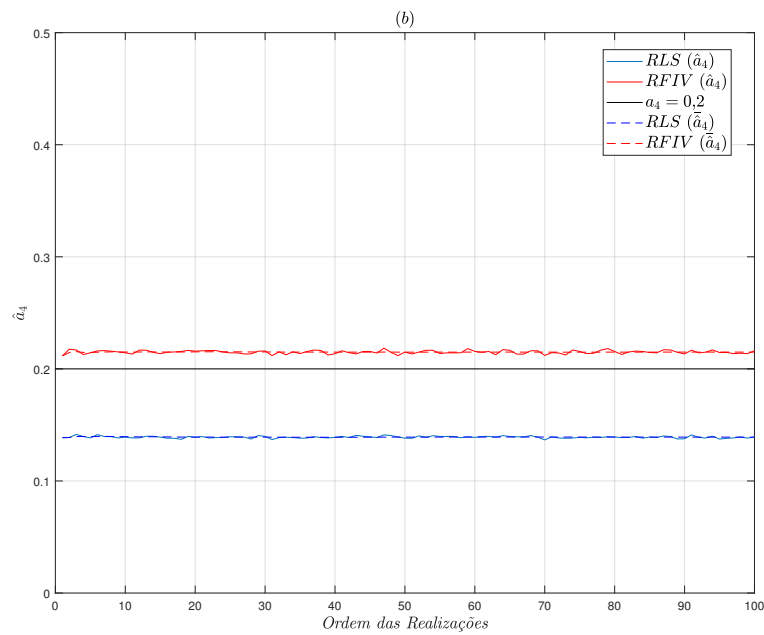
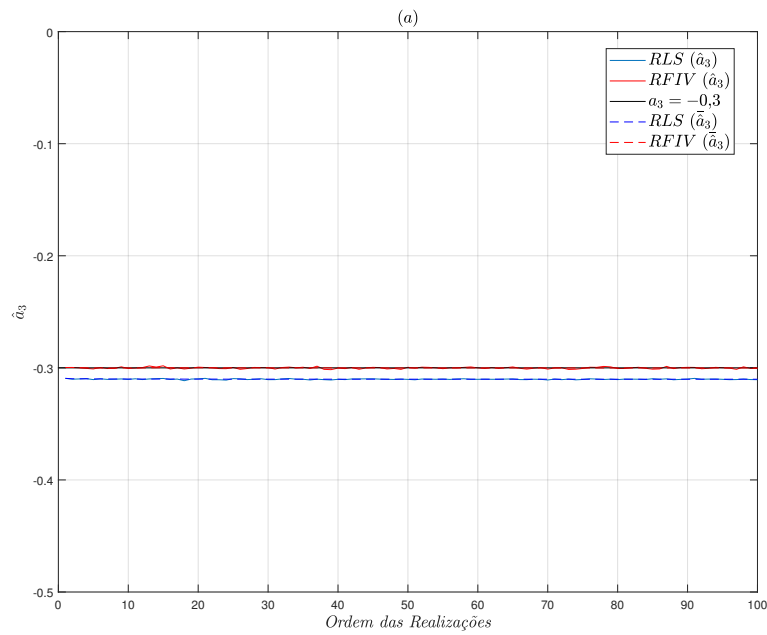
Fonte: Elaborado pelo autor.

Figura 54: Resultado da estimação dos parâmetros \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.



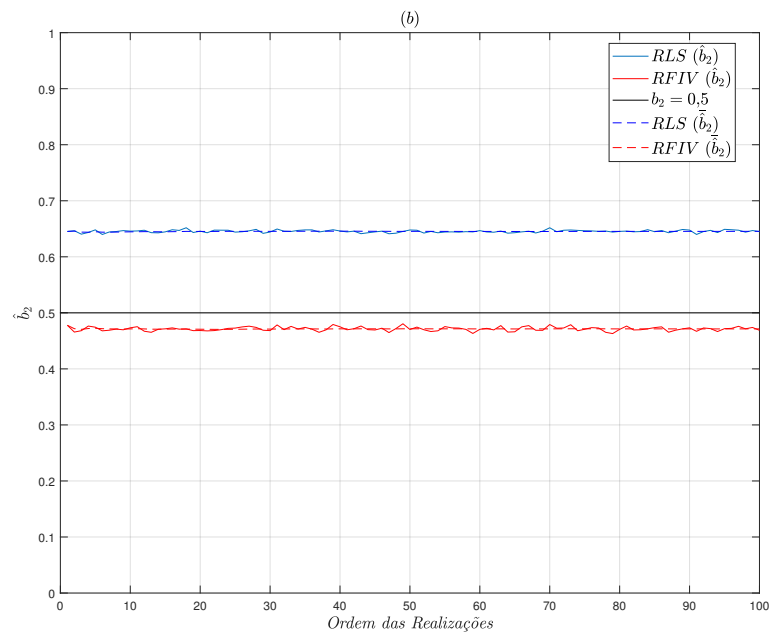
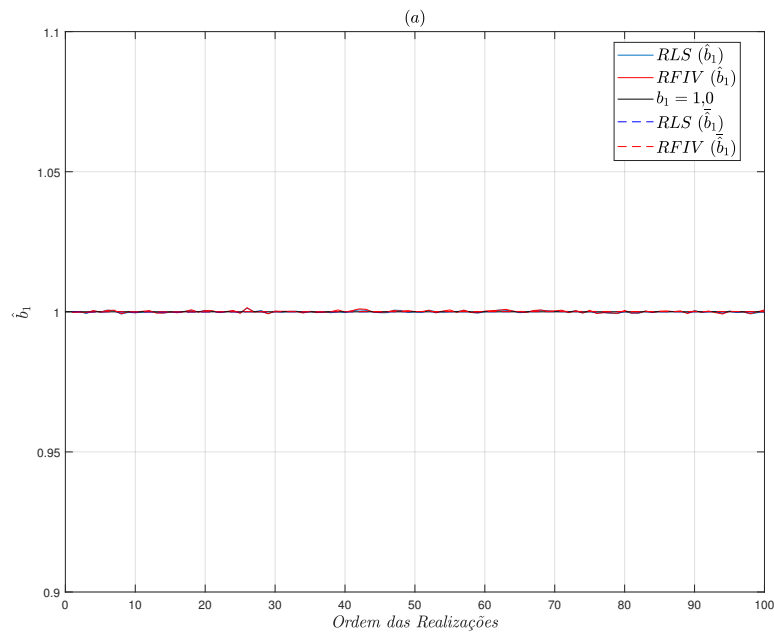
Fonte: Elaborado pelo autor.

Figura 55: Resultado da estimação dos parâmetros \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.



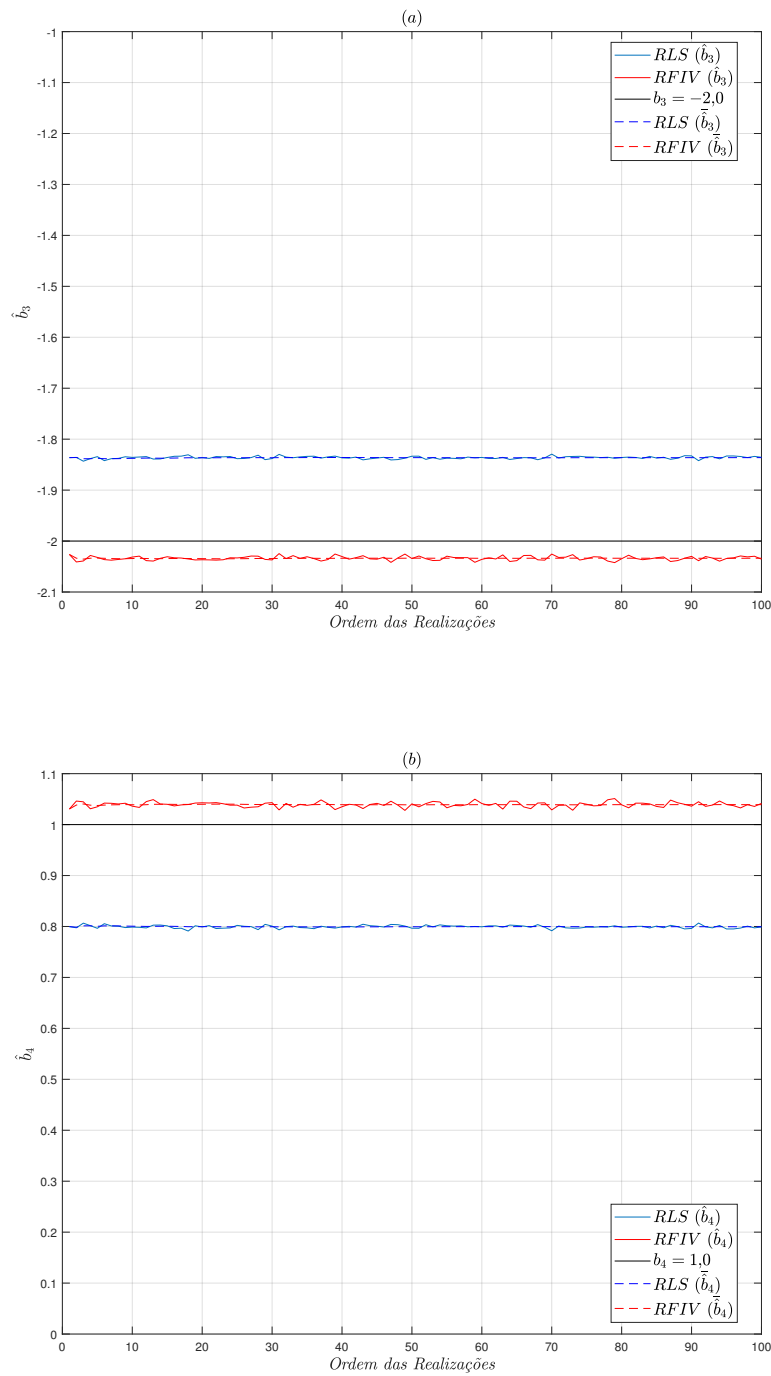
Fonte: Elaborado pelo autor.

Figura 56: Resultado da estimação dos parâmetros \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.



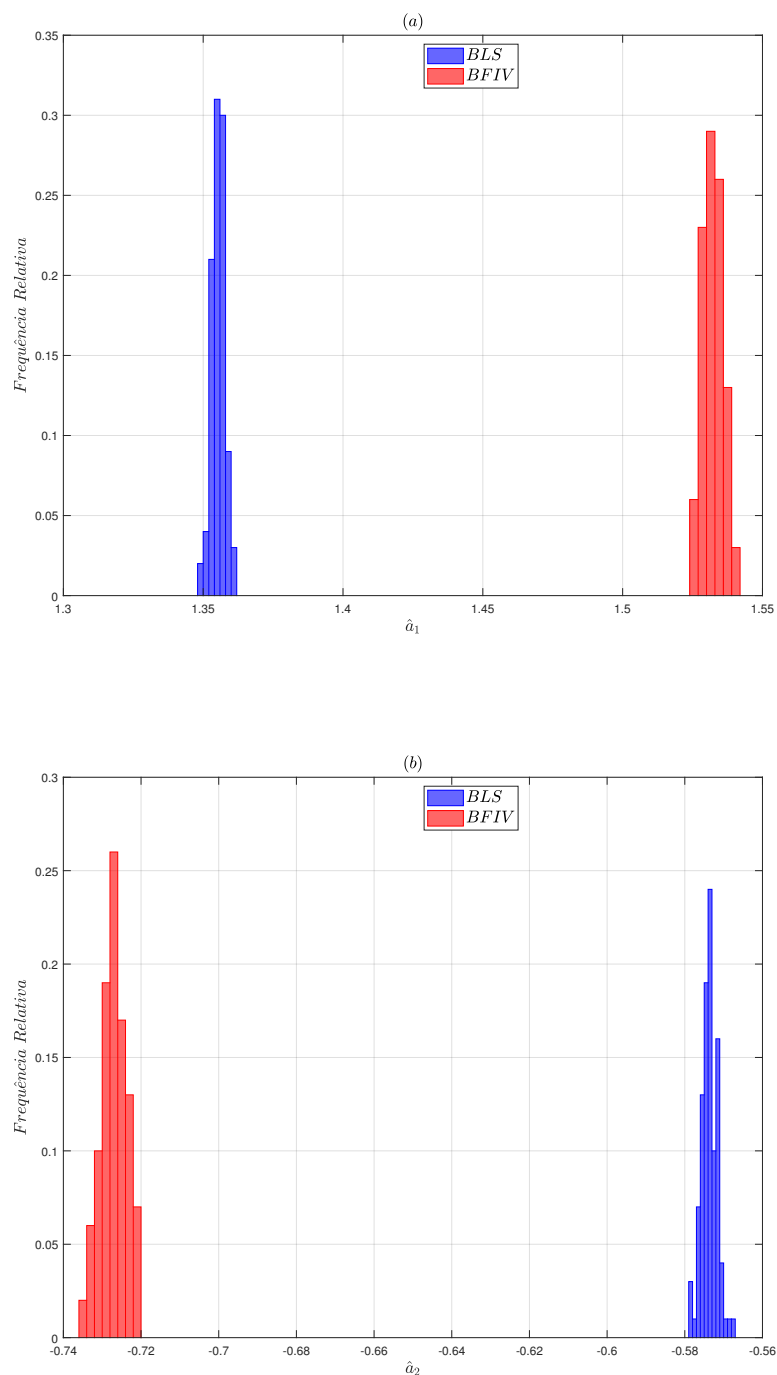
Fonte: Elaborado pelo autor.

Figura 57: Resultado da estimação dos parâmetros \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos dos mínimos quadrados recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV) em função do número de realizações.



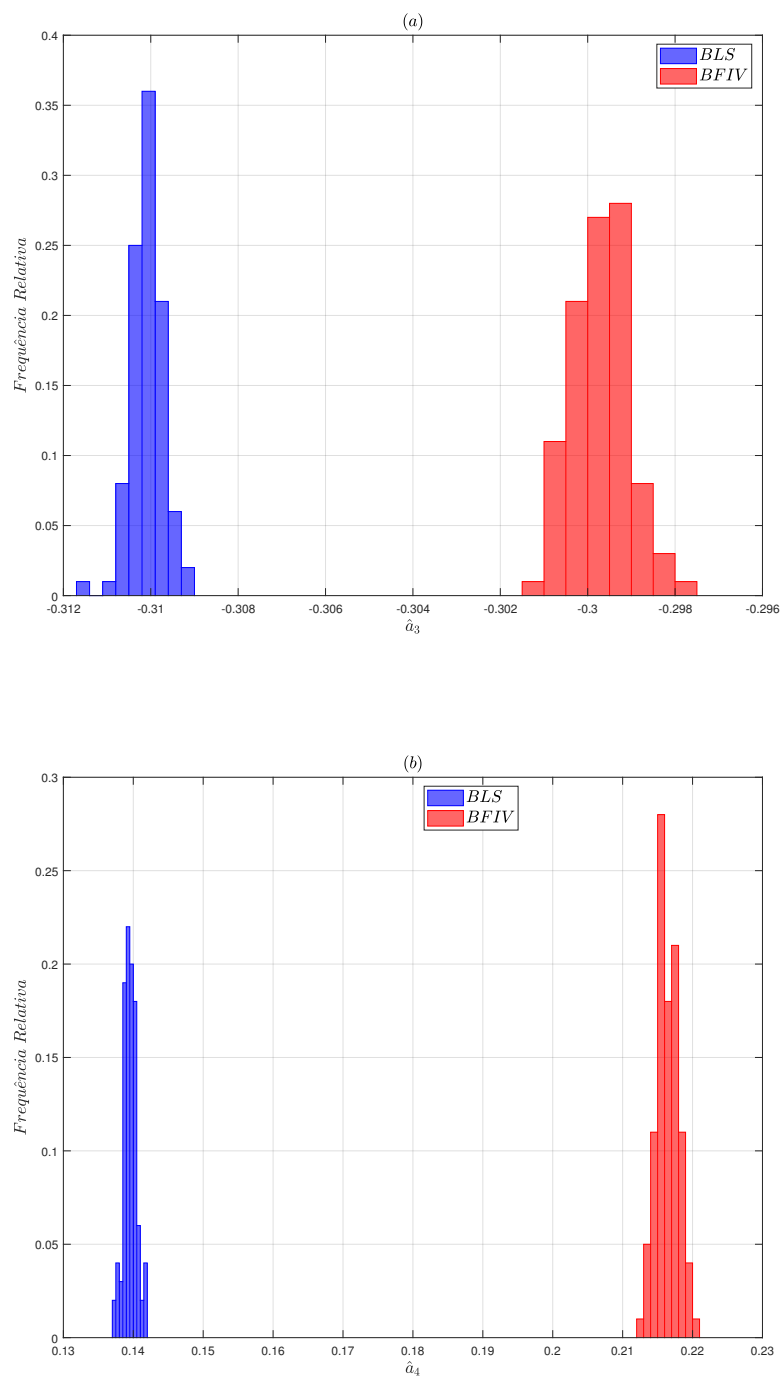
Fonte: Elaborado pelo autor.

Figura 58: Histograma de parâmetros estimados \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).



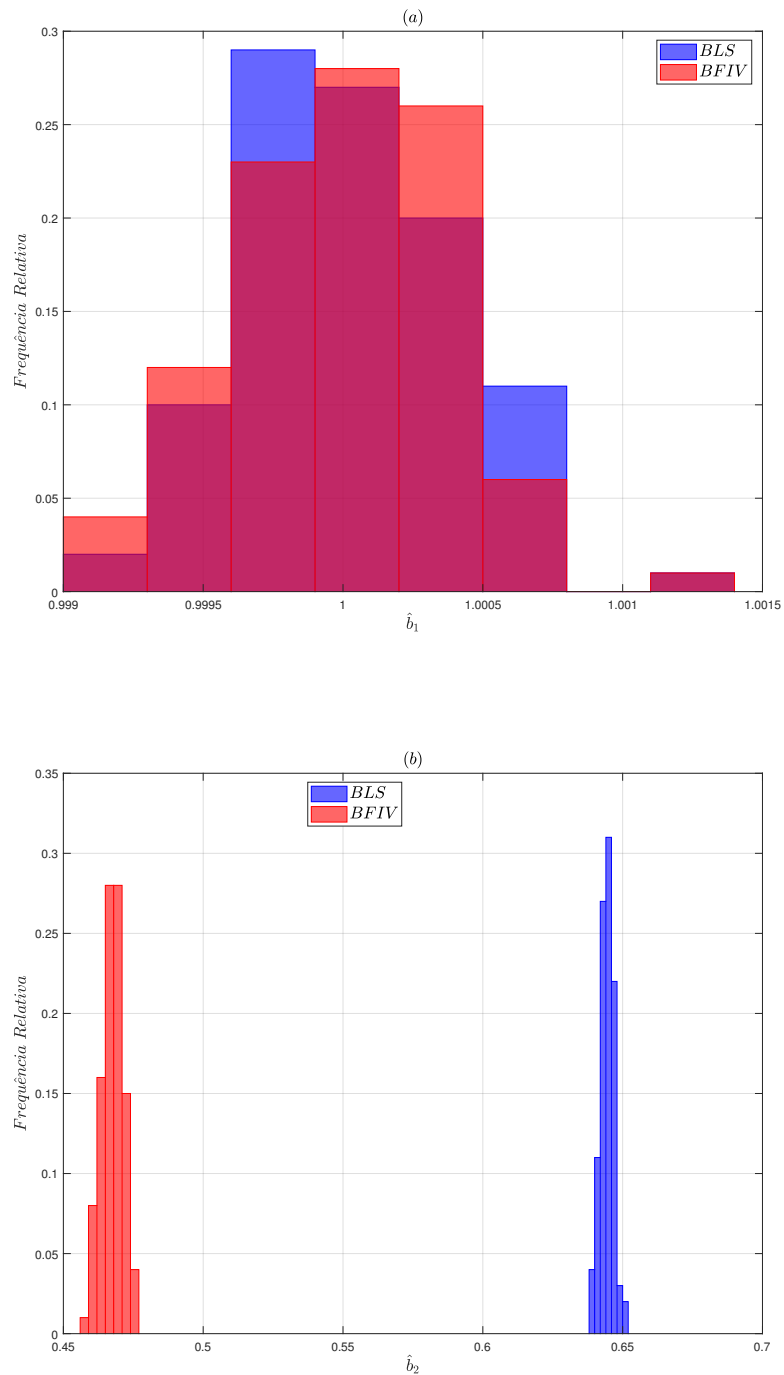
Fonte: Elaborado pelo autor.

Figura 59: Histograma de parâmetros estimados $\hat{\alpha}_3$ (a) e $\hat{\alpha}_4$ (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).



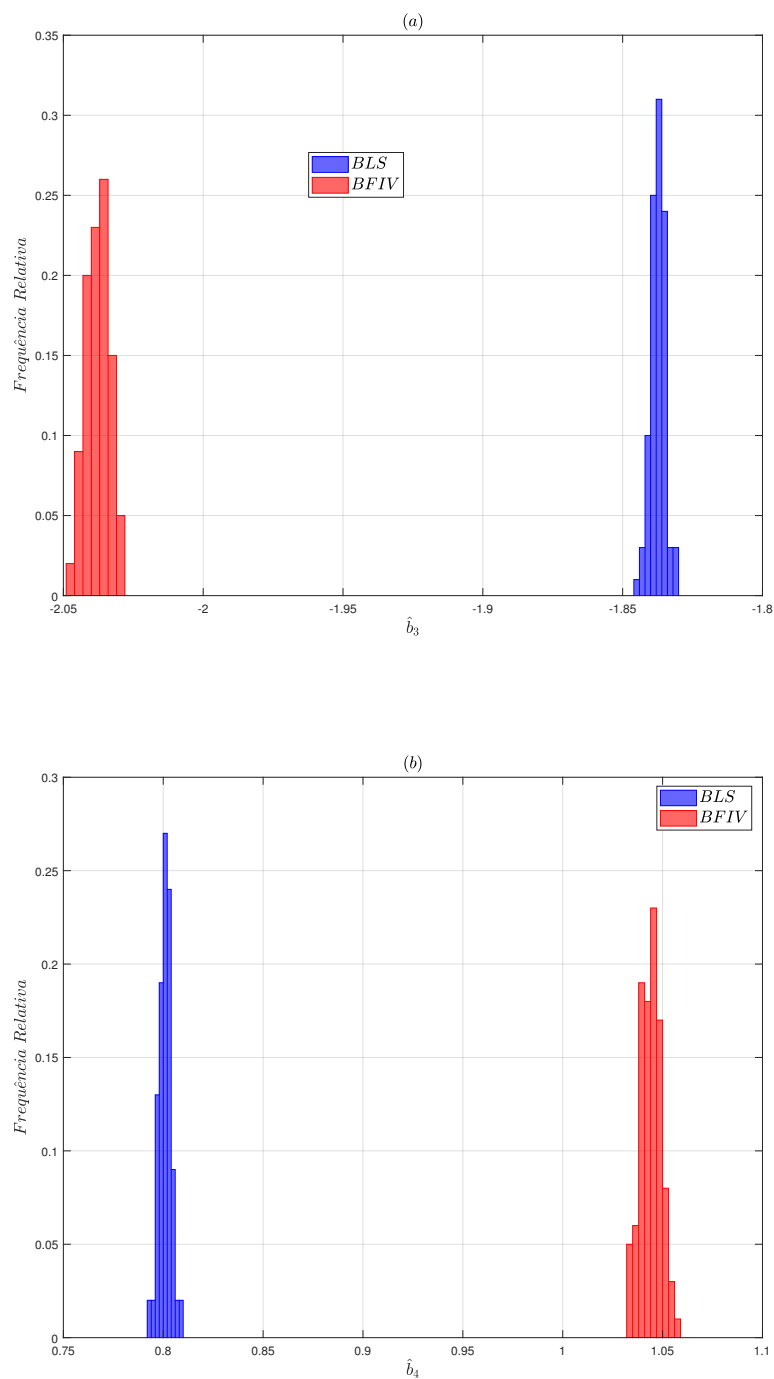
Fonte: Elaborado pelo autor.

Figura 60: Histograma de parâmetros estimados \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).



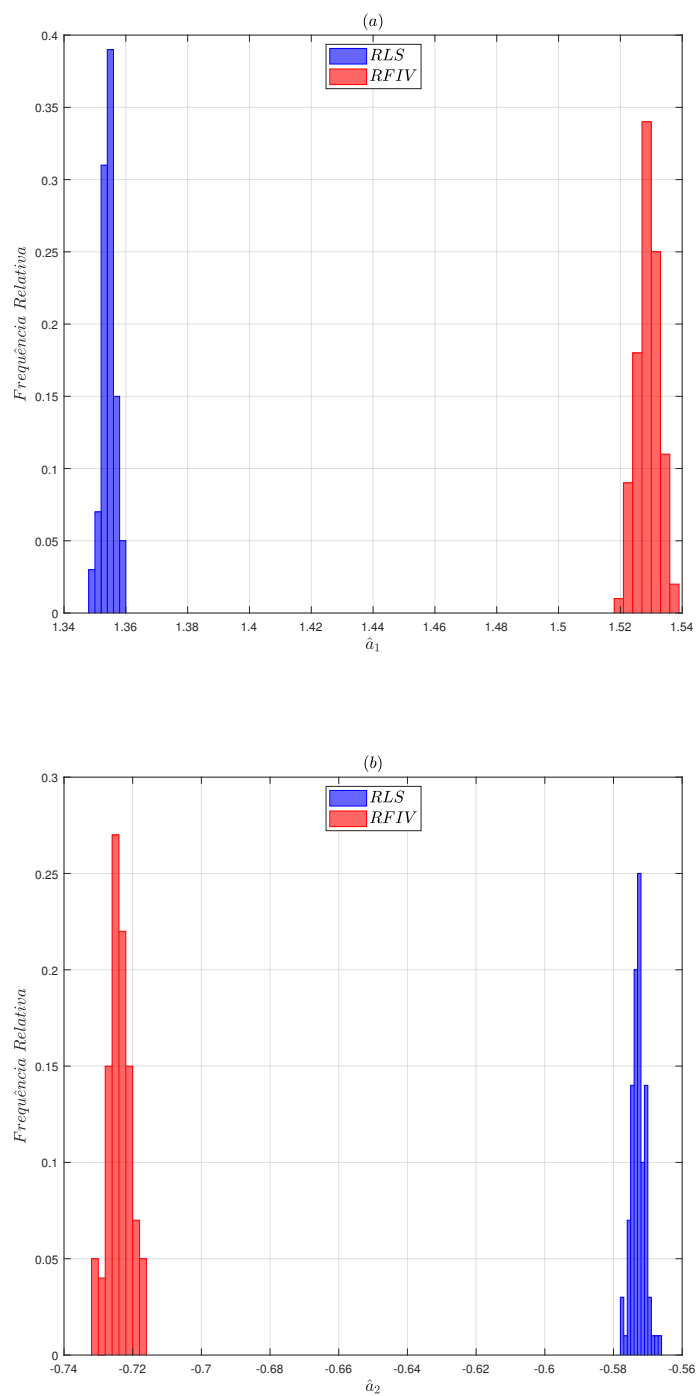
Fonte: Elaborado pelo autor.

Figura 61: Histograma de parâmetros estimados \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos não-recursivo (batelada) (BLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma não-recursiva (Batelada) (BFIV).



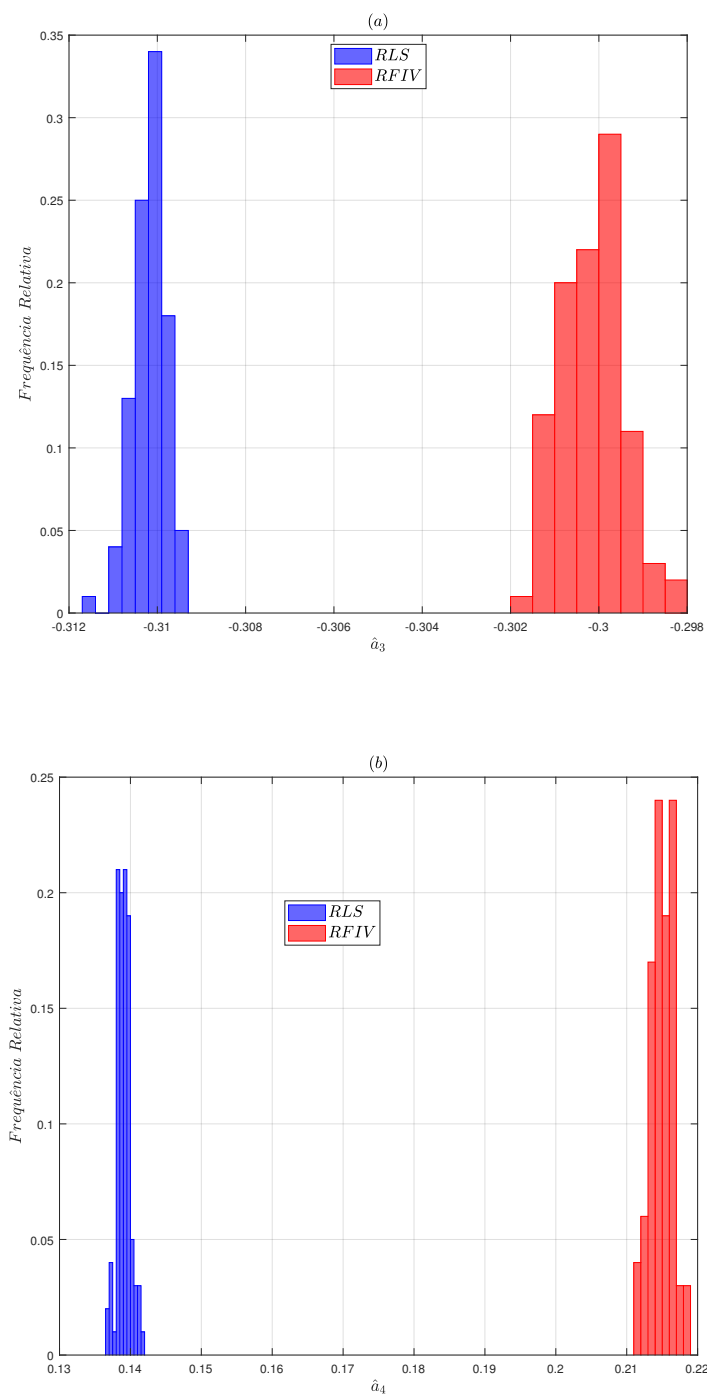
Fonte: Elaborado pelo autor.

Figura 62: Histograma de parâmetros estimados \hat{a}_1 (a) e \hat{a}_2 (b) a partir dos métodos recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).



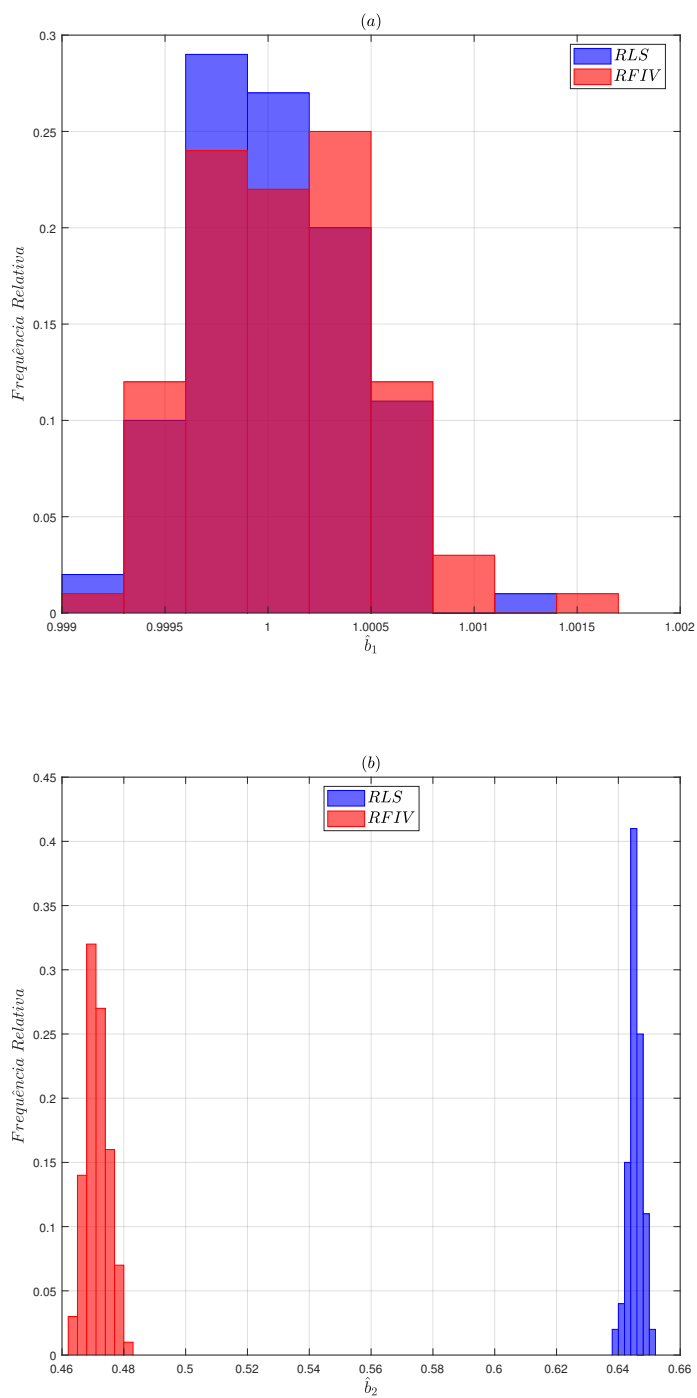
Fonte: Elaborado pelo autor.

Figura 63: Histograma de parâmetros estimados \hat{a}_3 (a) e \hat{a}_4 (b) a partir dos métodos recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).



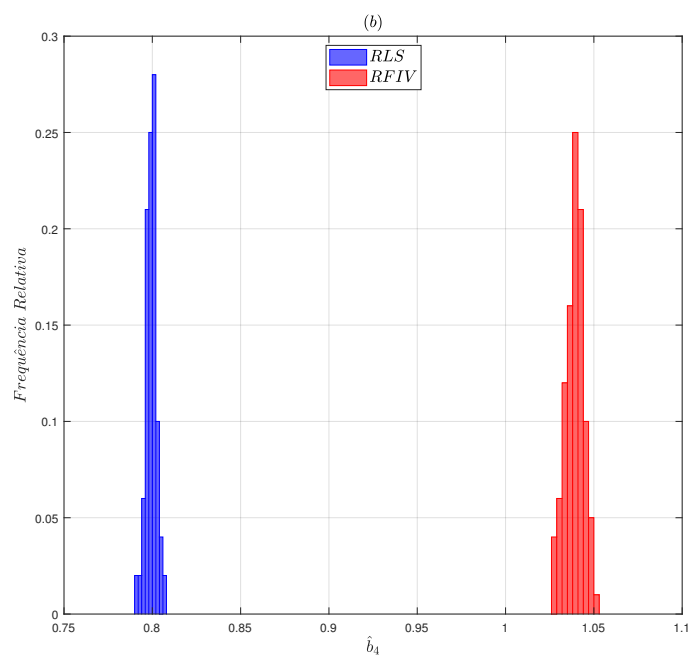
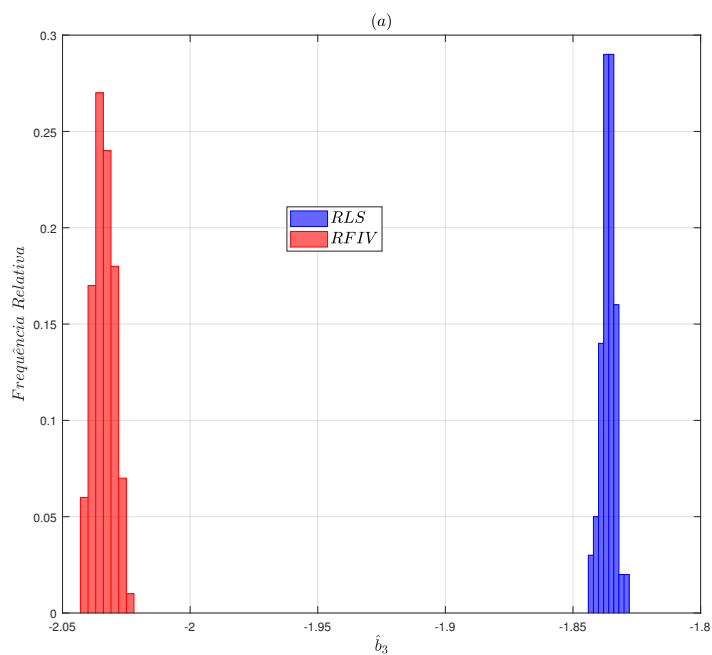
Fonte: Elaborado pelo autor.

Figura 64: Histograma de parâmetros estimados \hat{b}_1 (a) e \hat{b}_2 (b) a partir dos métodos recursivo (RLS) e do método dos mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).



Fonte: Elaborado pelo autor.

Figura 65: Histograma de parâmetros estimados \hat{b}_3 (a) e \hat{b}_4 (b) a partir dos métodos recursivo (RLS) e do método do mínimos quadrados com variável instrumental nebulosa de forma recursiva (RFIV).

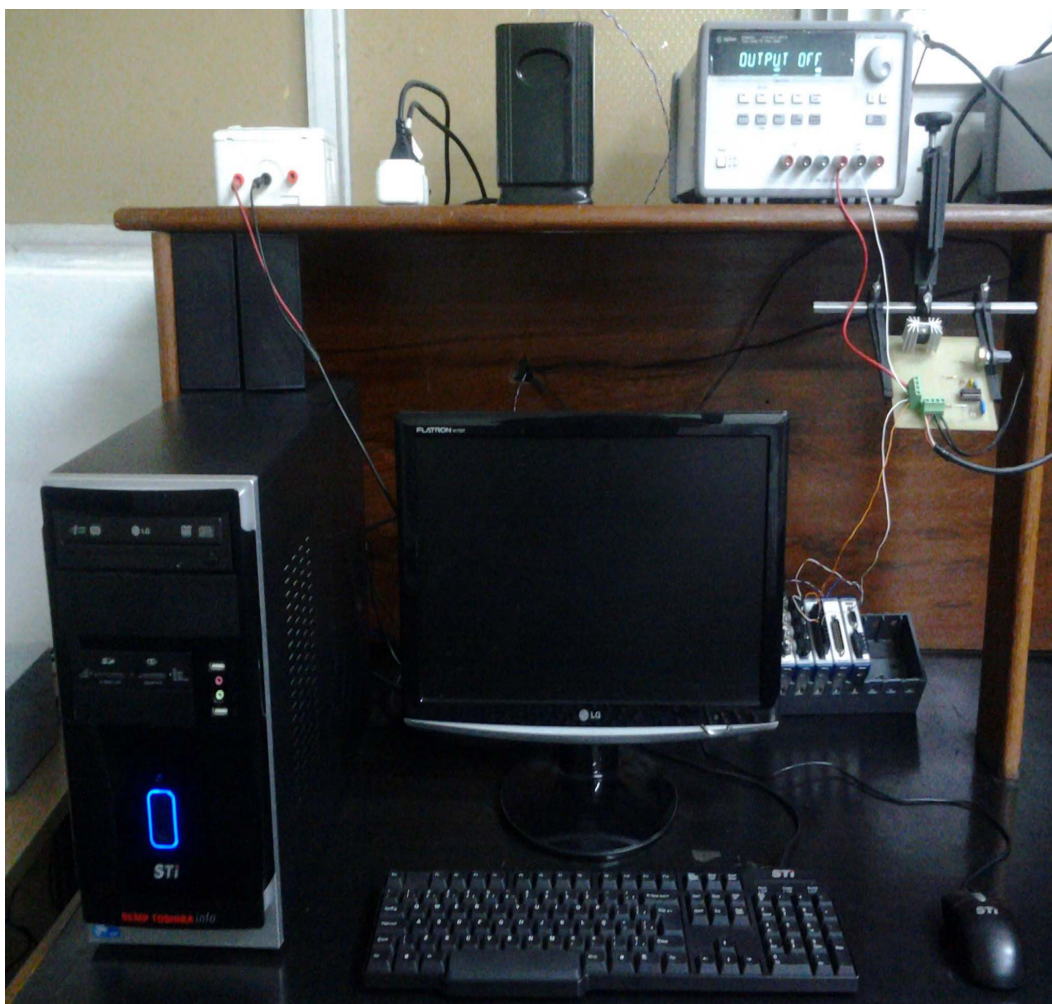


Fonte: Elaborado pelo autor.

5.4 Identificação de um Processo Térmico

Nesta seção, a metodologia proposta de algoritmo evolutivo são fundamentados na modelagem caixa preta de uma planta térmica (processo térmico). Os resultados experimentais da modelagem exemplificam o uso da metodologia de agrupamento nebuloso evolutivo com base no critério de máxima verossimilhança que utilizam o aprendizado participativo e algoritmo de junção de regras nebulosas evolutivas. Os dados experimentais foram obtidos a partir de uma plataforma de controle em tempo real que é composta de um processo térmico baseado em uma torradeira elétrica (PIRES; SERRA, 2014), conforme ilustrado na Figura .

Figura 66: Processo térmico baseado em uma torradeira elétrica. Processo em tempo real com sinal de entrada (tensão de entrada) e um sinal de saída (temperatura do processo térmico).



Fonte: (PIRES; SERRA, 2014).

Na Figura 67 são ilustrados os sinais de entrada $u^{(k)}$ e saída do processo térmico $y^{(k)}$, para $k = 1, 2, \dots, 1000$. Observa-se que há quatro regiões distintas deste processo para a variável de entrada como referência, que também é destacado no mapeamento de espaço \mathbb{R}^2 para $u^{(k)} \times y^{(k)}$, conforme ilustrado na Figura 68. O que pode caracterizar como quatro regiões em que suas respectivas amostras pertencentes mantêm uma similaridade entre elas, portanto, um algoritmo de agrupamento deve ser capaz de identificar tais regiões como grupos distintos, conforme observado também na Figura 68.

O modelo de inferência nebuloso Takagi–Sugeno utilizado na identificação e validação, é descrito a seguir:

$$R_{(i)} : \text{SE } u^{(k)} \text{ É } \Gamma_{u^{(k)}}^{(i)} \text{ ENTÃO}$$

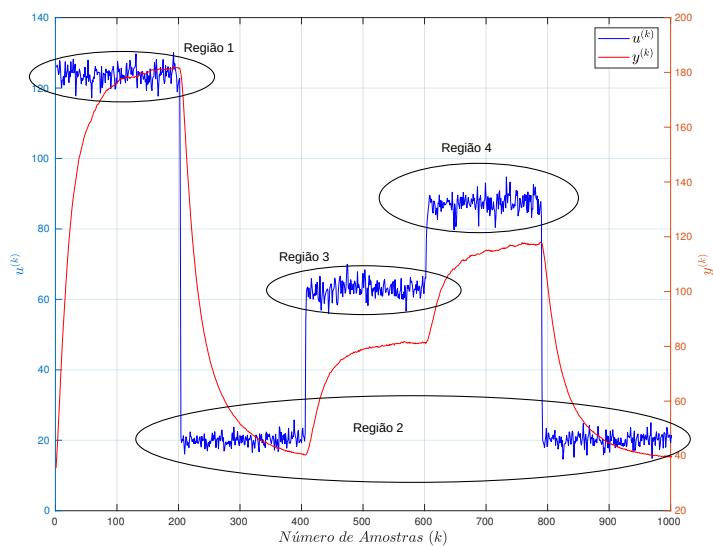
$$\hat{y}_{(i)}^{(k)} = \theta_{u^{(k-2)}}^{(i)} u^{(k-2)} + \theta_{u^{(k-1)}}^{(i)} u^{(k-1)} + \theta_{y^{(k-2)}}^{(i)} \hat{y}^{(k-2)} + \theta_{y^{(k-1)}}^{(i)} \hat{y}^{(k-1)}. \quad (5.14)$$

A identificação evolutiva da planta térmica é apresentada na Figura 69: os primeiros 250 pontos são aplicados ao algoritmo em batelada (estimação inicial), e os 750 pontos finais são aplicados ao algoritmo evolutivo (estimação evolutiva).

Após a última iteração o algoritmo nebuloso evolutivo de máxima verossimilhança utilizando o fator de aprendizagem com uma taxa fixa de aprendizado (GEORGIEVA; FILEV, 2009; ANGELOV; FILEV; KASABOV, 2010), o mesmo identificou apenas dois grupos similares ou seja, as regiões 2, 3 e 4 foram identificadas como grupo de amostras similares entre si, e as amostras pertencentes a região 1 identificada como o grupo 1, conforme ilustrado na Figura 70. Nesta tese, utiliza-se o fator de aprendizado participativo baseado numa métrica exponencial da máxima verossimilhança, após a última iteração do algoritmo percebe-se que este foi capaz de identificar quatro grupos relativos as quatro regiões distintas, conforme ilustrado na Figura 71. A análise comparativa entre ambas as metodologias propostas após a validação do modelo com base nas métricas de avaliação, seção 5.1, são ilustradas na Tabela 4. Os resultados demonstram que há um desempenho superior do algoritmo de agrupamento nebuloso evolutivo que utiliza o fator de aprendizado participativo em relação ao agrupamento realizado com o modelo proposto em (GEORGIEVA; FILEV, 2009; ANGELOV; FILEV; KASABOV, 2010) que utiliza uma taxa de aprendizado fixa.

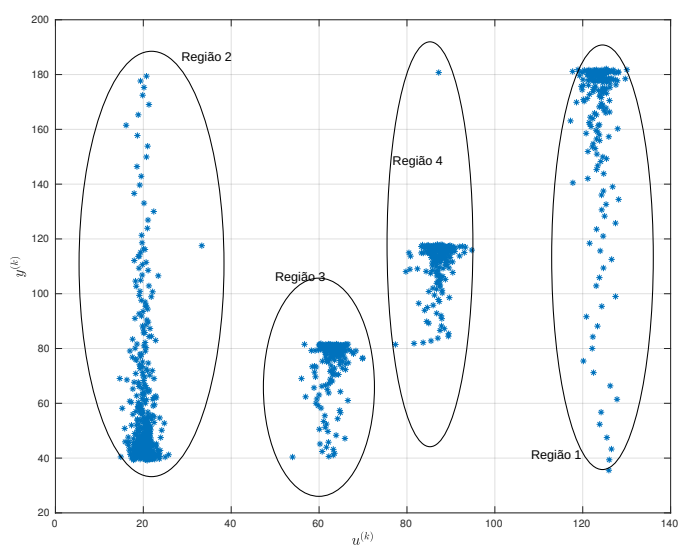
O total de regras avaliados após a leitura da $k_{\text{ésima}}$ amostra é ilustrado na Figura 72. Observe que após a leitura da $750_{\text{ésima}}$ amostra, o algoritmo de junção de regras identificou apenas quatro grupos similares, ocorrendo a junção de grupos próximos conforme proposto nos algoritmos de junção de grupos e algoritmo de aprendizado participativo.

Figura 67: Sinais de entrada $u^{(k)}$ e saída $y^{(k)}$ do processo térmico para 1000 amostras.



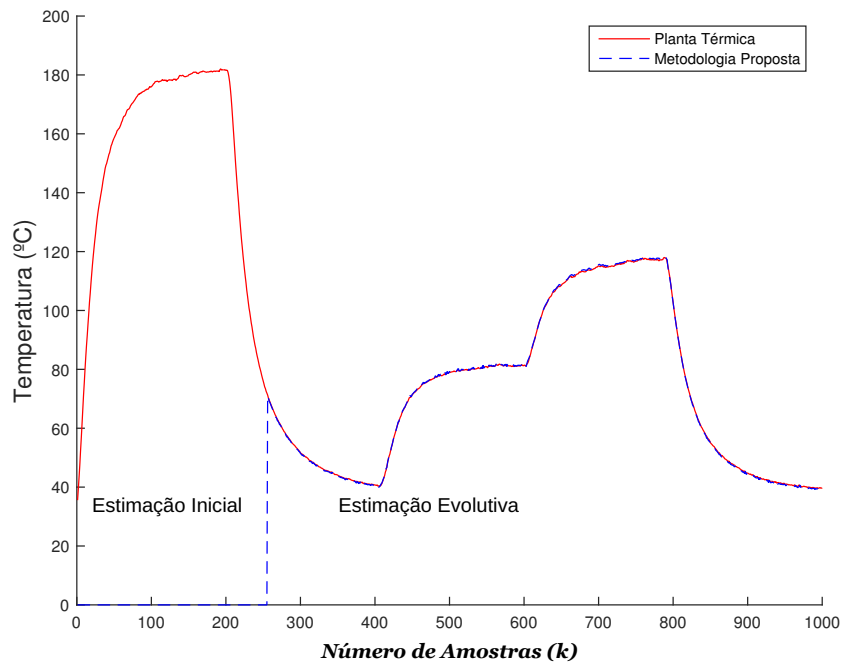
Fonte: Elaborado pelo Autor.

Figura 68: Sinal de entrada $u^{(k)}$ versus sinal saída $y^{(k)}$ do processo térmico no espaço \mathbb{R}^2 .



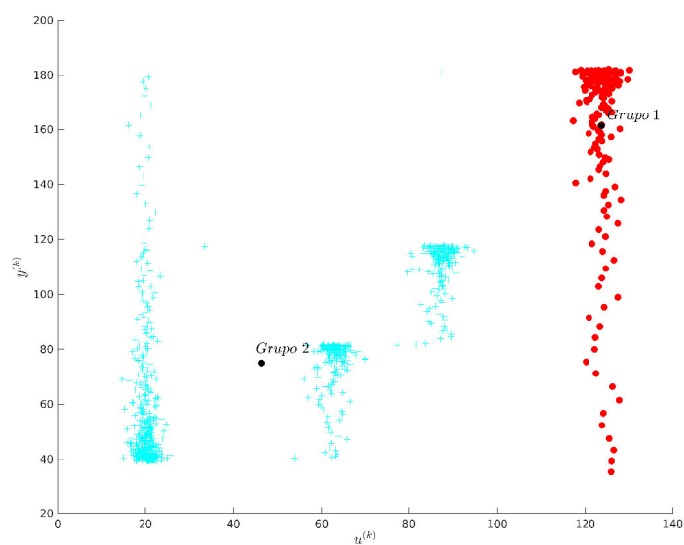
Fonte: Elaborado pelo Autor.

Figura 69: Sinal de saída utilizado para realizar a identificação do modelo.



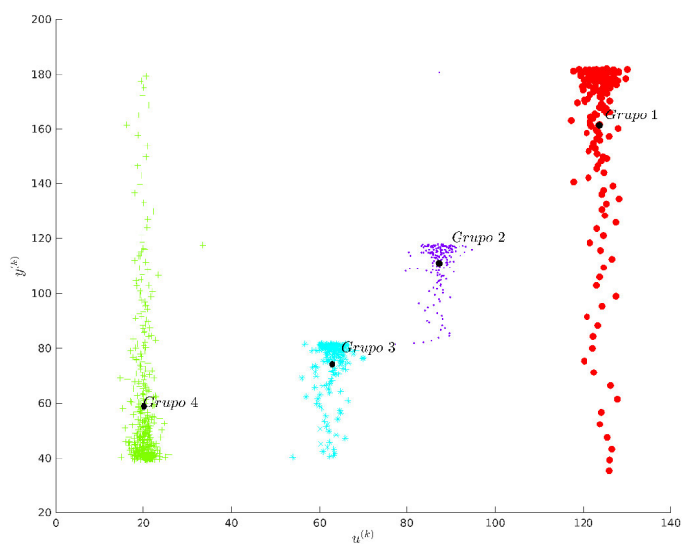
Fonte: Elaborado pelo Autor.

Figura 70: Grupos finais obtidos após a última iteração baseado no fator de aprendizado com taxa de aprendizado fixa, conforme (GEORGIEVA; FILEV, 2009; ANGELOV; FILEV; KASABOV, 2010).



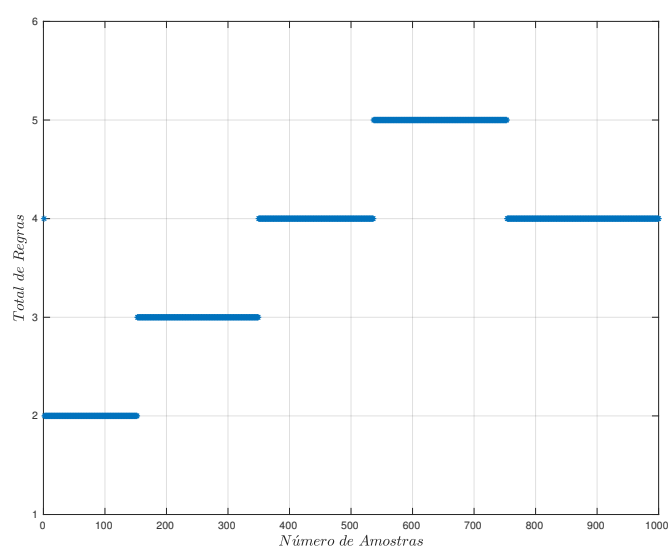
Fonte: Elaborado pelo Autor.

Figura 71: Grupos finais obtidos após a última iteração baseado no fator de aprendizado participativo utilizado nesta tese.



Fonte: Elaborado pelo Autor.

Figura 72: Número de regras evolutivas em função das amostras.



Fonte: Elaborado pelo Autor.

Tabela 4: Análise Comparativa das Taxas de Aprendizado utilizadas no Algoritmo de Agrupamento Nebuloso Evolutivo

Métrica de Avaliação	Fator de Aprendizado (GEORGIEVA; FILEV, 2009; ANGELOV; FILEV; KASABOV, 2010)	Fator de Aprendizado Participativo Proposto
Utheil	0,76338	0,38982
RMSE	7,2865	3,74712
VAF	98,19384	99,46730

5.5 Identificação de Sistemas Não Lineares

Nessa seção, a metodologia proposta de algoritmo evolutivo é avaliada a partir da identificação de um sistema não-linear é realizada uma comparação com metodologias vastamente citadas na literatura (ANGELOV; ZHOU, 2006; LUGHOFFER, 2008; RUBIO, 2009; LEMOS; CAMINHAS; GOMIDE, 2011).

O sistema não-linear a ser identificado é definido como:

$$y^{(k)} = \frac{y^{(k-1)}y^{(k-2)}(y^{(k-1)} - 0.5)}{1 + [y^{(k-1)}]^2 + [y^{(k-2)}]^2} + u^{(k-1)} \quad (5.15)$$

para $y^{(0)} = y^{(1)} = 0$.

A função parametrizada adotada nesta seção para fins de avaliação de desempenho é dada por

$$\hat{y}^{(k)} = f(u^{(k-2)}, u^{(k-1)}, \hat{y}^{(k-2)}, \hat{y}^{(k-1)}, \Theta) \quad (5.16)$$

onde

$$\Theta = [\theta_{u^{(k-2)}}, \theta_{u^{(k-1)}}, \theta_{\hat{y}^{(k-2)}}, \theta_{\hat{y}^{(k-1)}} \text{ e } \theta_0] \quad (5.17)$$

A equação (5.16) é a simulação livre ou predição de infinitos passos à frente, utilizada para avaliar o desempenho comparativo da metodologia proposta em relação a outras metodologias amplamente divulgadas na literatura.

Na etapa de identificação realizou-se o seguinte experimento: 1000 amostras foram geradas, sendo 50 para etapa em batelada e 950 para a etapa evolutiva. A entrada foi definida como um sinal aleatório de distribuição uniforme no intervalo de $u^{(k)} = [-1,5 + 1,5]$, o sinal das saídas $y^{(k)}$ e $\hat{y}^{(k)}$, real e estimados na etapa de identificação, respectivamente; são ilustrados na Figuras 73 e 74. Na etapa de validação foram geradas 300 amostras com $u^{(k)} = \sin\left(\frac{2\pi k}{25}\right)$.

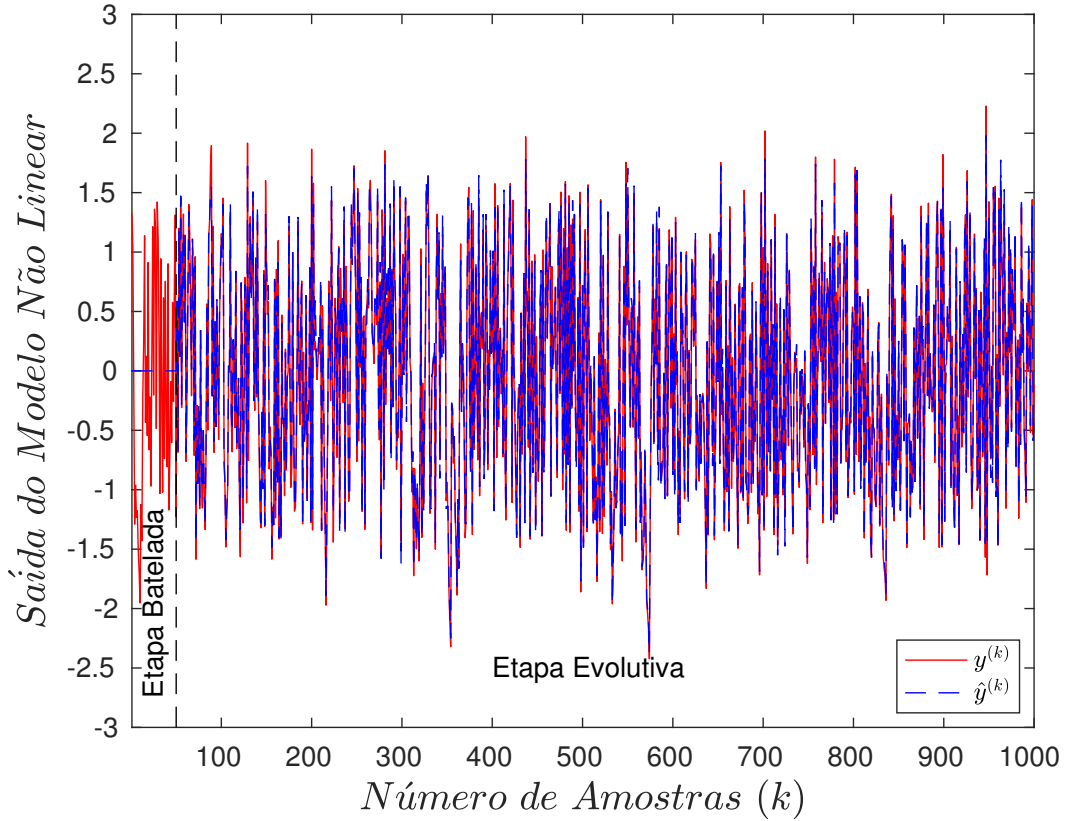
O modelo de inferência nebuloso Takagi–Sugeno utilizado na validação, é descrito a seguir:

$R_{(i)}$: SE $u^{(k-2)}$ É $\Gamma_{u^{(k-2)}}^{(i)}$ E $u^{(k-1)}$ É $\Gamma_{u^{(k-1)}}^{(i)}$ E $y^{(k-2)}$ É $\Gamma_{y^{(k-2)}}^{(i)}$ E $y^{(k-1)}$ É $\Gamma_{y^{(k-1)}}^{(i)}$ ENTÃO

$$\hat{y}_{(i)}^{(k)} = \theta_{u^{(k-2)}}^{(i)} u^{(k-2)} + \theta_{u^{(k-1)}}^{(i)} u^{(k-1)} + \theta_{y^{(k-2)}}^{(i)} \hat{y}^{(k-2)} + \theta_{y^{(k-1)}}^{(i)} \hat{y}^{(k-1)} + \theta_0 \quad (5.18)$$

onde $i = 1, 2, \dots, c$ e $k = 1, 2, \dots, N$; $\theta_{u^{(k-2)}}, \theta_{u^{(k-1)}}, \theta_{y^{(k-2)}}, \theta_{y^{(k-1)}}$ e θ_0 são os parâmetros do consequente a serem identificados na etapa evolutiva.

Figura 73: Identificação do Modelo Não-Linear. O eixo da ordenada representa a variação de amplitude das saídas reais e estimadas no processo de identificação, $y^{(k)}$ e $\hat{y}^{(k)}$, respectivamente.



Fonte: Elaborado pelo autor.

Após finalizar a etapa em batelada, estimação inicial do algoritmo, obteve-se o modelo nebuloso Takagi–Sugeno descrito a seguir:

$$R_{(1)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(1)} \\ \Gamma_{u^{(k-1)}}^{(1)} \\ \Gamma_{y^{(k-2)}}^{(1)} \\ \Gamma_{y^{(k-1)}}^{(1)} \end{bmatrix} \mathbf{ENTÃO} \hat{y}_{(1)}^{(k)} = \begin{bmatrix} 1,0287 \\ -0,0894 \\ 0,4621 \\ 0,1653 \\ 0,1267 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.19)$$

$$R_{(2)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(2)} \\ \Gamma_{u^{(k-1)}}^{(2)} \\ \Gamma_{y^{(k-2)}}^{(2)} \\ \Gamma_{y^{(k-1)}}^{(2)} \end{bmatrix} \text{ ENTÃO } \hat{y}_{(2)}^{(k)} = \begin{bmatrix} 0,9856 \\ 0,3202 \\ -0,3561 \\ 0,1314 \\ 0,0489 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.20)$$

À medida que as amostras são lidas, verifica-se que no final da etapa evolutiva o algoritmo identificou 6 grupos distintos, caracterizando-se como 6 regras do modelo nebuloso Takagi–Sugeno final. As regras de número 3, 4, 5 e 6 são criadas a partir das amostras 90, 356, 432, 674, respectivamente. A evolução do número de regras é ilustrado na Figura 74.

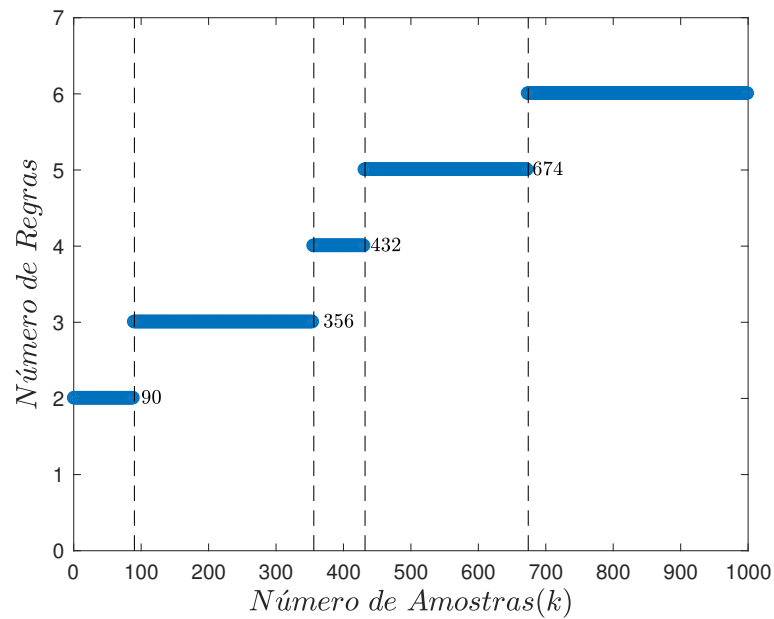
As amostras lidas são agrupadas e, à medida que o número de regras (grupo) é alterado, ocorre um redimensionamento das amostras nos grupos existentes. Consequentemente o número de amostras por grupo é alterado a cada novo conjunto de regras. O número de amostra por grupo é utilizado no algoritmo evolutivo, para que se possa identificar na etapa final deste, os grupos com uma única amostra, sendo um critério para se determinar quais regras deverão ser mescladas às regras existentes. Grupos com um único membro são mesclados aos grupos adjacentes. A evolução do número de amostras por grupos na etapa evolutiva é ilustrado na Figura 75.

As coordenadas do vetor de protótipos das regras, $V_{u^{(k-1)}}$, $V_{u^{(k-2)}}$, $V_{y^{(k-1)}}$, $V_{y^{(k-2)}}$, são apresentadas na Figura 76. A cada nova regra inserida o algoritmo ajusta as coordenadas do vetor de protótipos. A evolução dos parâmetros do consequente, à medida que o número de regras evolui em função das amostras lidas, é apresentado na Figura 77.

O desempenho comparativo para o problema da identificação do modelo nebuloso Takagi–Sugeno identificado é ilustrado na Tabela 5. O modelo proposto apresentou resultado satisfatório em relação aos modelos comparativos que são amplamente utilizados na literatura em problemas de identificação de sistemas.

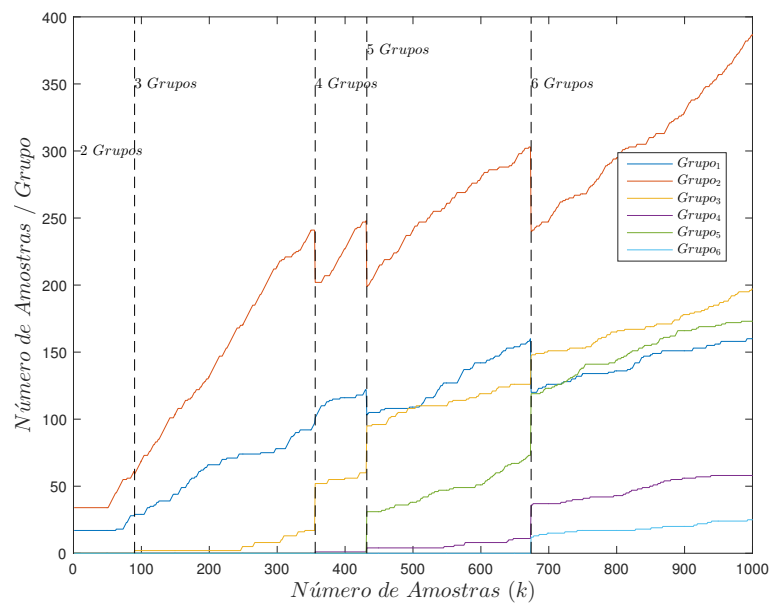
A validação do modelo identificado é ilustrada na Figura 78, este é obtido a partir do agrupamento nebuloso evolutivo baseado em máxima verossimilhança com o cálculo da probabilidade *a priori* mensurado a partir da equação (4.40). Porém, para fins de comprovação da proposta desenvolvida a respeito do cálculo desta probabilidade, na Figura 79 é ilustrado a validação do modelo identificado quando a probabilidade *a priori* é mensurada a partir da equação (4.39). O desempenho comparativo de ambas as metodologias são apresentadas na Tabela 6 a fim de demonstrar o efeito da mudança proposta na tese em relação ao desempenho final do modelo.

Figura 74: Evolução do número de regras em função do número de amostras lidas.



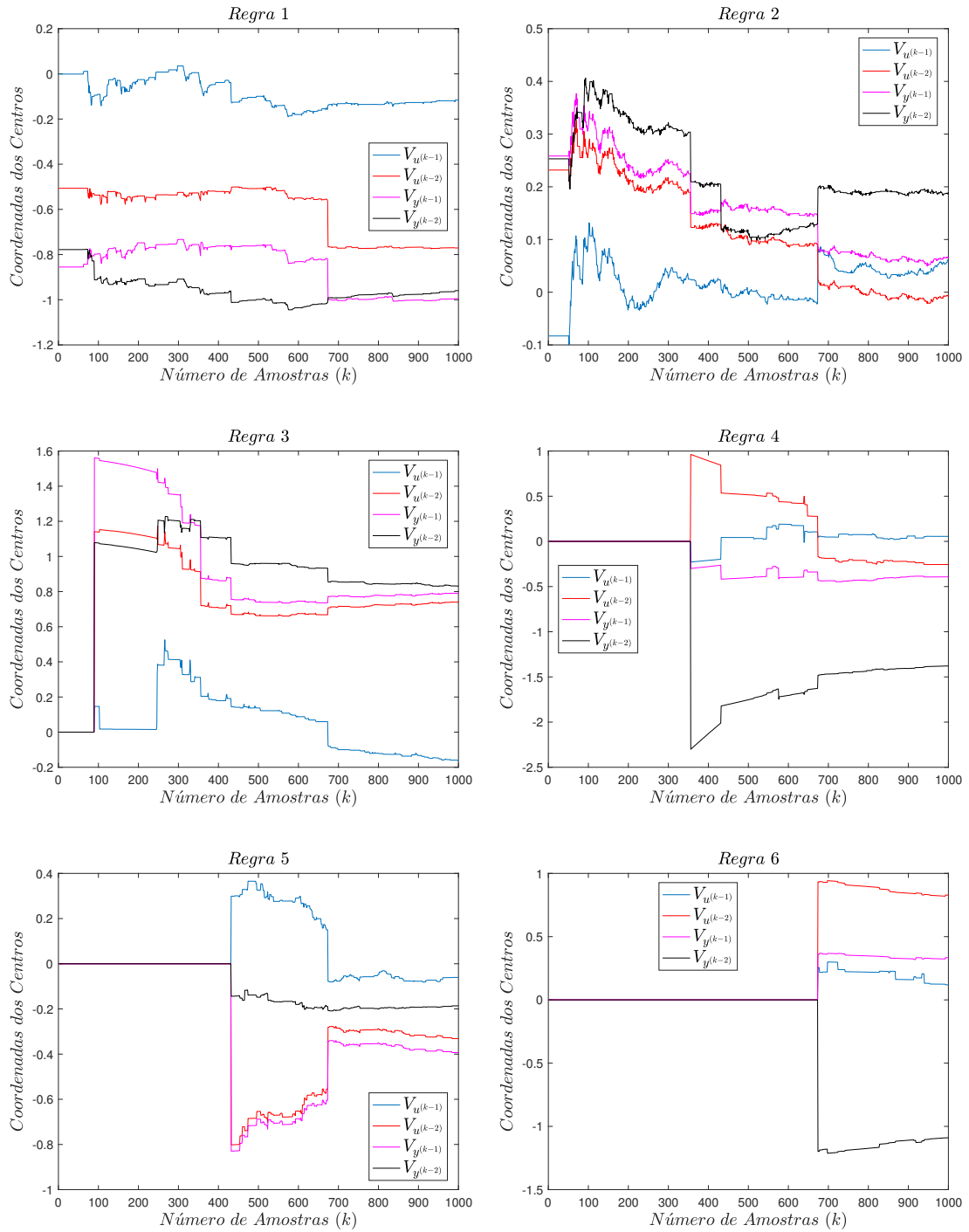
Fonte: Elaborado pelo autor.

Figura 75: Evolução do número de amostras/grupo em função do número de amostras lidas.



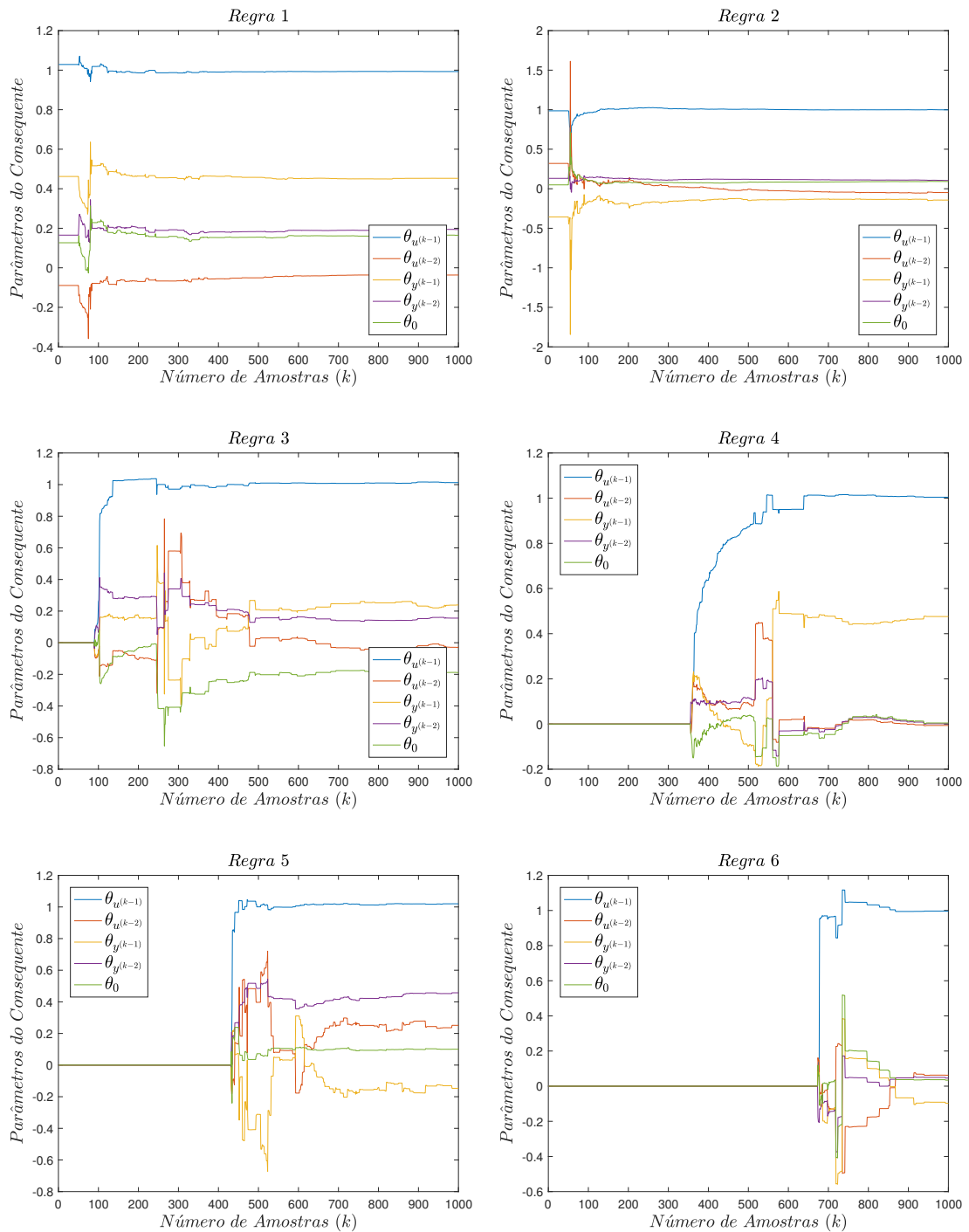
Fonte: Elaborado pelo autor.

Figura 76: Coordenadas do vetor de protótipos das Regras_(i) para $i = 1, 2, \dots, 6$.



Fonte: Elaborado pelo autor.

Figura 77: Evolução dos parâmetros do consequente para as regras das $Regras_i$ para $i = 1, 2, \dots, 6$.



Fonte: Elaborado pelo autor.

O modelo nebuloso Takagi–Sugeno após a estimação evolutiva, na última recursão, quando todos as amostras estão incorporados aos centros mais próximos, é dado por:

$$R_{(1)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(1)} \\ \Gamma_{u^{(k-1)}}^{(1)} \\ \Gamma_{y^{(k-2)}}^{(1)} \\ \Gamma_{y^{(k-1)}}^{(1)} \end{bmatrix} \text{ENTÃO } \hat{y}_{(1)}^{(k)} = \begin{bmatrix} 0,9904 \\ -0,3088 \\ 0,2096 \\ 0,2506 \\ 0,0187 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.21)$$

$$R_{(2)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(2)} \\ \Gamma_{u^{(k-1)}}^{(2)} \\ \Gamma_{y^{(k-2)}}^{(2)} \\ \Gamma_{y^{(k-1)}}^{(2)} \end{bmatrix} \text{ENTÃO } \hat{y}_{(2)}^{(k)} = \begin{bmatrix} 1,0075 \\ -0,1797 \\ 0,2384 \\ 0,1556 \\ -0,1892 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.22)$$

$$R_{(3)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(3)} \\ \Gamma_{u^{(k-1)}}^{(3)} \\ \Gamma_{y^{(k-2)}}^{(3)} \\ \Gamma_{y^{(k-1)}}^{(3)} \end{bmatrix} \text{ENTÃO } \hat{y}_{(3)}^{(k)} = \begin{bmatrix} 1,0123 \\ -0,0278 \\ 0,2384 \\ 0,1556 \\ -0,1892 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.23)$$

$$R_{(4)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(4)} \\ \Gamma_{u^{(k-1)}}^{(4)} \\ \Gamma_{y^{(k-2)}}^{(4)} \\ \Gamma_{y^{(k-1)}}^{(4)} \end{bmatrix} \text{ENTÃO } \hat{y}_{(4)}^{(k)} = \begin{bmatrix} 1,0042 \\ -0,0062 \\ 0,4759 \\ 0,0014 \\ 0,0046 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.24)$$

$$R_{(5)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(5)} \\ \Gamma_{u^{(k-1)}}^{(5)} \\ \Gamma_{y^{(k-2)}}^{(5)} \\ \Gamma_{y^{(k-1)}}^{(5)} \end{bmatrix} \text{ENTÃO } \hat{y}_{(5)}^{(k)} = \begin{bmatrix} 1,0202 \\ -0,2514 \\ -0,1470 \\ 0,4573 \\ 0,1004 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.25)$$

$$R_{(6)} : \mathbf{SE} \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ y^{(k-2)} \\ y^{(k-1)} \end{bmatrix} \hat{\mathbf{E}} \begin{bmatrix} \Gamma_{u^{(k-2)}}^{(6)} \\ \Gamma_{u^{(k-1)}}^{(6)} \\ \Gamma_{y^{(k-2)}}^{(6)} \\ \Gamma_{y^{(k-1)}}^{(6)} \end{bmatrix} \text{ENTÃO } \hat{y}_{(6)}^{(k)} = \begin{bmatrix} 1,9959 \\ +0,0625 \\ -0,0973 \\ 0,0460 \\ 0,0340 \end{bmatrix}^T \begin{bmatrix} u^{(k-2)} \\ u^{(k-1)} \\ \hat{y}^{(k-2)} \\ \hat{y}^{(k-1)} \\ 1 \end{bmatrix} \quad (5.26)$$

Tabela 5: Análise Comparativa para o Problema da Identificação

Modelo	Regras	NDEI
eTS (LUGHOFFER, 2008)	7	0,1036
xTS (ANGELOV; ZHOU, 2006)	7	0,0937
DENFIS (LEMO; CAMINHAS; GOMIDE, 2011)	7	0,0849
eFT (RUBIO, 2009)	7	0,0657
eMG (LEMO; CAMINHAS; GOMIDE, 2011)	7	0,0507
Modelo Proposto	6	0,0413

Tabela 6: Análise Comparativa para o Problema da Identificação em Função do Cálculo da Probabilidade *a Priori*.

Métricas de Avaliação	Probabilidade <i>a priori</i> (4.40)	Probabilidade <i>a priori</i> (4.39)
Utheil	0,35186	3,98778
RMSE	0,04329	0,53174
NDEI	0,04130	0,50679
VAF	99,90261	80,69425

Figura 78: Validação do modelo para o sistema não linear.

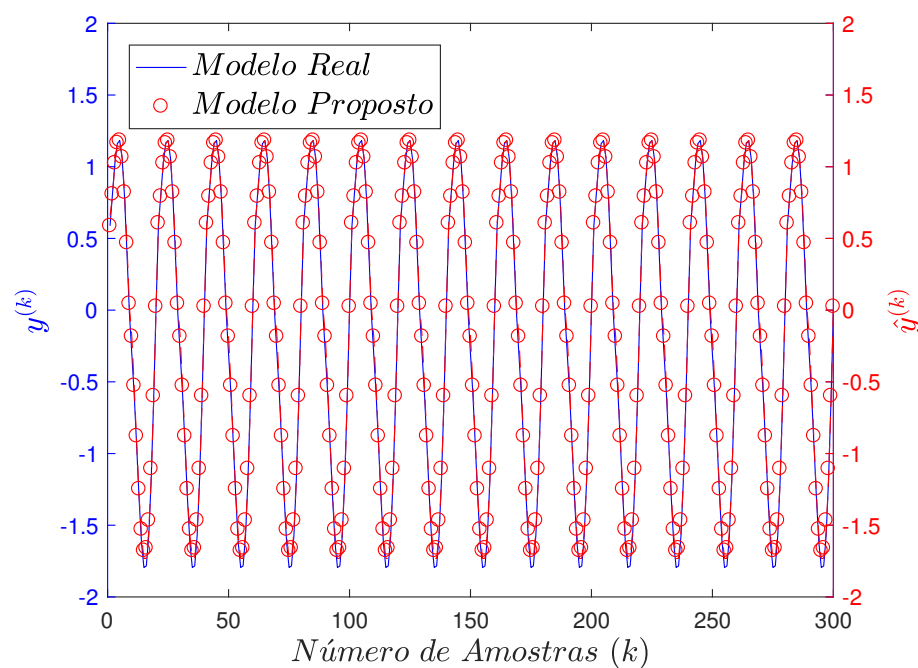
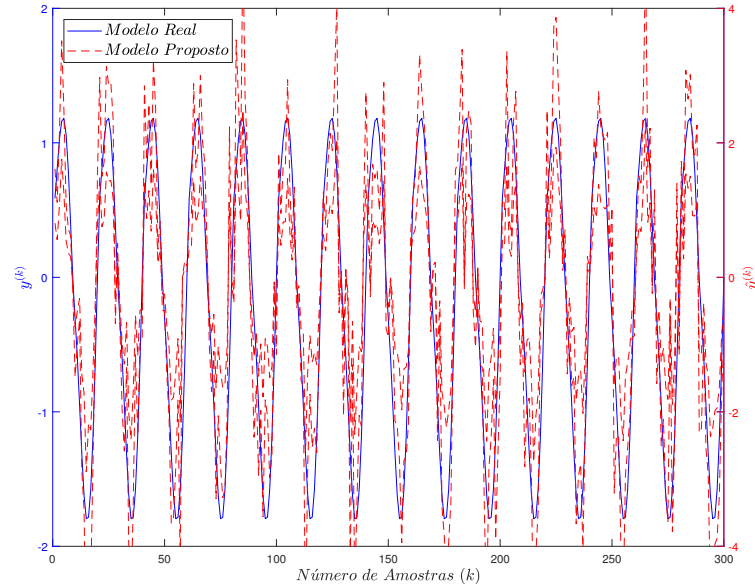


Figura 79: Gráfico comparativo do modelo proposto obtido a partir do cálculo da probabilidade *a priori* conforme proposto em (BABUSKA, 1998).



5.6 Identificação de um Helicóptero - 2DOF

Nesta seção, a identificação de um sistema dinâmico não linear experimental baseado na metodologia proposta será apresentada. O sistema dinâmico não linear experimental utilizado é o helicóptero 2-DOF da empresa Quanser Consulting Inc., como mostrado na Figura 80. O Helicóptero 2-DOF tem 2 graus de liberdade denominados de rotação *azimute* e rotação *elevação* com as seguintes características (GADEWADIKAR; HORVAT; KULJACA, 2013; QUANSER, 2014):

1. Intervalo da tensão do ângulo *azimute*, $\mathbf{V}_\vartheta \in [-24V, +24V]$.
2. Intervalo da tensão do ângulo *elevação*, $\mathbf{V}_\varphi \in [-15V, +15V]$.
3. Intervalo do ângulo *azimute*, $\vartheta \in [-40^\circ, +40^\circ]$.
4. Intervalo do ângulo *elevação*, $\varphi \in [-360^\circ, +360^\circ]$.

Figura 80: Helicóptero 2-DOF



Fonte: Quanser

Os sinais utilizados para modelagem do Helicóptero 2DOF são: ângulo de elevação $\vartheta^{(k)} > 0$ para $F_p > 0$; o ângulo de azimute $\varphi^{(k)} > 0$ para $F_y > 0$, no instante k (ZAERI et al., 2012; RAHMAN; SHOUMY, 2012). A modelagem caixa preta do helicóptero 2DOF, mostrado na Figura 81, com erro nas variáveis para ilustrar o desempenho do algoritmo de identificação *online* baseado na variável instrumental nebulosa evolutiva, é apresentada.

As estruturas utilizadas para gerar dados de identificação e validação são ilustradas nas Figuras 82 e 83, respectivamente. Observa-se que as tensões de controle de entrada e os ângulos de observação de saída foram corrompidos por um ruído autocorrelacionado, na etapa de identificação. Na etapa de validação os sinais de entrada e saída estão livres da perturbação para fins de avaliação do modelo obtido.

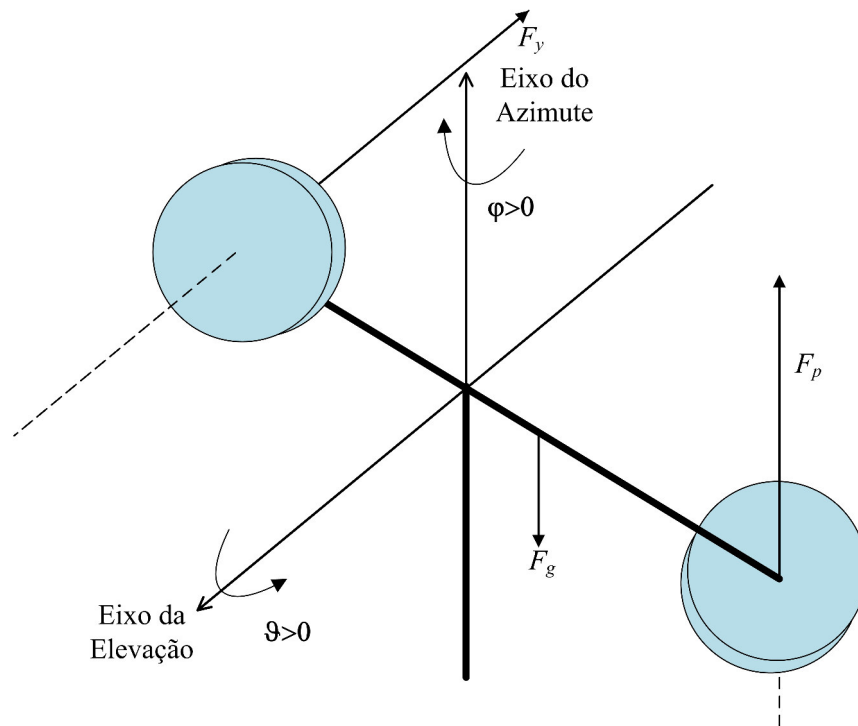
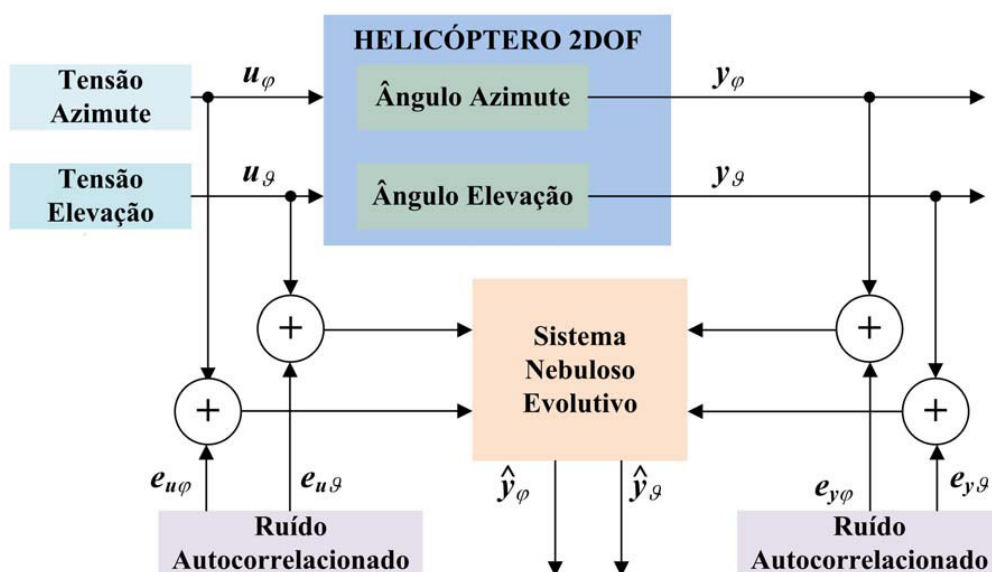


Figura 81: Diagrama de corpo livre do Helicóptero 2DOF.

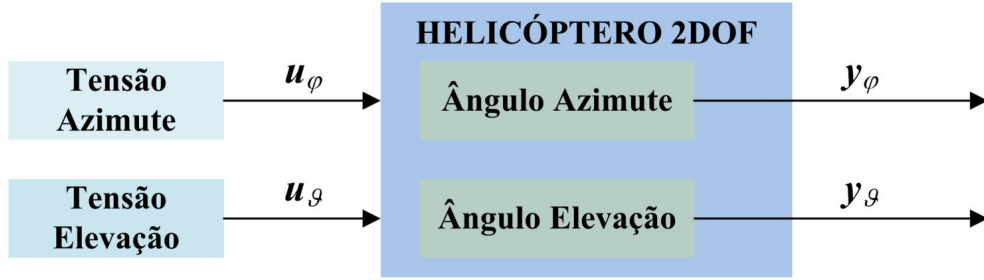
Fonte: Elaborado pelo autor.

Figura 82: Estrutura utilizada para gerar dados de identificação do modelo.



Fonte: Elaborado pelo autor.

Figura 83: Estrutura utilizada para gerar dados de validação do modelo.



Fonte: Elaborado pelo autor.

Os sinais de entrada e saída corrompidos com o ruído autocorrelacionado, são dados por,

$$\mathbf{u}_{e_\varphi} = \mathbf{u}_\varphi + \mathbf{e}_{\mathbf{u}_\varphi} \quad (5.27)$$

$$\mathbf{u}_{e_\vartheta} = \mathbf{u}_\vartheta + \mathbf{e}_{\mathbf{u}_\vartheta} \quad (5.28)$$

$$\mathbf{y}_{e_\varphi} = \mathbf{y}_\varphi + \mathbf{e}_{\mathbf{y}_\varphi} \quad (5.29)$$

$$\mathbf{y}_{e_\vartheta} = \mathbf{y}_\vartheta + \mathbf{e}_{\mathbf{y}_\vartheta} \quad (5.30)$$

de modo que,

$$e_{\mathbf{u}_\varphi}^{(k)} = \frac{1 - 0.6z^{-1} + 0.4z^{-2}}{1 - 0.95z^{-1} + 0.9506z^{-2}} e_1^{(k)} \quad (5.31)$$

$$e_{\mathbf{u}_\vartheta}^{(k)} = \frac{1 - 0.6z^{-1} + 0.4z^{-2}}{1 - 0.95z^{-1} + 0.9506z^{-2}} e_2^{(k)} \quad (5.32)$$

$$e_{\mathbf{y}_\varphi}^{(k)} = \frac{1 - 0.6z^{-1} + 0.4z^{-2}}{1 - 0.95z^{-1} + 0.9506z^{-2}} e_3^{(k)} \quad (5.33)$$

$$e_{\mathbf{y}_\vartheta}^{(k)} = \frac{1 - 0.6z^{-1} + 0.4z^{-2}}{1 - 0.95z^{-1} + 0.9506z^{-2}} e_4^{(k)} \quad (5.34)$$

com o sinal de ruído expresso por,

$$e_1^{(k)} = \mathbf{N}(0, \sigma_1^2) \quad (5.35)$$

$$e_2^{(k)} = \mathbf{N}(0, \sigma_2^2) \quad (5.36)$$

$$e_3^{(k)} = \mathbf{N}(0, \sigma_3^2) \quad (5.37)$$

$$e_4^{(k)} = \mathbf{N}(0, \sigma_4^2) \quad (5.38)$$

As funções parametrizadas adotadas nesta seção para fins de avaliação de desempenho dos sinais $y_\varphi^{(k)}$ e $y_\vartheta^{(k)}$ são expressas por

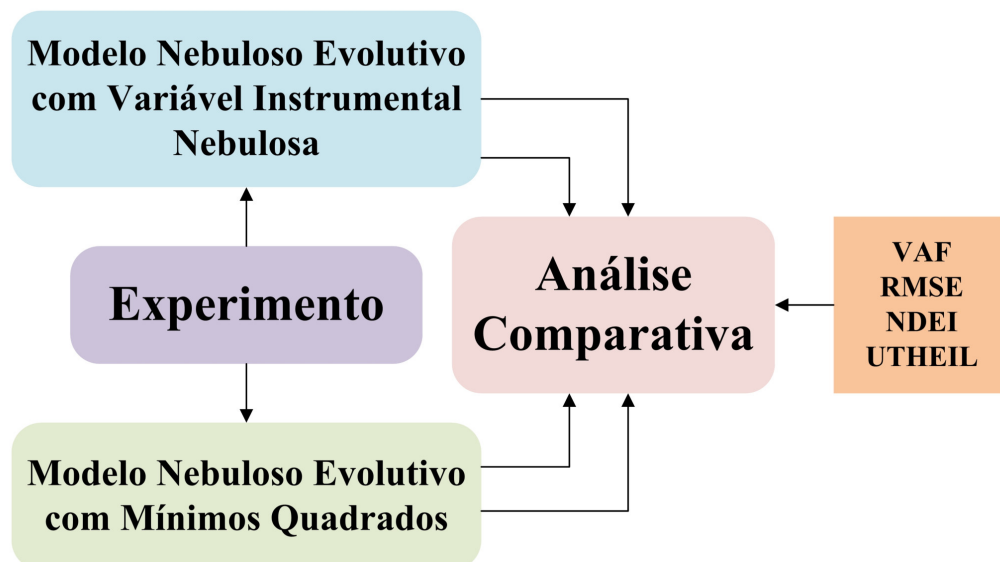
$$\begin{bmatrix} \hat{y}_\varphi^{(k)} \\ \hat{y}_\vartheta^{(k)} \end{bmatrix} = \begin{bmatrix} f(u_\varphi^{(k-2)}, u_\varphi^{(k-1)}, u_\vartheta^{(k-2)}, u_\vartheta^{(k-1)}, \hat{y}_\varphi^{(k-2)}, \hat{y}_\varphi^{(k-1)}, \hat{y}_\vartheta^{(k-2)}, \hat{y}_\vartheta^{(k-1)}, \Theta_\varphi) \\ f(u_\varphi^{(k-2)}, u_\varphi^{(k-1)}, u_\vartheta^{(k-2)}, u_\vartheta^{(k-1)}, \hat{y}_\varphi^{(k-2)}, \hat{y}_\varphi^{(k-1)}, \hat{y}_\vartheta^{(k-2)}, \hat{y}_\vartheta^{(k-1)}, \Theta_\vartheta) \end{bmatrix} \quad (5.39)$$

de modo que

$$\begin{bmatrix} \Theta_{\varphi} \\ \Theta_{\vartheta} \end{bmatrix} = \begin{bmatrix} \theta_{\varphi_1} & \theta_{\varphi_2} & \theta_{\varphi_3} & \theta_{\varphi_4} & \theta_{\varphi_5} & \theta_{\varphi_6} & \theta_{\varphi_7} & \theta_{\varphi_8} \\ \theta_{\vartheta_1} & \theta_{\vartheta_2} & \theta_{\vartheta_3} & \theta_{\vartheta_4} & \theta_{\vartheta_5} & \theta_{\vartheta_6} & \theta_{\vartheta_7} & \theta_{\vartheta_8} \end{bmatrix} \quad (5.40)$$

A identificação do helicóptero 2-DOF é realizada em um conjunto de dados de 220 amostras com os sinais entrada e saída corrompidos por ruído autocorrelacionado na faixa de $\sigma^2 \in [0,05 \ 0,20]$, o qual equivale a uma relação de sinal/ruído (SNR) de $[45, 5]$ dB. O experimento com os dados é composto por 1000 realizações e para cada bloco desta, o ruído é inserido nas variáveis de entrada e saída de acordo com a faixa citada. Os parâmetros do antecedente e a estrutura do modelo nebuloso evolutivo são os mesmos nos experimentos, enquanto os parâmetros do consequente são obtidos pela variável instrumental nebulosa inserido no contexto evolutivo e pelo método dos mínimos quadrados. Consequentemente, os resultados obtidos por ambos algoritmos de estimação paramétrica serão úteis para as conclusões a respeito da metodologia proposta diante de um ambiente com ruído. Os critérios que são aplicados para a validação do modelo proposto são: *Variance Accounted For (VAF)* (5.6), a Raiz do Erro Quadrático Médio (RMSE) (5.4), o Coeficiente Não-Dimensional (NDEI) (5.5) e o Coeficiente U-THEIL (5.1). O diagrama de bloco representativo na Figura 84 ilustra o procedimento utilizado para análise de desempenho do experimento realizado no helicóptero.

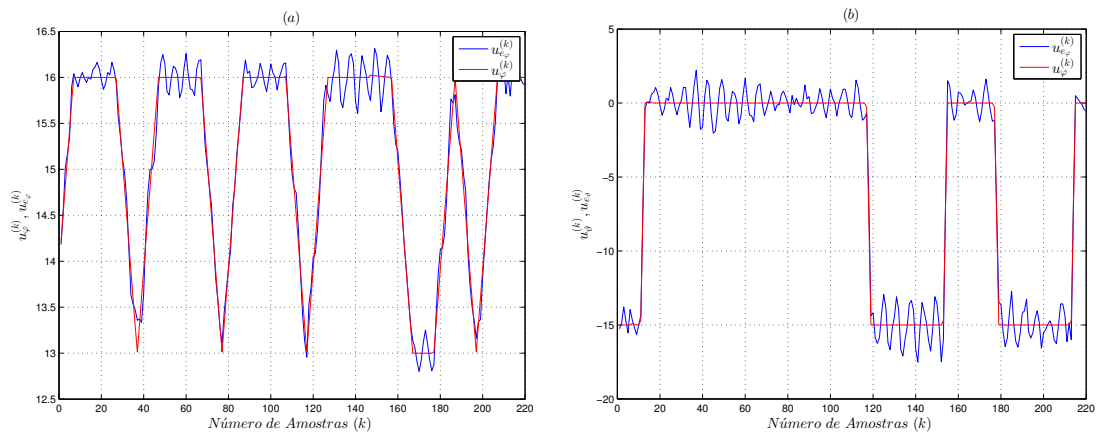
Figura 84: Diagrama de bloco da análise de desempenho do experimento realizado no Helicóptero 2DOF.



Fonte: Adaptado de (SERRA; BOTTURA, 2009)

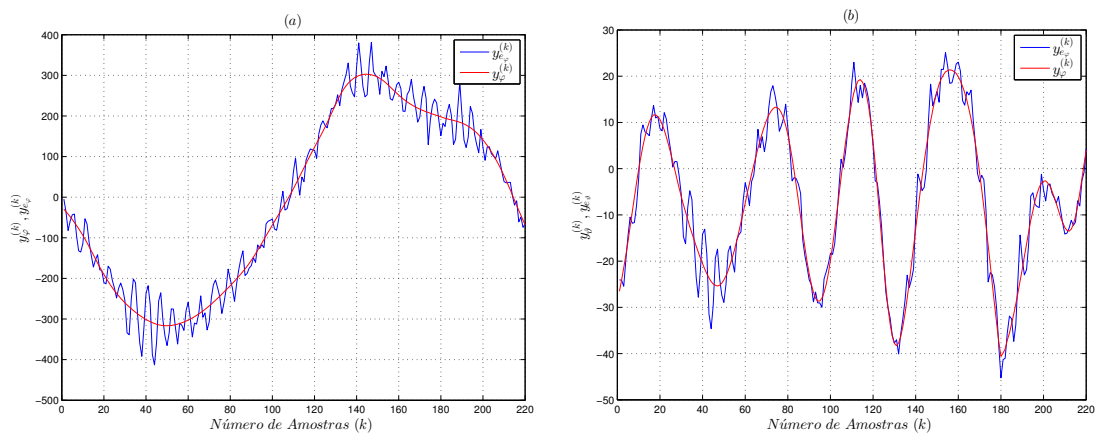
O sinais de entrada e saída utilizados no experimento e que foram corrompidos com um sinal de ruído de $\sigma^2 = 0,20$ são ilustrados na Figura 85.

Figura 85: Sinal de entrada para tensão de controle do ângulo de azimute (a), tensão de controle do ângulo de elevação (b).



Fonte: Elaborado pelo autor.

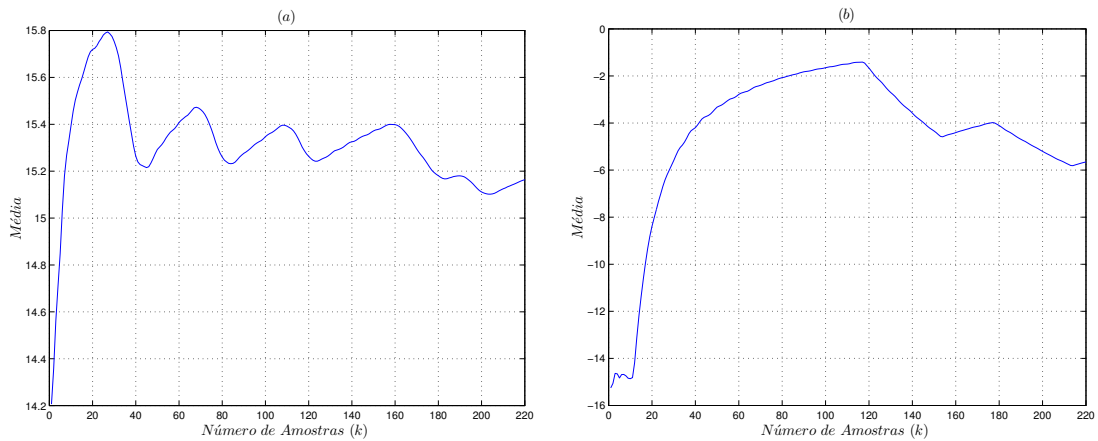
Figura 86: Sinal de entrada para ângulo de azimute (a), ângulo de elevação (b).



Fonte: Elaborado pelo autor.

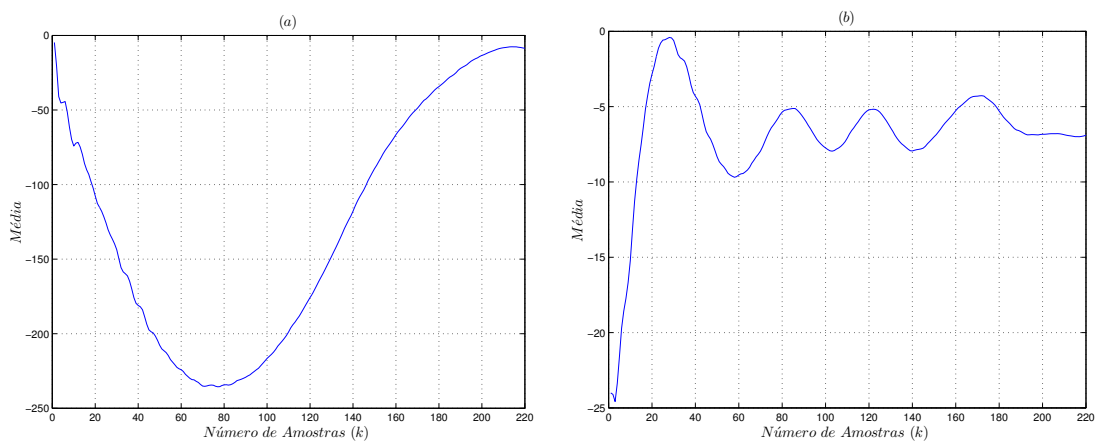
A verificação de não-estacionariedade dos dados do helicóptero foram baseadas nas avaliações das médias e variâncias dos sinais de entrada e saída; além disso, testou-se a hipótese nula de não-estacionariedade do teste de *Dickey-Fuller* (MORETTIN, 2006) com base na função de teste *Adftest* do Matlab versão 2016, resultados na Tabela 7. A avaliação da média e variância dos sinais ao longo do período k , são ilustrados nas Figuras 87–90.

Figura 87: Média móvel do sinal de entrada para tensão de controle do ângulo de azimute em (a) e tensão de controle do ângulo de elevação em (b).



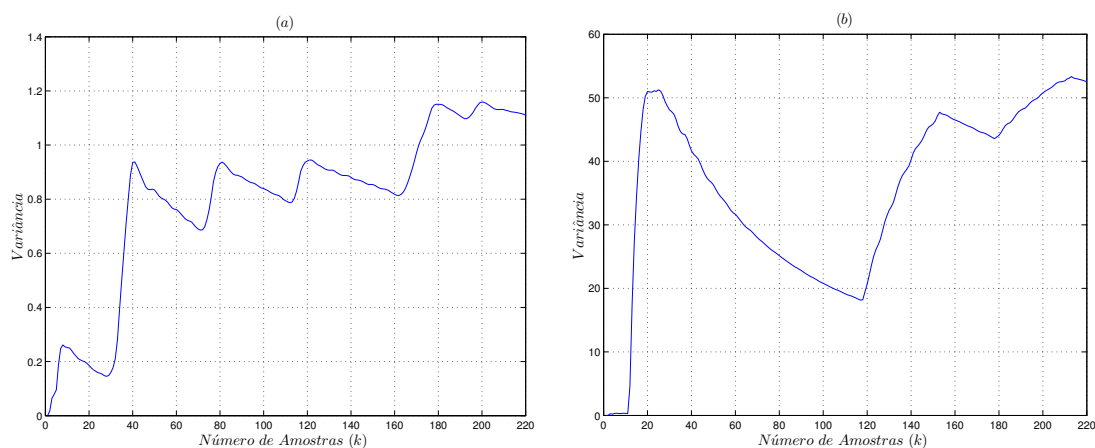
Fonte: Elaborado pelo autor.

Figura 88: Média móvel do sinal de entrada para ângulo de azimute em (a) e ângulo de elevação em (b).



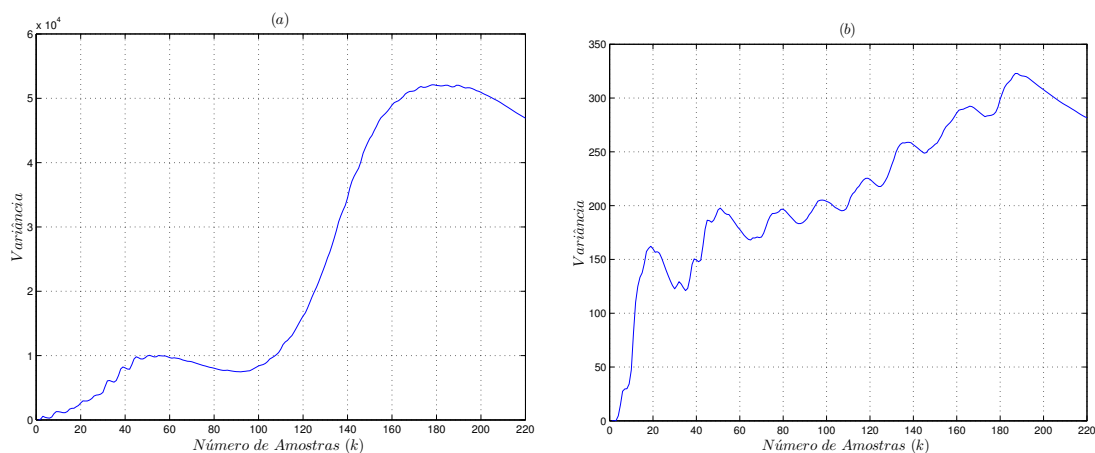
Fonte: Elaborado pelo autor.

Figura 89: Variância móvel do sinal de entrada para tensão de controle do ângulo de azimute em (a) e tensão de controle do ângulo de elevação em (b).



Fonte: Elaborado pelo autor.

Figura 90: Variância móvel do sinal de entrada para ângulo de azimute em (a) e ângulo de elevação em (b).



Fonte: Elaborado pelo autor.

Tabela 7: Teste de Hipótese nula do Teste *Dickey-Fuller* (MORETTIN, 2006).

Sinal	Resultado da Hipótese Nula
Tensão de Azimute	Não-Estacionária
Tensão de Elevação	Não-Estacionária
Ângulo de Azimute de Azimute	Não-Estacionária
Ângulo de Elevação de Azimute	Não-Estacionária

Fonte: Elaborado pelo autor.

Os resultados apresentados nas Figuras 87–90 indicam que a média e variância dos sinais utilizados para identificação tem uma características não-estacionária, fato este também observado no Teste de *Dickey-Fuller* (MORETTIN, 2006) na Tabela 7.

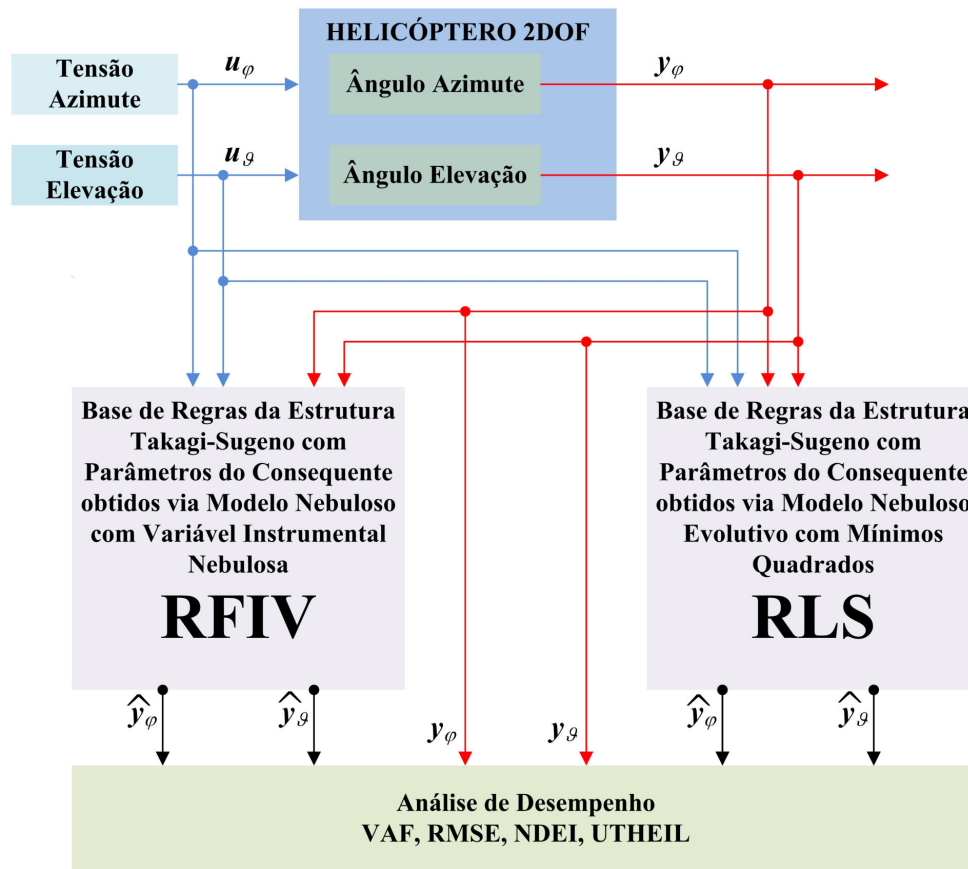
A base de regras do modelo nebuloso evolutivo inserido na estrutura de *Takagi-Sugeno*, é da forma

$$R_{(i)} : \text{SE} \begin{bmatrix} u_{\varphi}^{(k-2)} \\ u_{\varphi}^{(k-1)} \\ u_{\vartheta}^{(k-2)} \\ u_{\vartheta}^{(k-1)} \\ y_{\varphi}^{(k-2)} \\ y_{\varphi}^{(k-1)} \\ y_{\vartheta}^{(k-2)} \\ y_{\vartheta}^{(k-1)} \end{bmatrix} \hat{\text{É}} \begin{bmatrix} \Gamma_{u_{\varphi}}^{(i)(k-2)} \\ \Gamma_{u_{\varphi}}^{(i)(k-1)} \\ \Gamma_{u_{\vartheta}}^{(i)(k-2)} \\ \Gamma_{u_{\vartheta}}^{(i)(k-1)} \\ \Gamma_{y_{\varphi}}^{(i)(k-2)} \\ \Gamma_{y_{\varphi}}^{(i)(k-1)} \\ \Gamma_{y_{\vartheta}}^{(i)(k-2)} \\ \Gamma_{y_{\vartheta}}^{(i)(k-1)} \end{bmatrix} \text{ENTÃO} \begin{bmatrix} \hat{y}_{\varphi}^{(k)} \\ \hat{y}_{\vartheta}^{(k)} \end{bmatrix} = \begin{bmatrix} \theta_{\varphi_1}^{(i)} & \theta_{\vartheta_1}^{(i)} \\ \theta_{\varphi_2}^{(i)} & \theta_{\vartheta_2}^{(i)} \\ \theta_{\varphi_3}^{(i)} & \theta_{\vartheta_3}^{(i)} \\ \theta_{\varphi_4}^{(i)} & \theta_{\vartheta_4}^{(i)} \\ \theta_{\varphi_5}^{(i)} & \theta_{\vartheta_5}^{(i)} \\ \theta_{\varphi_6}^{(i)} & \theta_{\vartheta_6}^{(i)} \\ \theta_{\varphi_7}^{(i)} & \theta_{\vartheta_7}^{(i)} \\ \theta_{\varphi_8}^{(i)} & \theta_{\vartheta_8}^{(i)} \end{bmatrix}^T \begin{bmatrix} u_{\varphi}^{(k-2)} \\ u_{\varphi}^{(k-1)} \\ u_{\vartheta}^{(k-2)} \\ u_{\vartheta}^{(k-1)} \\ \hat{y}_{\varphi}^{(k-2)} \\ \hat{y}_{\varphi}^{(k-1)} \\ \hat{y}_{\vartheta}^{(k-2)} \\ \hat{y}_{\vartheta}^{(k-1)} \end{bmatrix} \quad (5.41)$$

onde $\Gamma_*^{(i)}$ são conjuntos nebulosos obtidos a partir do modelo nebuloso evolutivo associados à $(u_{\varphi}^{(k-2)}, u_{\varphi}^{(k-1)}, u_{\vartheta}^{(k-2)}, u_{\vartheta}^{(k-1)}, \hat{y}_{\varphi}^{(k-2)}, \hat{y}_{\varphi}^{(k-1)}, \hat{y}_{\vartheta}^{(k-2)}, \hat{y}_{\vartheta}^{(k-1)})$. Os parâmetros do consequente são obtidos via algoritmo de variável instrumental nebulosa inserido no contexto evolutivo baseado em máxima verossimilhança, conforme equações (4.100)–(4.105); e o método dos mínimos quadrados recursivo, conforme as equações (3.11)–(3.13).

Na etapa da validação obteve-se as saídas estimadas \hat{y}_{φ} e \hat{y}_{ϑ} por ambos algoritmos. Elas comparadas com as saídas reais do modelo experimental, y_{φ} e y_{ϑ} , para análise de desempenho baseados nas métricas VAF, RMSE, NDEI e U-THEIL, conforme ilustrado na Figura 91.

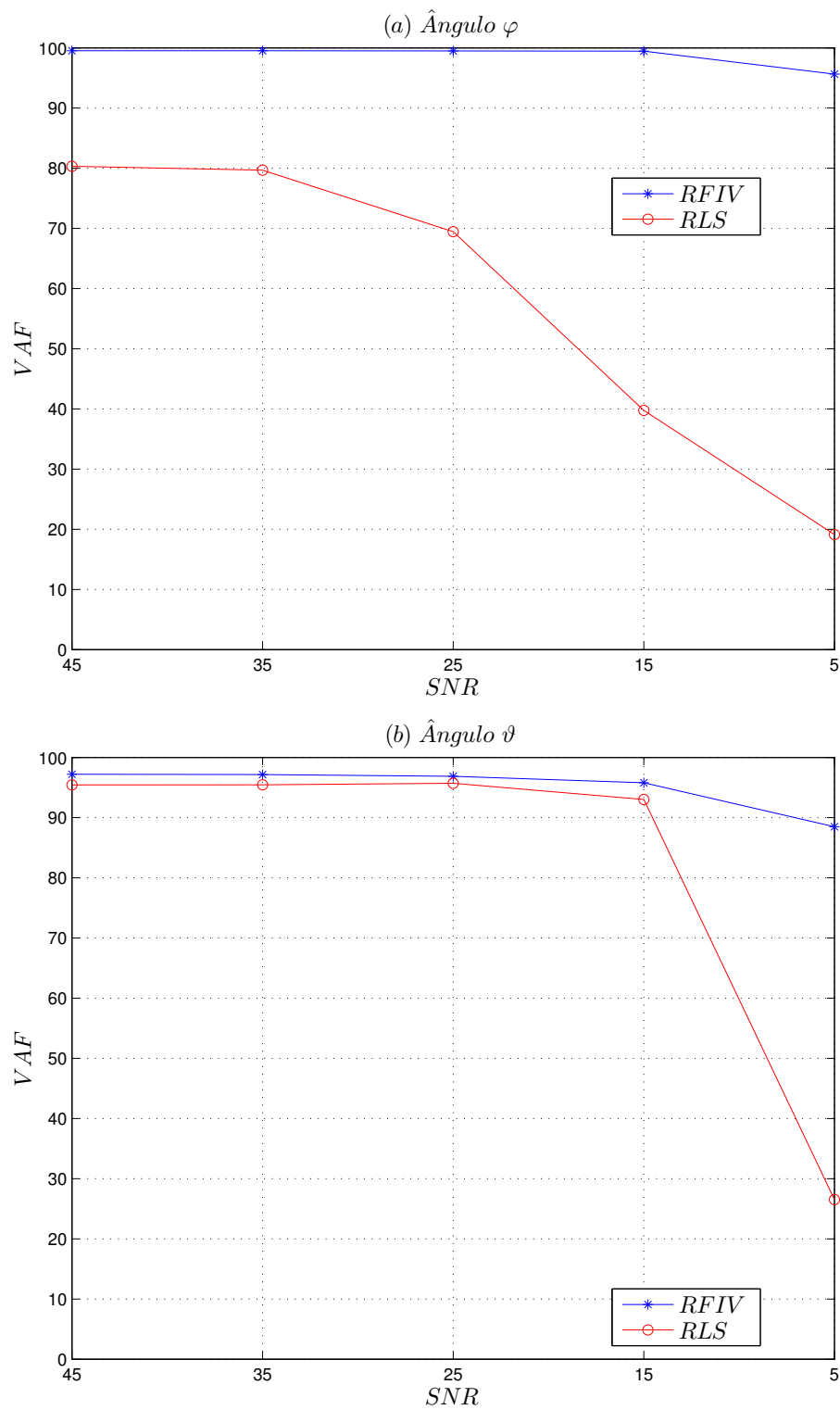
Figura 91: Diagrama de bloco da análise de desempenho na etapa de validação dos modelos obtidos.



Fonte: Elaborado pelo autor.

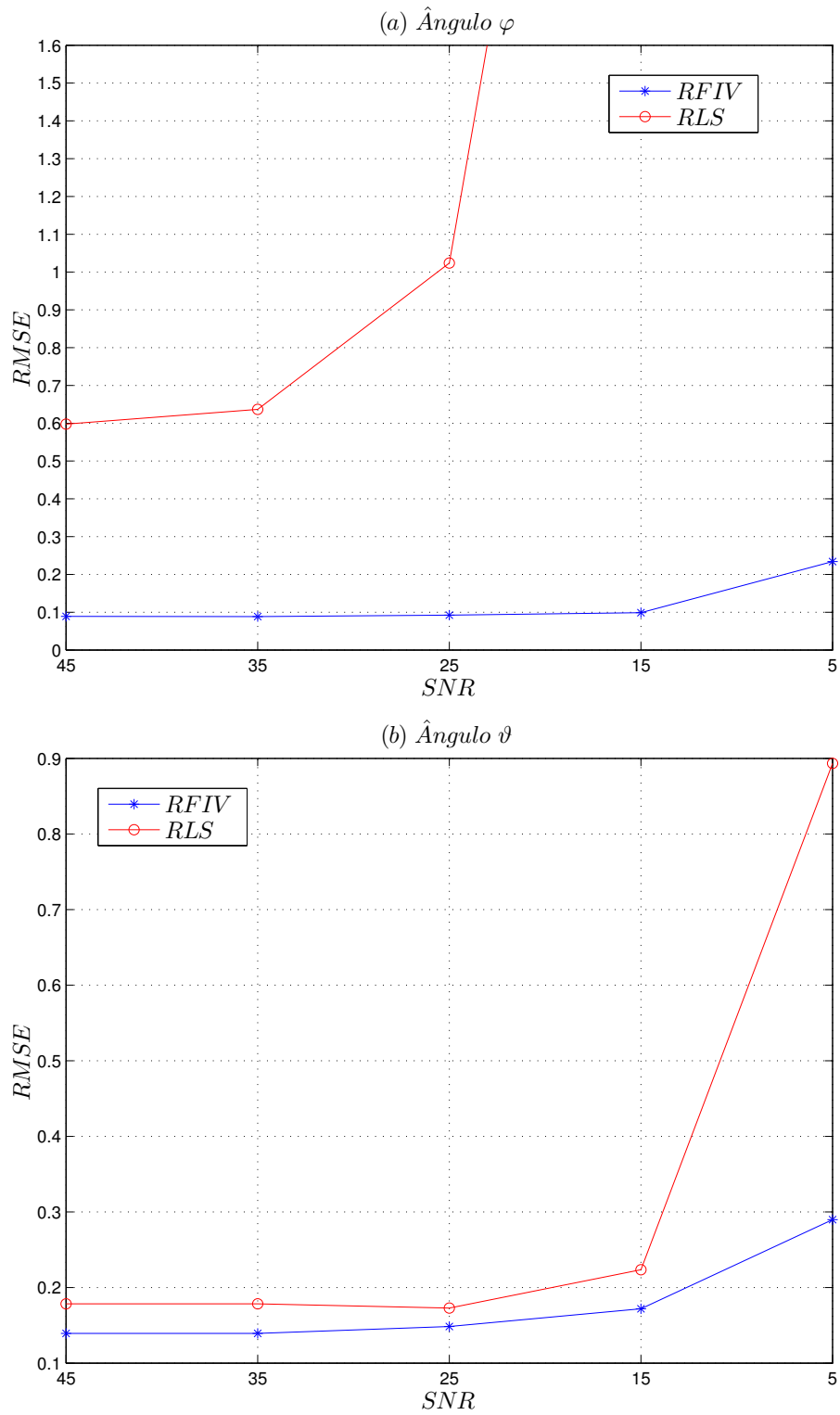
Os resultados da análise de desempenho dos modelos, observados nos diagramas representativos 84 e 91, baseado nas métricas de avaliação, são ilustradas nas Figuras 92 a 95. Perceba que o algoritmo proposto apresenta melhor desempenho que o algoritmo baseado em mínimos quadrados recursivos inserido neste contexto nebuloso evolutivo, sendo mais robusto ao ruído. Isto é devido à matriz de variável instrumental nebulosa inserida num contexto evolutivo de máxima verossimilhança, a qual satisfaz as condições de convergência propostas tão bem quando possível. Na presença de uma baixa variância ao ruído, ambos algoritmos apresentaram desempenho similares. Porém, quando a variância do ruído é incrementada, a matriz de variável instrumental nebulosa proposta satisfaz as condições de convergência, o que, conseqüentemente, torna o algoritmo mais robusto ao ruído autocorrelacionado.

Figura 92: Análise de robustez usada no VAF(%) (a) Ângulo de Azimute e (b) Ângulo de Elevação.



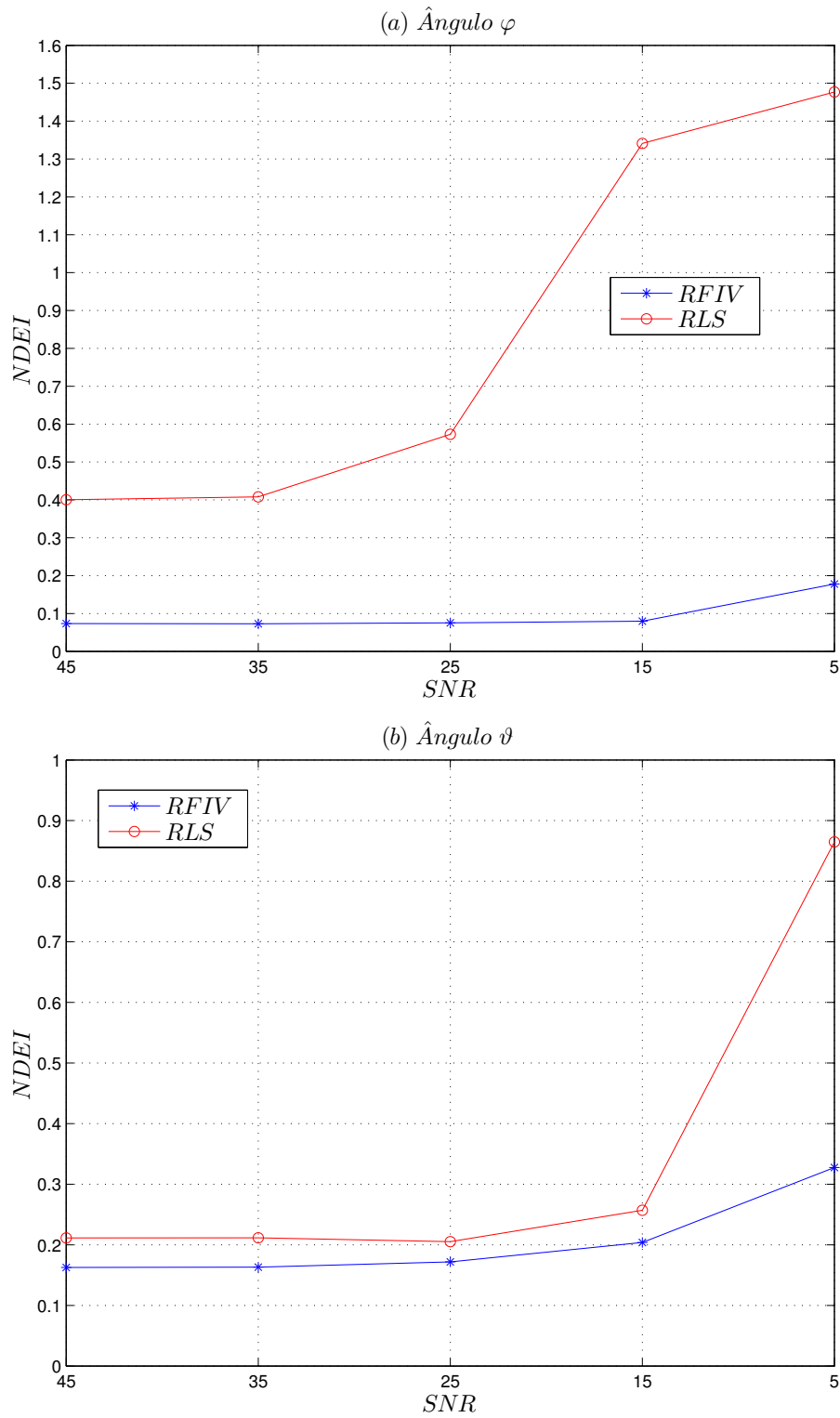
Fonte: Elaborado pelo autor.

Figura 93: Análise de robustez usada no RMSE (a) Ângulo de Azimute e (b) Ângulo de Elevação.



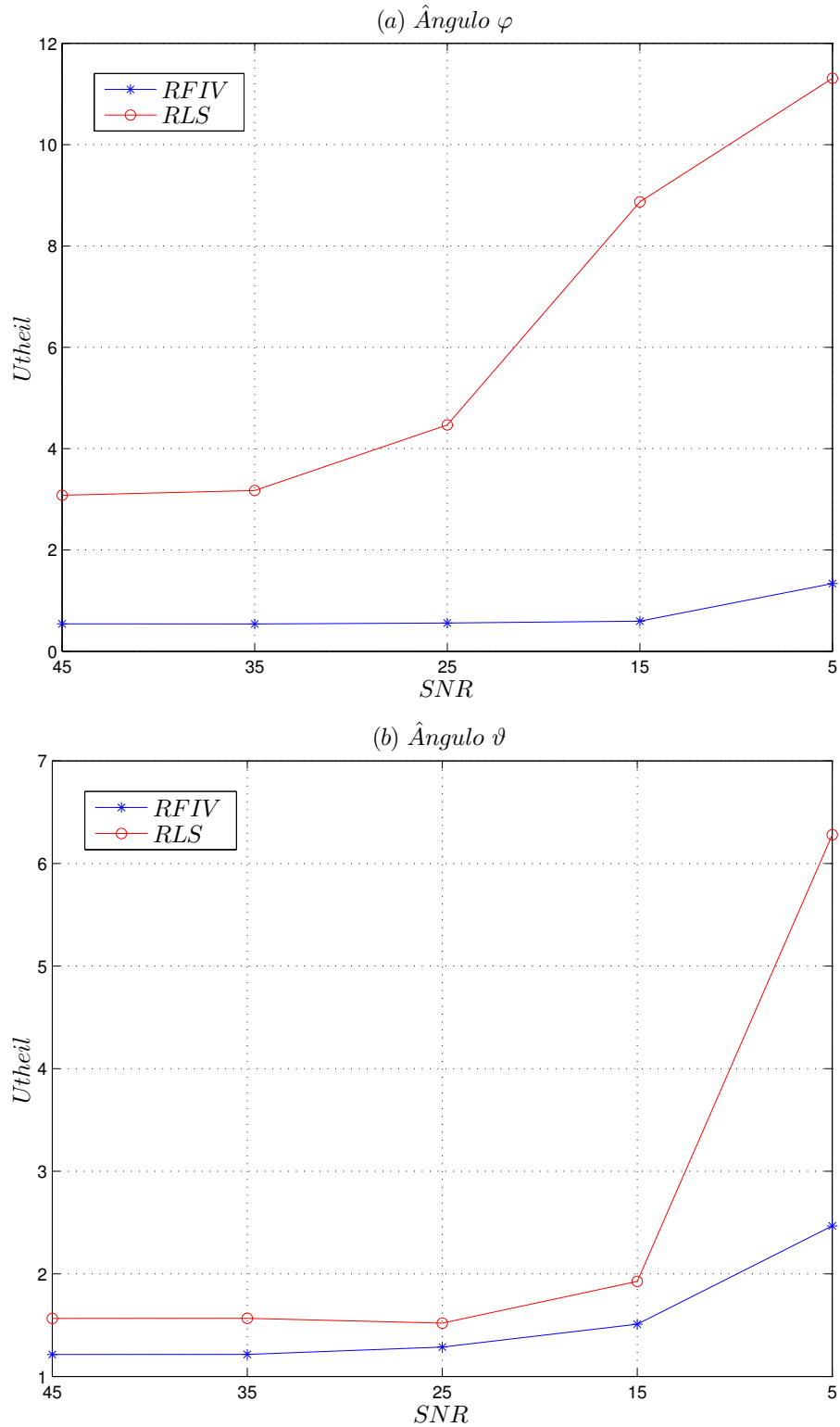
Fonte: Elaborado pelo autor.

Figura 94: Análise de robustez usada no NDEI (a) Ângulo de Azimute e (b) Ângulo de Elevação.



Fonte: Elaborado pelo autor.

Figura 95: Análise de robustez usada no UTHEIL (a) Ângulo de Azimute e (b) Ângulo de Elevação.

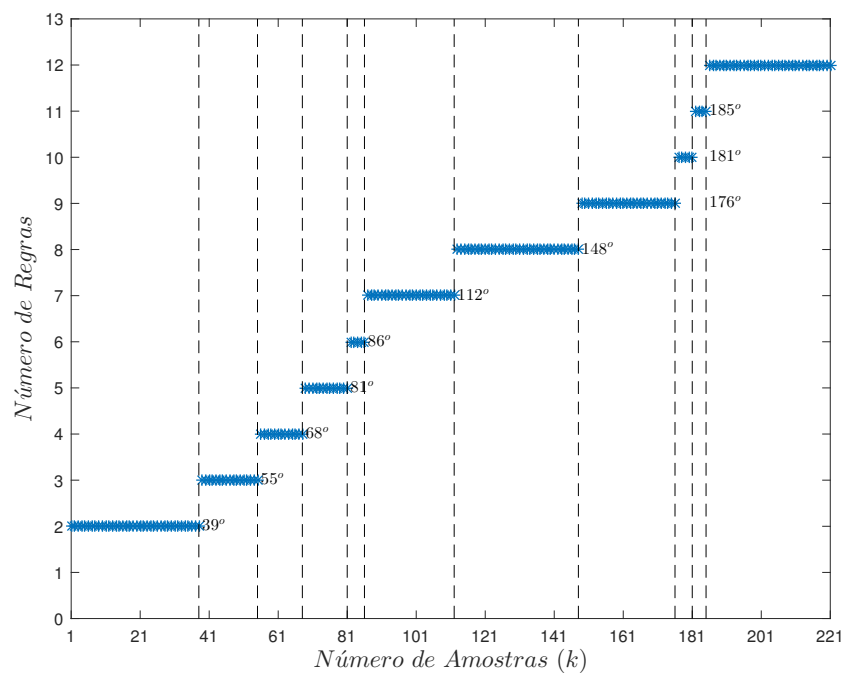


Fonte: Elaborado pelo autor.

A evolução do número de regras do modelo proposto, é ilustrado na Figura 96 e Tabela 8. O modelo finaliza com 12 regras após a aquisição da última amostra na etapa evolutiva.

O resultado médio da convergência dos parâmetros do consequente do modelo evolutivo para os ângulos de Azimute e Elevação, obtidos a partir das 1000 realizações do experimento, baseados na variável instrumental nebulosa inserida no contexto de máxima verossimilhança, com relação as 12 regras nebulosas, são ilustrados nas Figuras 97–104. Neste experimento verificou-se que o procedimento de agregação de regras não atuou, em virtude dos critérios de agregação de regras não serem satisfeitos, isto é, os grupos identificados não apresentam características de similaridades próximas que permitam que sejam agregados entre si.

Figura 96: Evolução do número de regras do modelo em função do número de amostras.



Fonte: Elaborado pelo autor.

Figura 97: Convergência dos parâmetros do consequente $\theta_{\varphi_1}^{(i)}$ e $\theta_{\vartheta_1}^{(i)}$ para $regras^{(i)}, i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.

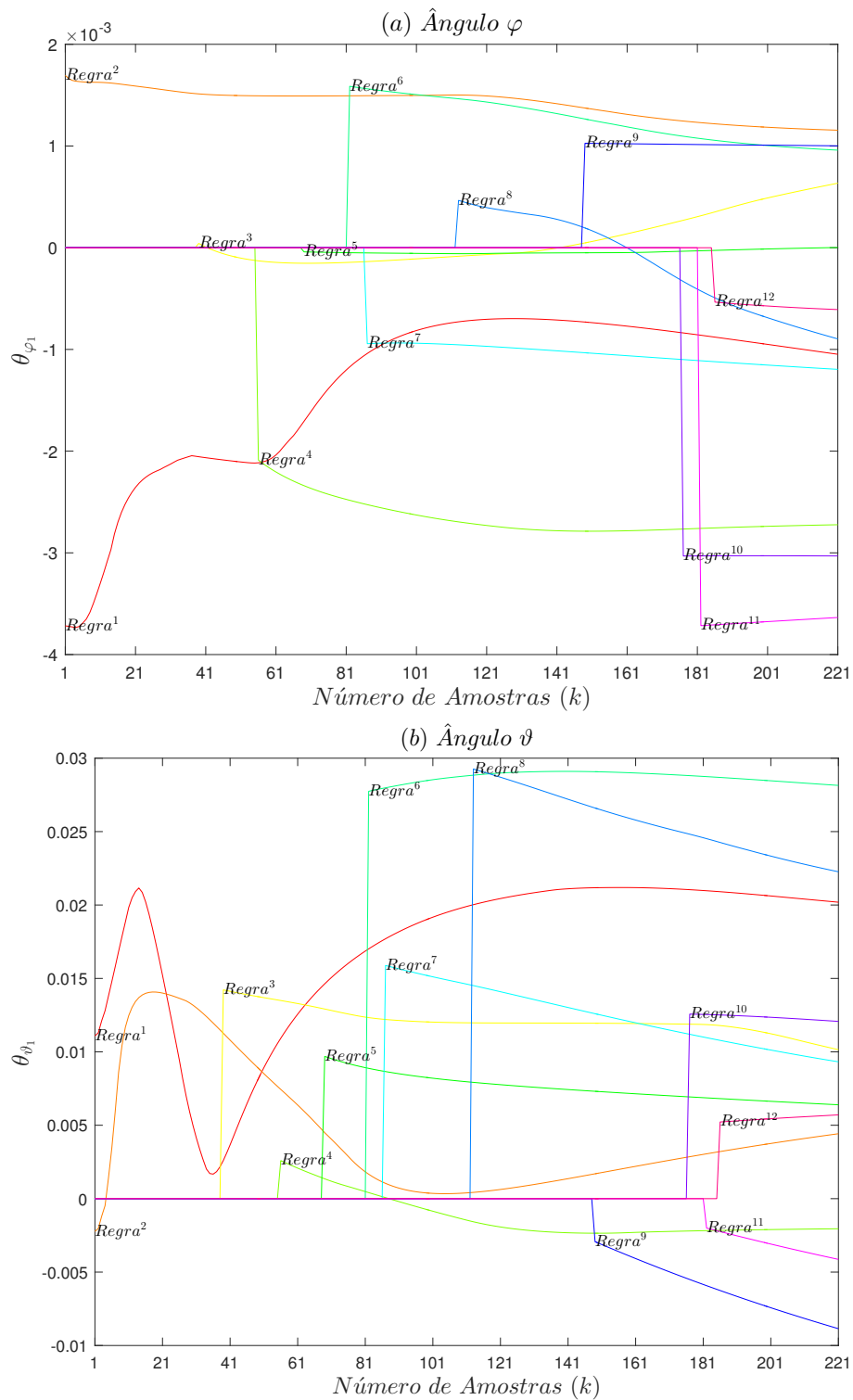


Figura 98: Convergência dos parâmetros do consequente $\theta_{\varphi_2}^{(i)}$ e $\theta_{\vartheta_2}^{(i)}$ para $regras^{(i)}, i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.

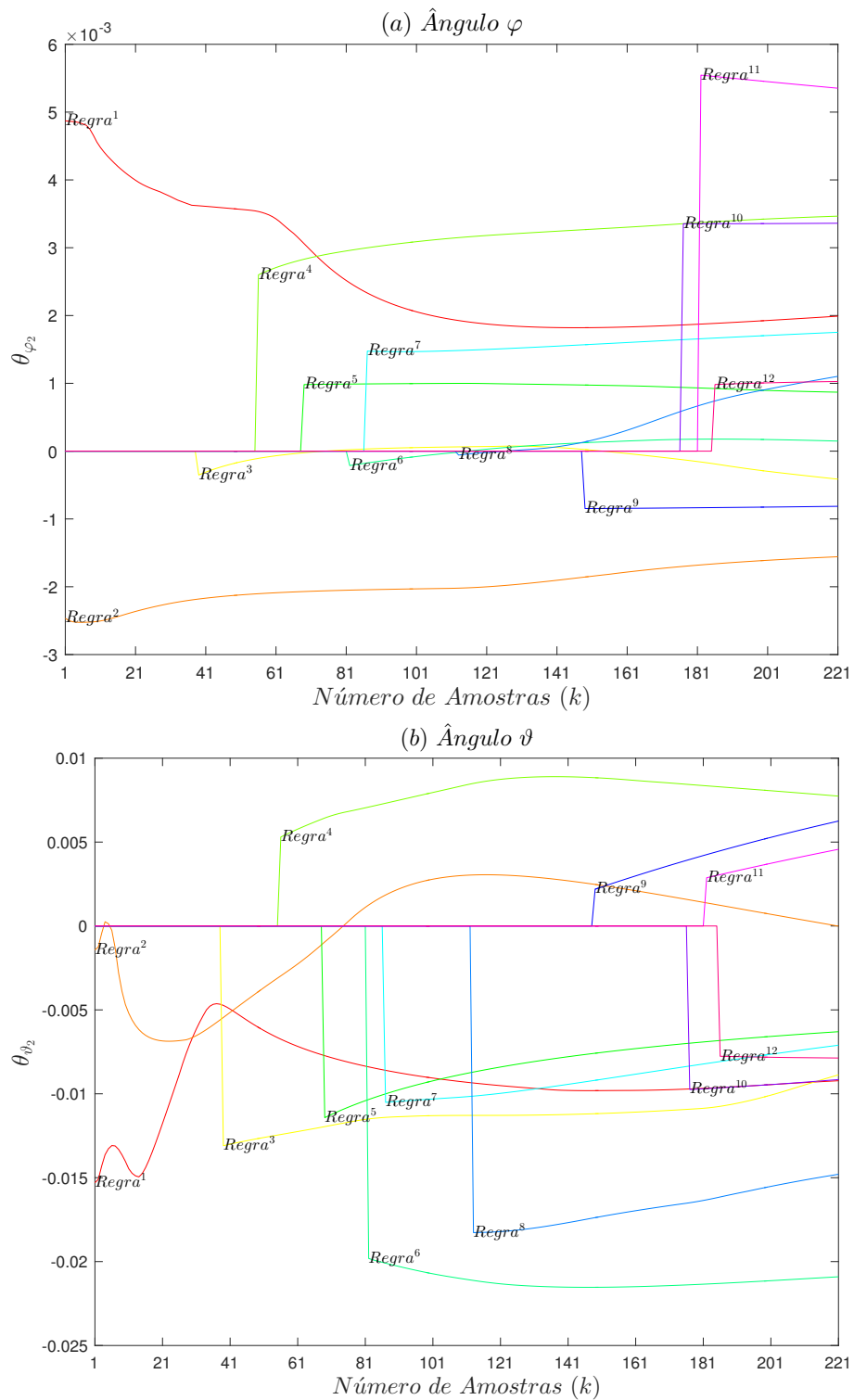


Figura 99: Convergência dos parâmetros do consequente $\theta_{\varphi_3}^{(i)}$ e $\theta_{\vartheta_3}^{(i)}$ para $regras^{(i)}, i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.

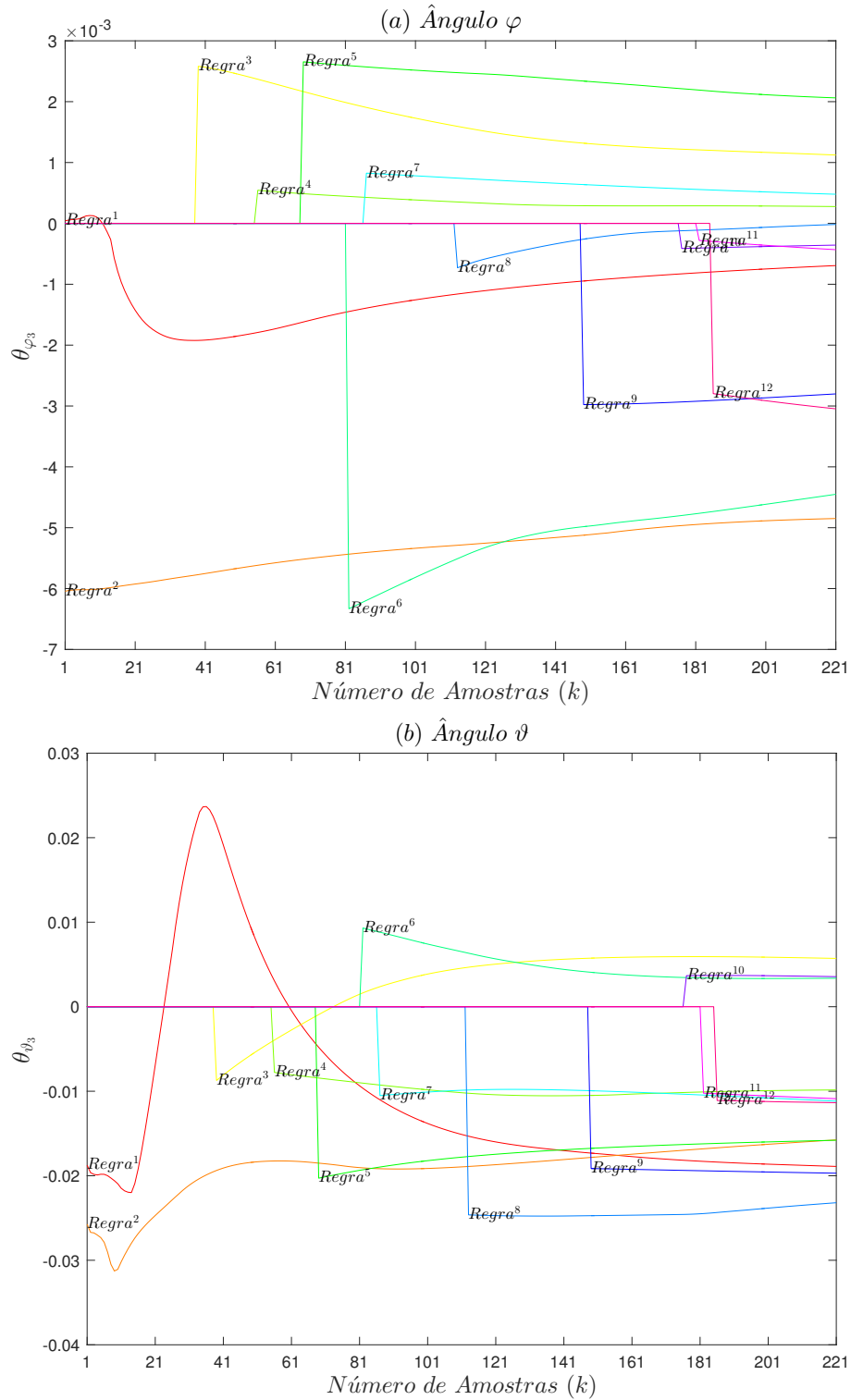


Figura 100: Convergência dos parâmetros do consequente $\theta_{\varphi_4}^{(i)}$ e $\theta_{\vartheta_4}^{(i)}$ para $regras^{(i)}, i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.

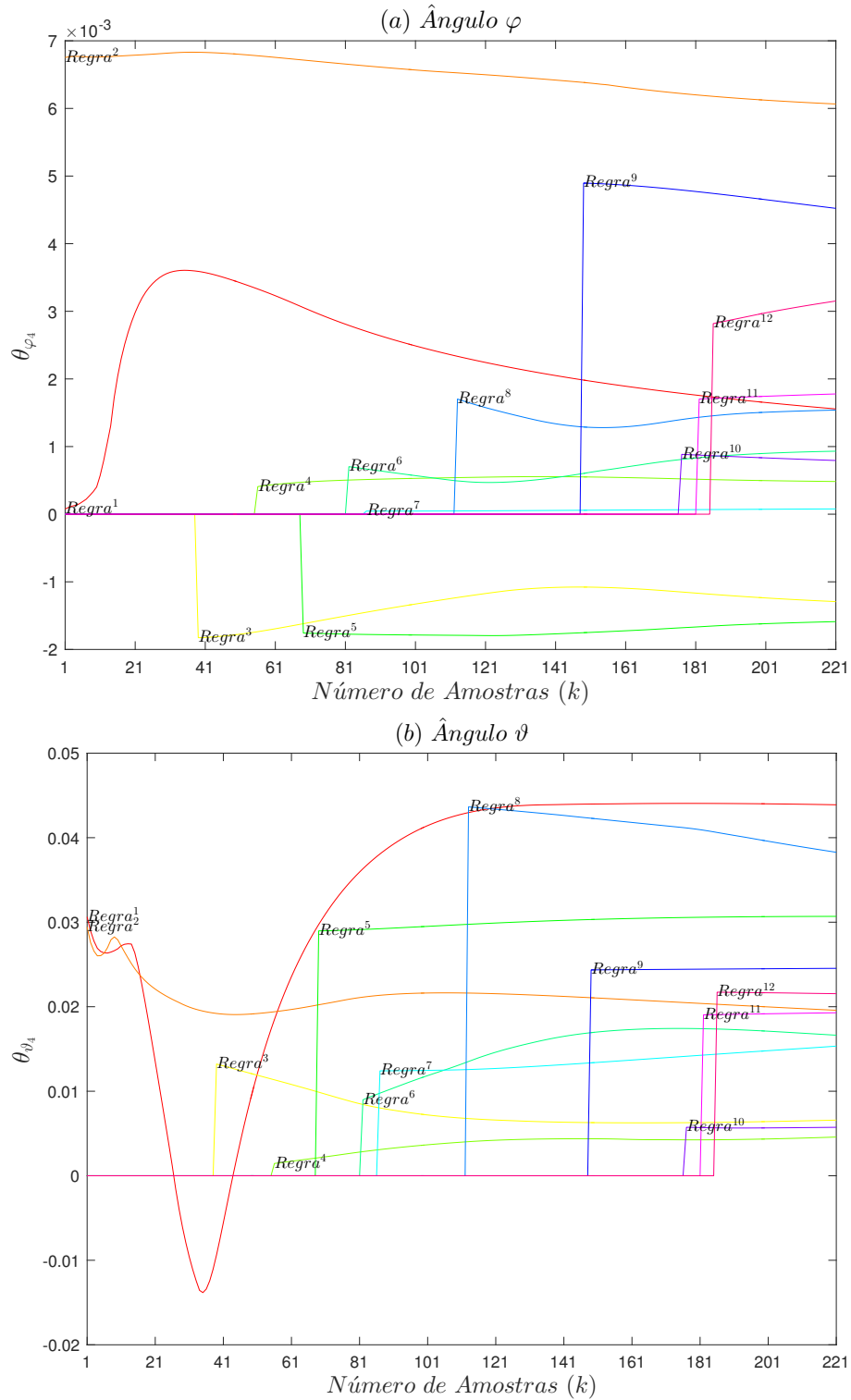
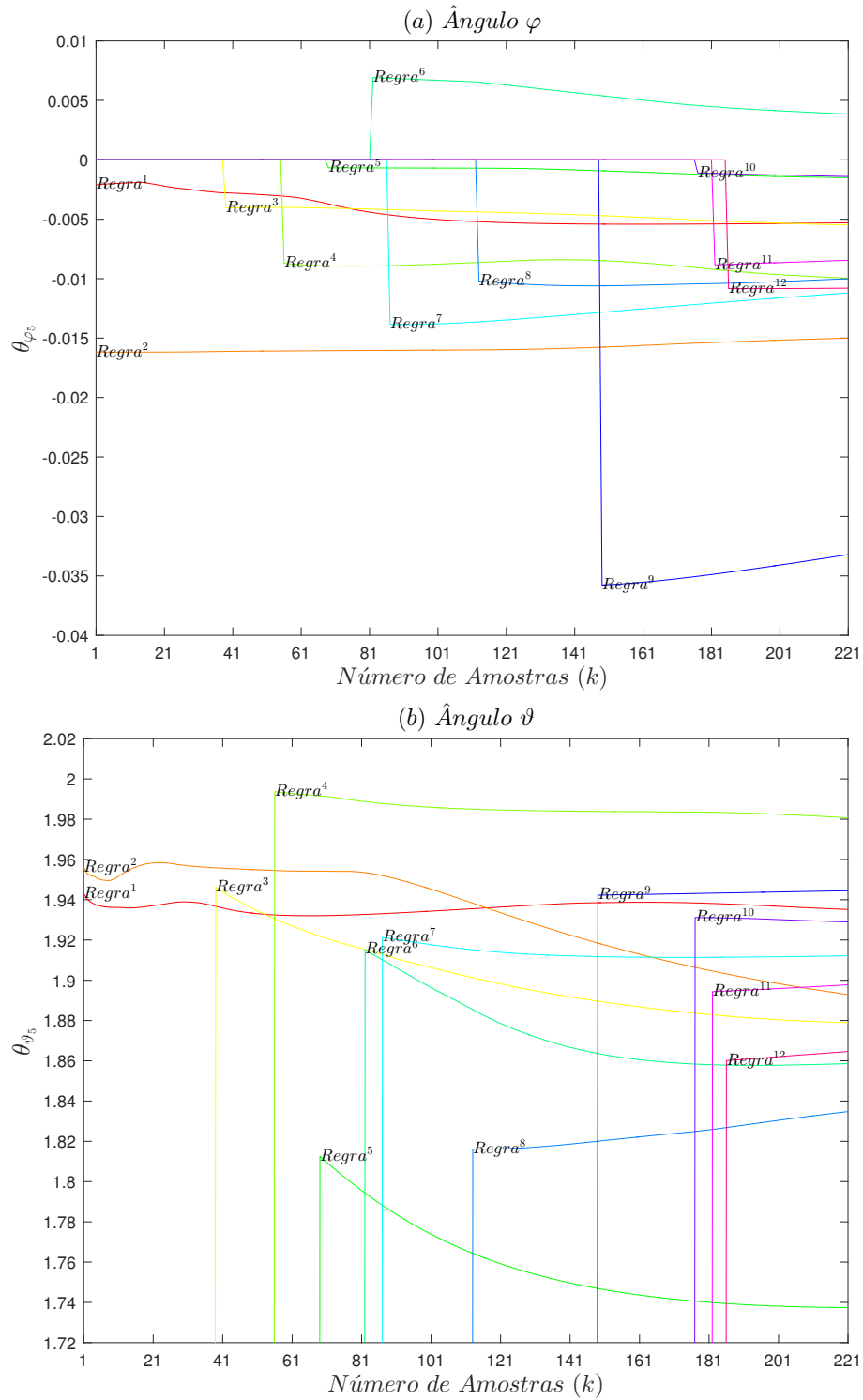
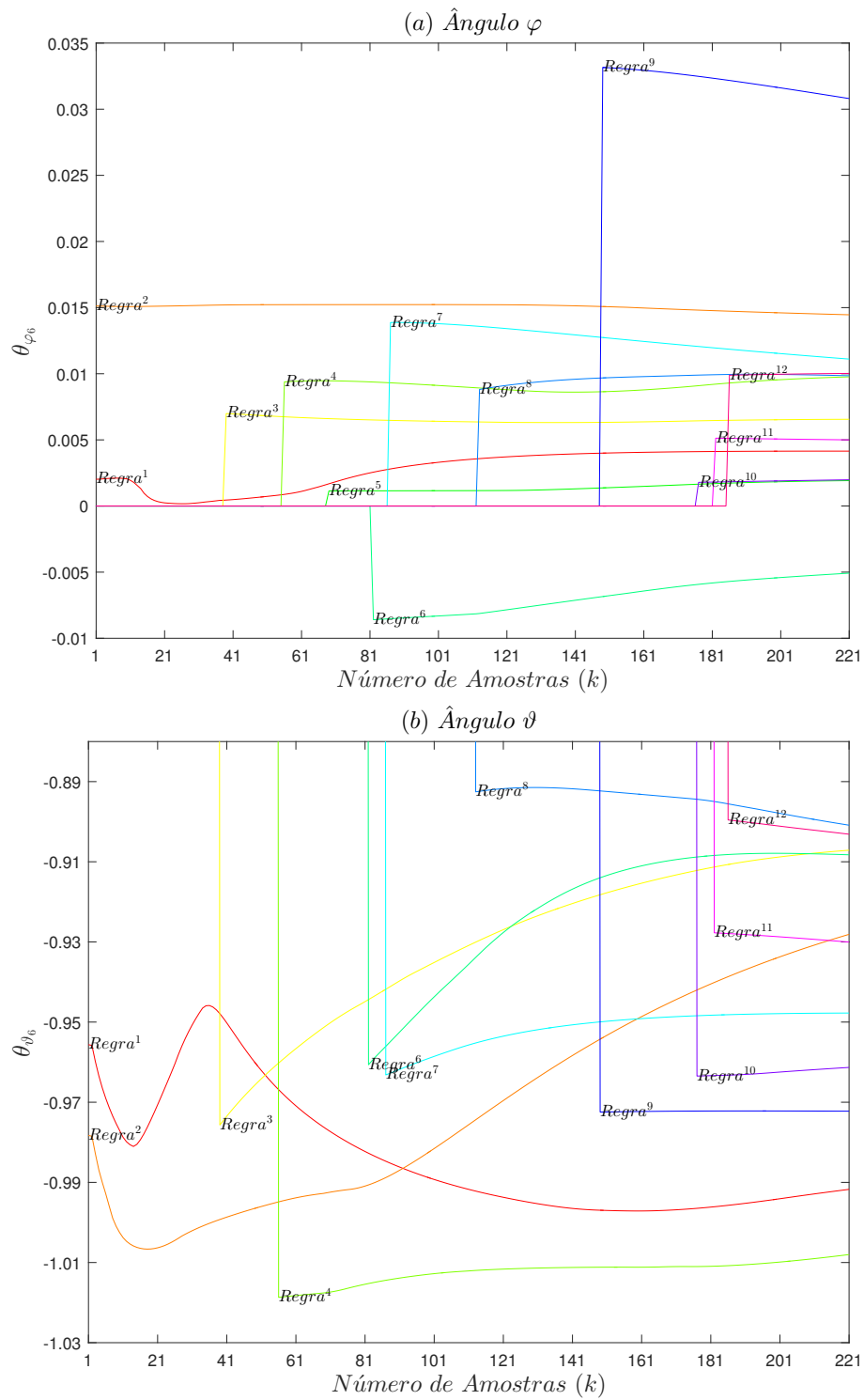


Figura 101: Convergência dos parâmetros do consequente $\theta_{\varphi_5}^{(i)}$ e $\theta_{\vartheta_5}^{(i)}$ para *regras*⁽ⁱ⁾, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.



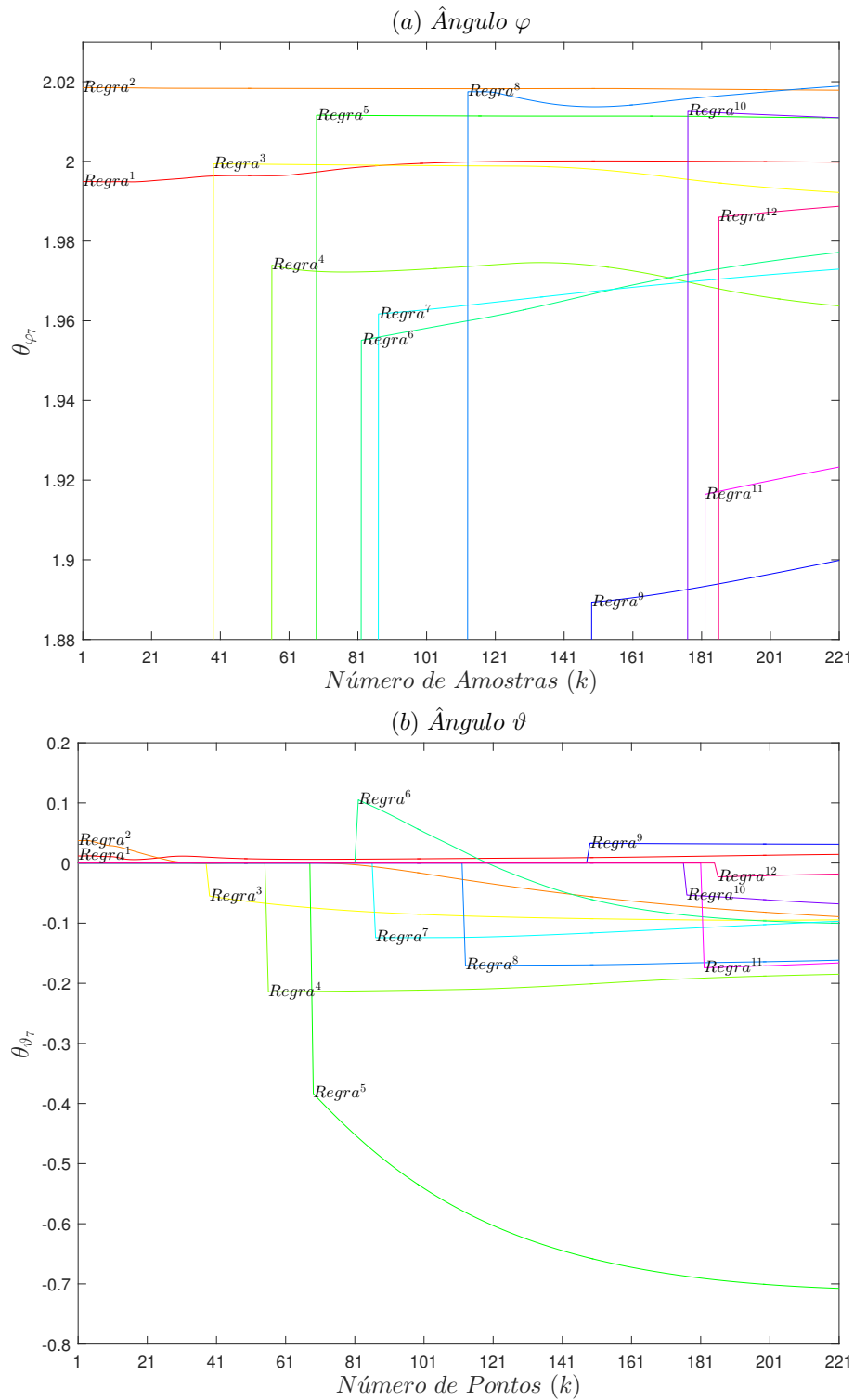
Fonte: Elaborado pelo autor.

Figura 102: Convergência dos parâmetros do consequente $\theta_{\varphi_6}^{(i)}$ e $\theta_{\vartheta_6}^{(i)}$ para $regras^{(i)}$, $i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.



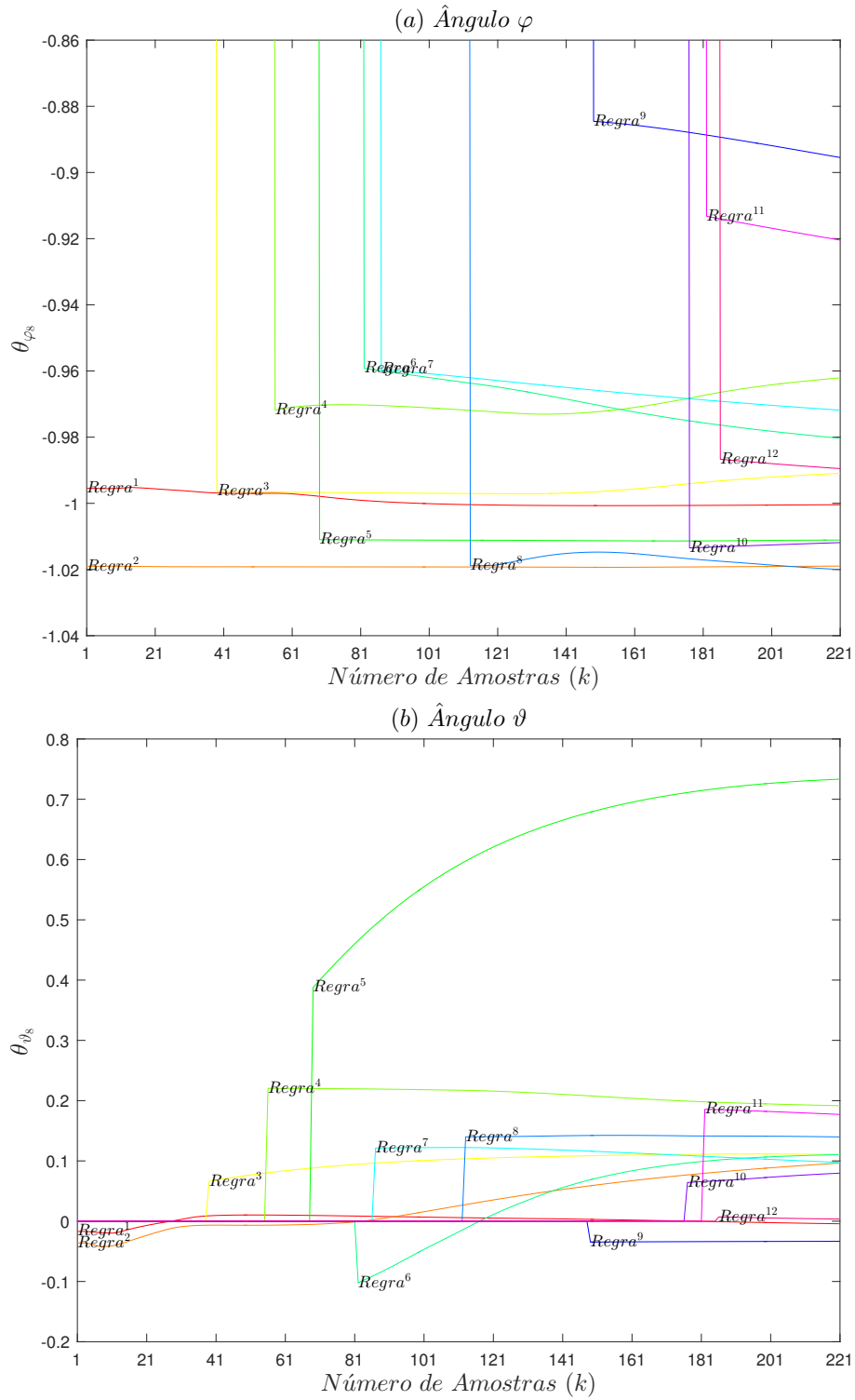
Fonte: Elaborado pelo autor.

Figura 103: Convergência dos parâmetros do consequente $\theta_{\varphi_7}^{(i)}$ e $\theta_{\vartheta_7}^{(i)}$ para $regras^{(i)}, i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.



Fonte: Elaborado pelo autor.

Figura 104: Convergência dos parâmetros do consequente $\theta_{\varphi_s}^{(i)}$ e $\theta_{\vartheta_s}^{(i)}$ para $regras^{(i)}, i = 1, 2, \dots, 12$, (a) Ângulo de Azimute e (b) Ângulo de Elevação.



Fonte: Elaborado pelo autor.

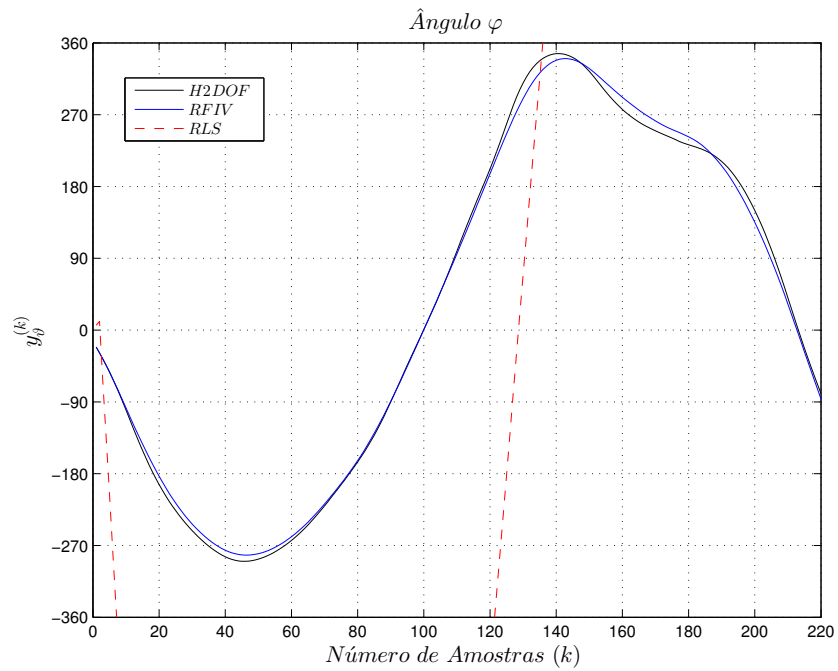
Tabela 8: Evolução do número de regras em função do número de amostras.

Número da Amostra	Número de Regras
38	3
55	4
68	5
81	6
86	7
112	8
148	9
176	10
181	11
185	12

Fonte: Elaborado pelo autor.

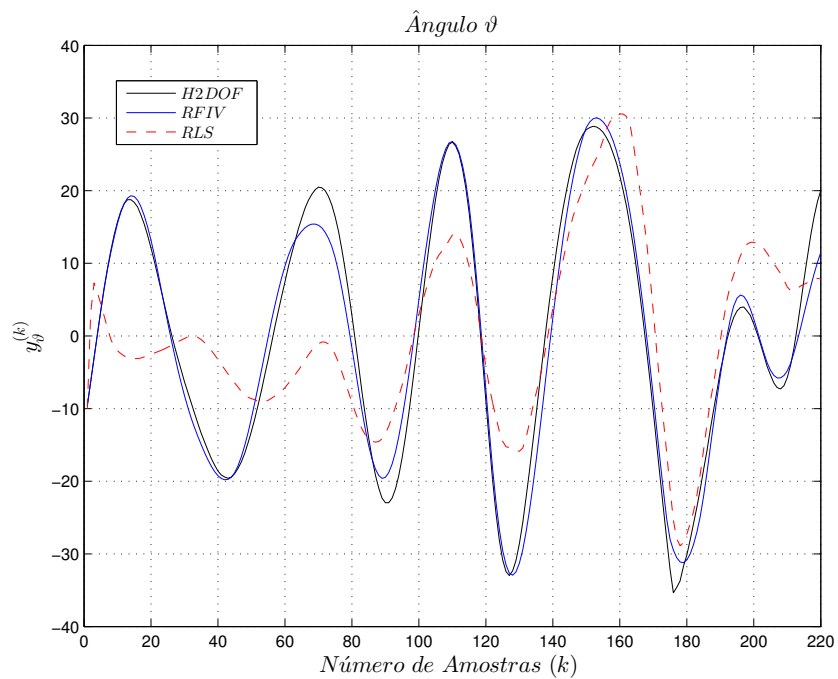
O modelo validado foi avaliado a partir de um conjunto novo de dados experimentais. O objetivo era comparar a metodologia proposta e o método dos mínimos quadrados recursivo inserido no modelo nebuloso evolutivo, a fim de garantir estruturas idênticas no antecedente de cada um dos modelos. Para decidir quão satisfatório foi o modelo proposto utilizou-se as métricas de avaliação na situação de pior caso, ou seja, um sistema corrompido com um ruído autocorrelacionado com uma relação sinal/ruído de 5 dB. A análise de robustez do modelo validado para os Ângulos de Azimute e Elevação, é apresentado na Figura 105 e na Tabela 9. Utilizou-se os dados experimentais do helicóptero 2-DOF (**H2DOF**), o resultado do modelo proposto utilizando predição livre (**RFIV**) e o modelo utilizando o método dos mínimos quadrados recursivo (**RLS**). Perceba-se, que a metodologia proposta apresenta melhor desempenho em relação à metodologia comparada relativas aos sinais de saída do ângulo de azimute e ângulo de elevação quando comparado ao método dos mínimos quadrados recursivo inserido na estrutura do modelo nebuloso evolutivo.

Figura 105: Validação do modelo para o Ângulo de Azimute φ com SNR de 5 dB.



Fonte: Elaborado pelo autor.

Figura 106: Validação do modelo para o Ângulo de Elevação ϑ com SNR de 5 dB.



Fonte: Elaborado pelo autor.

Tabela 9: Análise Comparativa das Métricas VAF, NDEI, RMSE e UTHEIL com SNR de 5 dB.

Métricas	Ângulo de Azimute φ		Ângulo de Elevação ϑ	
	RFIV	RLS	RFIV	RLS
VAF	95,645	19,134	88,488	26,525
NDEI	0,178	1,477	0,327	0,865
RMSE	0,234	10,119	0,290	0,890
UTHEIL	1,338	11,310	2,467	6,279

Este capítulo apresentou os resultados computacionais e experimentais que foram utilizados para a avaliação de desempenho da metodologia proposta, a partir da comparação com outras metodologias amplamente apresentadas na literatura e a exemplificação de uma modelagem caixa preta inserida em um ambiente cujos conjuntos de dados foram comprometidos por ruído autocorrelacionado.

6 CONCLUSÕES E PROPOSTAS FUTURAS

6.1 Conclusões

Há várias metodologias para identificação de sistemas dinâmicos não-lineares usando modelos evolutivos e destacam-se as estruturas baseadas em métricas de distâncias não adaptativas que utilizam a distância *Euclidiana*, métricas de distâncias adaptativas com uso de *Mahalanobis*, além das novas estruturas granulares evolutivas. Também, verifica-se que na literatura há vários modelos de estimadores paramétricos utilizados na estimação paramétrica de agrupamentos nebulosos evolutivos, porém, percebe-se que muitos deles lidam com estimadores polarizados. A natureza de um modelo nebuloso evolutivo baseado em métricas de distância adaptativas que se utiliza da máxima verossimilhança, aliado ao uso da variável instrumental nebulosa inserida neste mesmo contexto foi apresentado ao longo do trabalho, com os principais resultados obtidos para fins de validação da metodologia proposta.

Um algoritmo de agrupamento nebuloso evolutivo recursivo baseado na estimação de máxima verossimilhança com variável instrumental nebulosa inserida no contexto evolutivo aplicado a modelagem caixa preta de sistemas dinâmicos operando em ambiente não-estacionário, foi proposto.

O desempenho da metodologia proposta foi avaliada em simulações computacionais e identificação de sistemas dinâmicos baseados em plantas reais a serem descritos:

- a) **Análise de Convergência da Variável Instrumental:** a análise de convergência e estatística da variável instrumental nebulosa no contexto evolutivo baseada em máxima verossimilhança foi objeto de estudo desta tese. O desempenho da variável instrumental nebulosa de forma não-recursiva (batelada) e de forma recursiva foi avaliado e comparado sob as mesmas condições de testes com o desempenho do método dos mínimos quadrados para fins de análise e validação dos modelos de variável instrumental;
- b) **Sistemas dinâmico não-linear:** o desempenho final da modelagem de um sistema dinâmico não-linear de segunda ordem amplamente utilizado na literatura permitiu a comparação com os desempenhos de outras metodologias de sistemas evolutivos existentes. Além disso, permitiu a avaliar o efeito do procedimento de cálculo da probabilidade *a priori* proposta nesta tese em relação à metodologia apresentada nos estimadores de máxima verossimilhança;

- c) **Modelagem de um Sistema Térmico:** na modelagem caixa preta de um sistema térmico (planta térmica) avaliou-se o efeito da metodologia proposta para cálculo da taxa de aprendizado participativo utilizada nesta tese e comparada com metodologia de taxa de aprendizado fixa que é utilizada em outras metodologias de algoritmos evolutivos;
- d) **Modelagem de um Helicóptero 2DOF:** a modelagem caixa-preta de um sistema experimental baseado em um helicóptero com 2 graus de liberdade, cujo conjunto de dados contém a presença de ruído autocorrelacionado para avaliar o desempenho da metodologia proposta em um sistema dinâmico operando em um ambiente não-estacionário.

Nesta tese, uma estrutura autorregressiva com entrada exógena (ARX) inserida no método de inferência de *Takagi-Sugeno* é utilizada, dentre às várias estruturas existentes na literatura, pois apresenta uma formulação simples para representar o comportamento de qualquer sistema dinâmico, além de ser usada em conjunto com a metodologia de variável instrumental nebulosa no contexto evolutivo. A estrutura ARX em combinação com a inferência de *Takagi-Sugeno* possibilitou identificação de modelos não-lineares a partir de parâmetros obtidos do modelo nebuloso evolutivo baseado na métrica de máxima verossimilhança com taxa de aprendizagem participativa utilizando um fator exponencial.

Dentre as principais conclusões sobre a metodologia apresentada nesta tese, podemos destacar:

- a) a eficiência da metodologia proposta na identificação de um sistema dinâmico não-linear amplamente utilizado na literatura, foi avaliada a partir do desempenho desta em relação às outras metodologias existentes na literatura, quando comparada as métricas de avaliação: raiz do erro quadrático médio (RMSE), índice de erro não-dimensional (NDEI) e o total de regras nebulosas identificadas na iteração final do algoritmo. O que possibilitou avaliar a estratégia utilizada no procedimento adotado para o cálculo da probabilidade *a priori* associado à métrica de máxima verossimilhança utilizada no algoritmo de agrupamento nebuloso evolutivo;
- b) a metodologia proposta na identificação de um sistema dinâmico experimental, um processo térmico (planta térmica), no tempo discreto, nos permitiu avaliar o efeito da metodologia utilizada para cálculo do fator de aprendizado participativo utilizado nesta proposta em relação às metodologias disponíveis na literatura, além de avaliar o mecanismo adotado para junção de regras que apresentam similaridade;
- c) a metodologia proposta na identificação de um sistema dinâmico experimental, um helicóptero com dois graus de liberdade, no tempo discreto, corrompidos por ruído autocorrelacionado, quando utilizada em conjunto com a variável instrumental

nebulosa inserida no contexto evolutivo foi comparada com o método dos mínimos quadrados recursivos, também inseridos no mesmo contexto evolutivo. O desempenho favorável obtido nesta identificação é devido ao uso de uma variável instrumental nebulosa inserida em um contexto evolutivo, que possibilitou o ajuste do vetor de parâmetros utilizado na estimação paramétrica recursiva à medida que o modelo evolui e modifica a estrutura interna do modelo de inferência *Takagi–Sugeno*;

- d) a partir de simulações computacionais percebeu-se que, para uma melhor eficiência do modelo nebuloso evolutivo baseado em máxima verossimilhança na modelagem de sistemas dinâmicos operando em ambiente não-estacionário, foi proposto um novo procedimento para cálculo da probabilidade *a priori* da máxima verossimilhança e seu desempenho foi evidenciado nos resultados computacionais e experimentais;
- e) um problema característico de sistemas nebulosos evolutivos é o crescimento exponencial do número de regras nebulosas em função do aumento da dimensão do seu espaço de entrada. Este crescimento do número de regras é denominado de *maldição da dimensionalidade* e é um dos principais problemas envolvidos na identificação de sistemas dinâmicos com um espaço de entrada relativamente grande. Adotou-se uma estratégia de verificação e busca de regras similares a cada nova iteração; como consequência da estratégia adotada para mesclar as regras similares é a obtenção do menor número destas e garantiu uma melhor transparência e interpretabilidade do modelo nebuloso evolutivo;
- f) a análise de convergência da variável instrumental nebulosa inserida no contexto evolutivo, utilizada na estimação paramétrica do consequente do modelo *Takagi–Sugeno*, demonstrou uma garantia de robustez na estimação paramétrica de sistemas dinâmicos na presença de ruído autocorrelacionado. O refinamento desta variável instrumental diante do uso de etapas de pré-filtragem demonstraram eficiência da metodologia quando utilizada em sistemas dinâmicos não-lineares amplamente divulgados na literatura.

6.2 Propostas Futuras

Como propostas de continuidade para este trabalho sugere-se:

- a) Analisar os parâmetros iniciais da estimação inicial e evolutiva do algoritmo proposto. Estes são escolhidos por um especialista e propor metodologias para ajuste automático destes parâmetros, proporcionando uma maior autonomia à metodologia proposta nesta tese;
- b) Desenvolver o uso de uma variável instrumental nebulosa inserida no contexto evolutivo baseada na metodologia *Bayesiana*, voltada para identificação de sistemas

dinâmicos não-lineares no tempo discreto e em ambiente com ruído autocorrelacionado.

- c) Implementar algoritmos nebulosos evolutivos que sejam baseados em modelos probabilísticos *Bayesianos*. Portanto, a apresentação de uma nova estratégia para modelos nebulosos evolutivos, para identificação de sistemas dinâmicos não-estacionários, é de grande importância.
- d) Reavaliar o método proposto para ser aplicado a controladores nebulosos inserido no contexto evolutivo;
- e) Analisar o procedimento de junção de regras a partir da comparação dos parâmetros do consequente obtidos durante a estimação evolutiva.

Apêndices

APÊNDICE A – PROCESSOS ESTOCÁSTICOS

Neste apêndice objetiva-se fornecer alguns conceitos básicos sobre processos estocásticos e estacionariedade. Estes conceitos permitem entender os estudos desenvolvidos nesta tese e para um estudo mais completo, as seguintes referências são recomendadas: (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008).

A.1 Definição

Seja uma variável aleatória como sendo uma função real z que associa cada amostra ω , pertencentes ao espaço de amostras Ω um número $z(\omega) \in \mathbb{R}$. Seja um vetor aleatório (ou variável n -dimensional) definido por uma função vetorial z que associa a cada amostra $\omega \in \Omega$ um vetor n -dimensional $\mathbf{z} \in \mathbb{R}^n$ (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008).

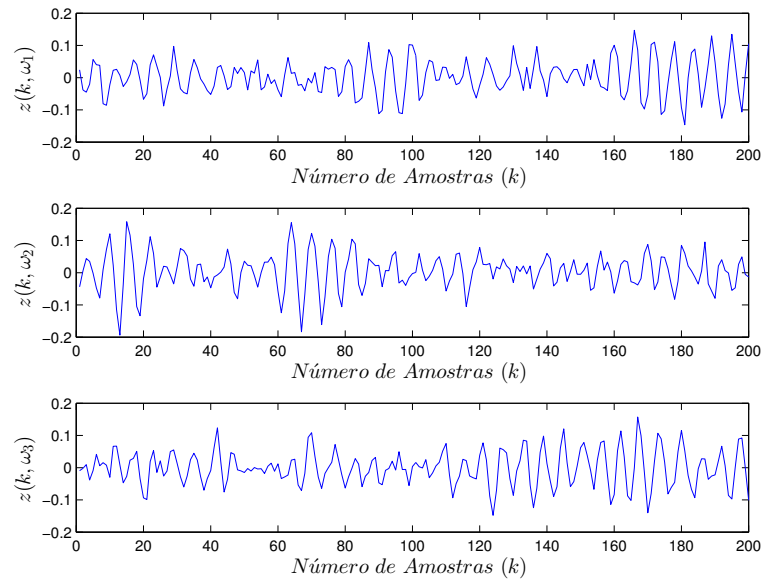
Considere então um mapa que associa a cada amostra $\omega \in \Omega$ uma função real de um parâmetro k pertencente a um conjunto N (na maioria dos processos estocásticos, o parâmetro k está associado a um instante de tempo). Cria-se desta maneira uma família \mathbb{F} de funções de k , ($k \in N$). A este mapeamento denomina-se de processo estocástico ou processo aleatório (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008).

Deste modo, pode-se dizer que um processo estocástico é o mapa definido por

$$\begin{aligned} z : \quad \Omega &\longrightarrow \mathbb{F} \\ \omega &\longrightarrow z(k, \omega), k \in N \end{aligned} \tag{A.1}$$

Baseado nestas condições, pode-se dizer que um processo estocástico é um conjunto de variáveis aleatórias, que supõem-se definidas num espaço de probabilidades. Ou seja, um processo estocástico é uma função de dois argumentos $z(k, \omega_i)$, onde $\omega_i, i = 1, 2, \dots$ representa o valor de uma i ésima variável aleatória em cada instante $k = 1, 2, \dots N$, ilustrado na Figura 107 que representa um processo estocástico produzido pelo mapeamento de três parâmetros aleatórios em funções dos instantes k , observa-se que para um instante k tem-se distintos valores para z .

Figura 107: Processo estocástico produzido pelo mapeamento das amostras em funções de um parâmetro k .



Fonte: Elaborada pelo autor

A.2 Classificação de Processos Estocásticos

Os processos estocásticos podem ser classificados tanto de acordo com os valores que ele assume quanto de acordo com os valores que seu parâmetro pode assumir. Assim, um processo estocástico que toma valores em conjunto discreto de amostras no eixo real é dito um processo estocástico discreto. Por outro lado, se o processo toma valores em uniões de subconjuntos contínuos de \mathbb{R} , ele é considerado um processo estocástico contínuo. De maneira análogo, processos cujo parâmetro toma valores em um subconjunto discreto dos reais são ditos processos estocásticos de parâmetro discreto enquanto aqueles cujo parâmetro toma valores em uniões de subconjuntos contínuos de \mathbb{R} são ditos processos estocásticos de parâmetro contínuo (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008). Esta tese baseia-se em processos estocásticos de parâmetros discretos.

A.3 Especificação de Processos Estocásticos

Seja um processo estocásticos $z(k)$, cujo o valor do parâmetro k está fixo, então, obtém-se a variável aleatória z_k . Associada a esta variável aleatória tem-se, uma função de distribuição de probabilidade $F_{z(k)}(\mathbf{Z})$ e, conseqüentemente, uma função de densidade de probabilidade $p_{z(k)}(\mathbf{Z})$ (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008). Note que para valores distintos de variáveis aleatórias, tem-se as seguintes definições:

Definição A.3.1 (Especificação de 1ª Ordem de um Processo Estocástico) : diz-se que um processo estocástico está especificado até a primeira ordem quando a função densidade de probabilidade $p_{z^{(k)}}(\mathbf{Z})$ é conhecida para qualquer valor de k . As funções densidade de probabilidade são denominadas funções densidades de probabilidade de primeira ordem do processo estocástico $z(k)$.

Definição A.3.2 (Especificação de 2ª Ordem de um Processo Estocástico): diz-se que um processo estocástico está especificado até a segunda ordem quando a função de densidade de probabilidade conjunta $p_{z_1^{(k)}, z_2^{(k)}}(\mathbf{Z}_1, \mathbf{Z}_2)$ é conhecida. A função densidade de probabilidade conjunta é denominada função densidade de probabilidade de segunda ordem do processo estocástico $z(k)$.

Definição A.3.3 (Especificação de Ordem m de um Processo Estocástico): diz-se que um processo estocástico está especificado até a ordem m quando a função de probabilidade conjunta $p_{z_1^{(k)}, z_2^{(k)}, \dots, z_m^{(k)}}(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_m)$ é conhecida. A função densidade de probabilidade conjunta é denominada função densidade de probabilidade de ordem m do processo estocástico $z(k)$.

A.4 Momentos de Processos Estocásticos

Os momentos de um processo estocástico são os momentos de variáveis aleatórias definidas em quaisquer instantes do processo (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008). Alguns desses momentos, de maior importância, são definidos a seguir.

Definição A.4.1 (Média de um processo estocástico): a média de um processo estocástico $\mathbf{Z} = \{z^{(k)}, k\}$, representada por $\mu_f = E_{\mathbf{Z}} [z^{(k)}]$, é definida como a média da variável aleatória $\{z^{(k)}\}$ associada a um instante k , ou seja,

$$\mu_f = E_{\mathbf{Z}} [z^{(k)}], \quad k \in N \quad (\text{A.2})$$

Definição A.4.2 (Autocorrelação de um processo estocástico): a função de autocorrelação de um processo estocástico $z(k)$, representada por $R_f(\mathbf{Z}_1, \mathbf{Z}_2) = E_{\mathbf{Z}} [\mathbf{Z}_1 \mathbf{Z}_2]$, é definida como correlação entre as variáveis aleatórias \mathbf{Z}_1 e \mathbf{Z}_2 , associada a quaisquer valores de $k = 1, 2, \dots, N$;

$$R_f(\mathbf{Z}_1, \mathbf{Z}_2) = E_{\mathbf{Z}} [\mathbf{Z}_1 \mathbf{Z}_2], \quad k \in N \quad (\text{A.3})$$

Definição A.4.3 (Autocovariância de um processo estocástico): a função autocovariância de um processo estocástico $z(k)$, representada por $C_f(\mathbf{Z}_1, \mathbf{Z}_2) = E_{\mathbf{Z}} [(\mathbf{Z}_1 - \mu_{z_1})(\mathbf{Z}_2 - \mu_{z_2})]$, é definida como a covariância entre as variáveis aleatórias \mathbf{Z}_1 e \mathbf{Z}_2 , associada a quaisquer valores de $k = 1, 2, \dots, N$;

$$C_f(\mathbf{Z}_1, \mathbf{Z}_2) = E_{\mathbf{Z}} [(\mathbf{Z}_1 - \mu_{z_1})(\mathbf{Z}_2 - \mu_{z_2})], \quad k \in N \quad (\text{A.4})$$

A.5 Estacionariedade de Processos Estocásticos

Um processo estocástico pode ser estacionário em diversos graus de estacionariedade. Os conceitos de estacionariedade de ordem m , estacionariedade estrita e estacionariedade no sentido amplo são definidos a seguir (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008).

Definição A.5.1 (Estacionariedade de um processo estocástico de ordem m): um processo estocástico $z(k)$ é dito estacionário de ordem m quando a função densidade de probabilidade de ordem m não se altera com o deslocamento de tempo, um processo estocástico é dito estacionário de ordem m , como segue,

$$p_{z_1^{(k)}, z_2^{(k)}, \dots, z_m^{(k)}}(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_m) = p_{z_1^{(k+\tau)}, z_2^{(k+\tau)}, \dots, z_m^{(k+\tau)}}(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_m) \quad \forall \tau \quad (\text{A.5})$$

Definição A.5.2 (Estacionariedade de um processo estocástico no sentido estrito): um processo estocástico $z(k)$ é dito estacionário no sentido estrito, quando suas estatísticas são invariantes no tempo, ou seja,

$$E_{\mathbf{Z}} [z^{(k)}] = E_{\mathbf{Z}} [z^{(k+\tau)}] \quad \forall \tau \quad (\text{A.6})$$

$$R_f(z_1^{(k)}, z_2^{(k)}) = E[z_1^{(k)} z_2^{(k)}] = E[z_1^{(k+\tau)} z_2^{(k+\tau)}] = R_f(z_1^{(k+\tau)}, z_2^{(k+\tau)}) \quad \forall \tau \quad (\text{A.7})$$

$$E_{\mathbf{Z}} [(z_1^{(k)} - \mu_{z_1})(z_2^{(k)} - \mu_{z_2})] = E_{\mathbf{Z}} [(z_1^{(k+\tau)} - \mu_{z_1})(z_2^{(k+\tau)} - \mu_{z_2})] \quad \forall \tau \quad (\text{A.8})$$

Definição A.5.3 (Estacionariedade de um processo estocástico no sentido amplo): um processo estocástico $z(k)$ é dito ser estacionário no sentido amplo, quando sua média é constante, mas a sua autocorrelação dependente do intervalo de deslocamento entre as variáveis aleatórias (A.PAPOULIS; PILAI, 2002; ALBUQUERQUE J. M. P. FORTES, 2008).

No contexto de processos estocásticos e processos estacionários, seja o seguinte sistema dinâmico:

$$\mathbf{x}^{(k+1)} = f(\mathbf{x}^{(k)}, \mathbf{u}^{(k)}) + erro^{(k)} \quad k = 0, 1, 2, \dots \quad (\text{A.9})$$

onde $\mathbf{x}^{(k)}$ e $\mathbf{u}^{(k)}$ representam o estado e entrada do sistema, respectivamente, e $erro^{(k)}$ é o erro estocástico, no instante k .

A.6 Sistema Dinâmico Estocástico

Há sistemas cujo comportamento é determinístico, ou seja, os argumentos e parâmetros estruturais da função $f(\bullet)$ (equação (A.9)) são suficientes para se determinar

o estado \mathbf{x} . No entanto, quando os argumentos e parâmetros estruturais da função $f(\bullet)$ (equação (A.9)) são estabelecidos por comportamentos estocásticos cujas propriedades são regidas por distribuições aleatórias em função do tempo, a dinâmica do sistema torna-se estocásticas. Na literatura, estudos de sistemas dinâmicos estocásticos são bastante abordos em teoria de filtros de *Kalman* e sistemas dinâmicos baseados em inferências (MORETTIN, 2006).

Seja a média do sistema definido pela equação (A.9), expresso por

$$E_{\mathbf{X}}(\mathbf{x}^{(k+1)}) = E_{\mathbf{X}}(f(\mathbf{x}^{(k)}, \mathbf{u}^{(k)}) + erro^{(k)}) = \mu_f \quad (\text{A.10})$$

Seja a variância do sistema definido pela equação (A.9), expresso por

$$Var_{\mathbf{X}}(\mathbf{x}^{(k+1)}) = Var_{\mathbf{X}}(f(\mathbf{x}^{(k)}, \mathbf{u}^{(k)}) + erro^{(k)}) = \sigma_f^2 \quad (\text{A.11})$$

Definição A.6.1: um sistema dinâmico estocástico é dito estacionário quando apresenta as seguintes características em seus momentos:

- Média constante do conjunto de dados;
- Variância constante do conjunto de dados;
- Covariância independe do instante k .

A.7 Sistemas Dinâmicos Não-Estacionários

Definição A.7.1: um conjunto de dados é dito não-estacionário quando a distribuição conjunta de quaisquer variáveis variam e se modificam a cada instante k (MORETTIN, 2006).

Definição A.7.2: um conjunto de dados não-estacionários apresentam médias e variâncias distintas em intervalos de tempo k , isto é, a média e/ou variância são variantes no instante k .

Os parâmetros estruturais da função $f(\bullet)$ da equação (A.9) variam a cada instante k , isto é, a não-estacionariedade de um sistema dinâmico será refletida nos parâmetros variantes que compõe a matriz de estados do sistema dinâmico, portanto, é dito ser não-estacionário quando

$$E_{\mathbf{X}}(\mathbf{x}^{(k+1)}) = \mu_f^{(k+1)} \neq constante \quad k = 1, 2, \dots \quad (\text{A.12})$$

e

$$Var_{\mathbf{X}}(\mathbf{x}^{(k+1)}) = \sigma_f^{(k+1)} \neq constante \text{ e incrementada com } k \quad k = 1, 2, \dots \quad (\text{A.13})$$

onde $E_{\mathbf{X}}(\bullet)$ e $Var_{\mathbf{X},k}(\bullet)$ são a média e variância do sistema dinâmico da equação (A.9), respectivamente.

A maioria das séries de dados são do tipo não-estacionárias, apresentam uma raiz unitária, a exemplo a séries de caminhos aleatórios (do inglês, *random walks*), que apresenta amostras dispersa em um padrão. O problema da raiz unitária, de não-estacionariedade, aparece quando o modelo da auto-regressivo da série de dados apresente uma raiz sobre o círculo unitário, pode ser observado no seguinte modelo auto-regressivo (MORETTIN, 2006; ALBUQUERQUE J. M. P. FORTES, 2008):

$$y^{(k)} = \rho_e y^{(k-1)} + erro^{(k)} \quad (\text{A.14})$$

onde $erro^{(k)}$ é o termo do erro estocástico e a raiz unitária existe para $\rho_e = 1$.

De acordo com (MORETTIN, 2006), o teste de *Dickey-Fuller* baseia-se nas seguintes hipóteses:

- $H_0: \rho_e = 1$ Hipótese nula, ou seja, conjunto de dados não-estacionários.
- $H_1: \rho_e < 1$ Hipótese alternativa, ou seja, conjunto de dados estacionários.

Conforme observado em (MORETTIN, 2006), o teste de *Dickey-Fuller*, se a hipótese nula for rejeitada, a série de dados não tem uma raiz unitária, consequentemente a mesma é estacionária.

APÊNDICE B – MÁXIMA VEROSSIMILHANÇA

Neste apêndice objetiva-se fornecer alguns conceitos básicos sobre verossimilhança e máxima verossimilhança. Estes conceitos permitem entender os estudos desenvolvidos nesta tese e para um estudo mais completo, as seguintes referências são recomendadas: (ISERMANN; NCHHOF, 2011).

A discussão a respeito dos estimadores de máxima verossimilhança é precedida pelos termos da verossimilhança. Seja um modelo linear expresso por:

$$\mathbf{Y} = \mathbf{\Psi}^T \mathbf{\Theta} + \mathbf{E} \quad (\text{B.1})$$

Neste modelo, \mathbf{Y} , cuja dimensão é $N \times 1$, é chamado de vetor de medidas; $\mathbf{\Theta}$, cuja dimensão é $p \times 1$, é chamado de vetor de parâmetros; $\mathbf{\Psi}$, cuja dimensão é $N \times p$, é chamado de vetor de observações.

B.1 Verossimilhança

Considere o vetor de parâmetros desconhecidos $\mathbf{\Theta}$ que descreve um conjunto de dados de N observações distribuídas identicamente independentemente $y^{(k)}$, $k = 1, 2, \dots, N$ e seja o vetor de medidas $\mathbf{Y} = \{z^{(1)}, z^{(2)}, \dots, z^{(N)}\}$ (ISERMANN; NCHHOF, 2011).

A verossimilhança de $\mathbf{\Theta}$, dado um conjunto de observações \mathbf{Y} , define-se ser proporcional (\propto) ao valor da função de densidade de probabilidade das observações dada pelo vetor de parâmetros, expresso por:

$$\tilde{l}(\mathbf{\Theta}|\mathbf{Y}) \propto \tilde{f}_{\mathbf{Y}}(\mathbf{Y}|\mathbf{\Theta}) \quad (\text{B.2})$$

onde \tilde{l} é a função de verossimilhança e $\tilde{f}_{\mathbf{Y}}$ a função densidade de probabilidade condicional. Admitindo-se que as observações nos instantes k são distribuídas independentemente e identicamente, tem-se:

$$\tilde{l}(\mathbf{\Theta}|\mathbf{Y}) \propto \tilde{f}_{\mathbf{Y}}(y^{(1)}|\mathbf{\Theta}) \tilde{f}_{\mathbf{Y}}(y^{(2)}|\mathbf{\Theta}) \dots \tilde{f}_{\mathbf{Y}}(y^{(N)}|\mathbf{\Theta}) \quad (\text{B.3})$$

Na maioria das aplicações observa-se que $\tilde{f}_{\mathbf{Y}}(\mathbf{Y}|\mathbf{\Theta})$ é exponencial, portanto, o uso do logaritmo natural se faz presente, ou seja,

$$\tilde{L}(\mathbf{\Theta}|\mathbf{Y}) = \ln [\tilde{l}(\mathbf{\Theta}|\mathbf{Y})] \quad (\text{B.4})$$

A função \tilde{L} é referida como sendo a função de log-verossimilhança.

B.2 Métodos e Estimadores de Máxima–Verossimilhança

O método de máxima–verossimilhança é baseado em uma ideia relativamente simples onde diferentes populações geram diferentes amostras e que uma amostra (isto é, um conjunto de dados) é mais semelhante a algumas populações em detrimento a outras amostras (??ISERMANN; NCHHOF, 2011).

O estimador de máxima–verossimilhança (EMV) $\hat{\Theta}_{mv}$ é o valor determinado de Θ que maximiza a função \tilde{L} , como segue:

$$\tilde{L}(\hat{\Theta}_{mv}|\mathbf{Y}) = \max [\tilde{L}(\Theta|\mathbf{Y})] \quad (\text{B.5})$$

A ideia é selecionar os parâmetros do estimador de máxima–verossimilhança $\hat{\Theta}_{mv}$, de modo que, estes maximizem a função de verossimilhança para valores verdadeiros dos estimadores Θ ; conseqüentemente, tem-se

$$\tilde{L}(\Theta|\mathbf{Y})\Big|_{\Theta=\hat{\Theta}_{mv}} \rightarrow \max \quad (\text{B.6})$$

Para determinar os valores a partir da estimação de máxima–verossimilhança a partir da função de verossimilhança relacionados aos parâmetros desconhecidos e obtendo-se a primeira derivada da função de máxima verossimilhança em relação aos parâmetros desconhecidos $\Theta = \{\theta_1, \theta_2, \dots, \theta_p\}$ e igualando-se a zero, tem-se (??ISERMANN; NCHHOF, 2011).

$$\frac{\partial \tilde{L}(\Theta|\mathbf{Y})}{\partial \theta_i} \Big|_{\Theta=\hat{\Theta}_{mv}} = 0 \quad \forall i = 1, 2, \dots, p \quad (\text{B.7})$$

B.3 Variância

Neste contexto, seja o seguinte modelo de função de transferência,

$$y^{(k)} = \frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}\psi^{(k)} + \frac{\mathbf{D}(z^{-1})}{\mathbf{C}(z^{-1})}e^{(k)} \quad (\text{B.8})$$

Conforme (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015), por simplicidade, ainda que neste tese não será necessário informação *a priori* da estatística do ruído, seja os polinômios $\mathbf{C}(z^{-1})$ e $\mathbf{D}(z^{-1})$, conhecido. De acordo com (YOUNG, 1984; YOUNG, 2011; YOUNG, 2015), a formulação da abordagem dos estimadores de máxima verossimilhança que requer a função de log–verossimilhança, para as N observações, é expressa por

$$\begin{aligned} \tilde{L}(\mathbf{a}, \mathbf{b}, \sigma^2, \mathbf{y}, \psi) = & -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) \dots \\ & \dots - \frac{1}{2\sigma^2} \left[\frac{\mathbf{C}}{\mathbf{D}}\mathbf{y} - \frac{\mathbf{BC}}{\mathbf{AD}}\psi \right]^T \left[\frac{\mathbf{C}}{\mathbf{D}}\mathbf{y} - \frac{\mathbf{BC}}{\mathbf{AD}}\psi \right] \end{aligned} \quad (\text{B.9})$$

onde $\mathbf{a} = [a_1, a_2, \dots, a_p]^T$, $\mathbf{b} = [b_1, b_2, \dots, b_p]^T$, $\mathbf{y} = [y_1, y_2, \dots, y_N]^T$ e $\boldsymbol{\psi} = [\psi_1, \psi_2, \dots, \psi_N]^T$ e a ordem dos vetores \mathbf{a} e \mathbf{b} , dada por p . O vetor de parâmetros estimados é determinado a partir da seguinte formulação:

$$\hat{\boldsymbol{\Theta}} = \arg \min_{\boldsymbol{\theta}} \tilde{L}(\boldsymbol{\Theta}, \mathbf{y}, \boldsymbol{\psi}) \quad (\text{B.10})$$

a partir das derivadas parciais dos termos da equação (B.9), obtém-se as seguintes equações:

$$\begin{aligned} \frac{\partial \tilde{L}}{\partial a_i} &= \frac{1}{\sigma^2} \sum_{k=1}^N \left[\frac{\mathbf{C}}{\mathbf{D}} \mathbf{y}^{(k)} - \frac{\mathbf{BC}}{\mathbf{AD}} \boldsymbol{\psi}^{(k)} \right] \\ \frac{\mathbf{BC}}{\mathbf{A}^2 \mathbf{D}} z^{(-i)} \boldsymbol{\psi}^{(k)} &= 0; i = 1, 2, \dots, p \end{aligned} \quad (\text{B.11})$$

$$\begin{aligned} \frac{\partial \tilde{L}}{\partial b_i} &= \frac{1}{\sigma^2} \sum_{k=1}^N \left[\frac{\mathbf{C}}{\mathbf{D}} \mathbf{y}^{(k)} - \frac{\mathbf{BC}}{\mathbf{AD}} \boldsymbol{\psi}^{(k)} \right] \\ \frac{\mathbf{C}}{\mathbf{AD}} z^{(-i)} \boldsymbol{\psi}^{(k)} &= 0; i = 1, 2, \dots, p \end{aligned} \quad (\text{B.12})$$

$$\begin{aligned} \frac{\partial L}{\partial \hat{\sigma}^2} &= -\frac{N}{\sigma^2} + \dots \\ \frac{1}{\sigma^4} \sum_{k=1}^N \left[\frac{\mathbf{C}}{\mathbf{D}} \mathbf{y}^{(k)} - \frac{\mathbf{BC}}{\mathbf{AD}} \boldsymbol{\psi}^{(k)} \right]^2 &= 0 \end{aligned} \quad (\text{B.13})$$

APÊNDICE C – PARÂMETROS INICIAIS DOS ALGORITMOS

- **Parâmetros de Filtragem**

$$\omega_u \ 0,90$$

$$\omega_y \ 0,90$$

- **Parâmetros de Agrupamentos**

$$\mu_h \ 0,50$$

$$P_{tol} \ 50$$

$$m \ 2,10$$

$$\omega_g \ 2,1$$

$$\epsilon \ 10^{-5}$$

- **Parâmetros de Agregação de Grupos**

$$\gamma_c \ 0,85$$

$$k_1 \ 0,98$$

$$k_2 \ 0,20$$

- **Parâmetros de Taxa de Aprendizado**

$$\lambda_0 \ 0,90$$

$$\lambda \ 0,90$$

$$\beta \ 0,05$$

$$\gamma_0 \ 0,99$$

$$\gamma \ 0,09$$

REFERÊNCIAS

ABONYI, J. *Fuzzy Model Identification for Control*. [S.l.]: Birkhauser Basel, 2003. Citado 7 vezes nas páginas 10, 43, 45, 81, 82, 83 e 94.

ABONYI, J.; FEIL, B. *Cluster Analysis for Data Mining and System Identification*. [S.l.]: Springer, 2007. ISBN 9783764379872. Citado 9 vezes nas páginas 10, 40, 43, 45, 55, 81, 82, 83 e 94.

ABONYI, J. et al. Modified gath–geva clustering for fuzzy segmentation of multivariate time-series. *Fuzzy Sets and Systems*, Elsevier, v. 149, n. 1, p. 39–56, 2005. Citado 9 vezes nas páginas 10, 32, 43, 44, 45, 81, 82, 83 e 95.

AGUIRRE, L. *Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. [S.l.]: Editora UFMG, 2015. Citado 12 vezes nas páginas 25, 27, 31, 32, 59, 60, 61, 62, 63, 64, 65 e 66.

ALBUQUERQUE J. M. P. FORTES, W. A. F. J. P. A. *Probabilidade, Variáveis Aleatórias e Processos Estocásticos*. [S.l.]: PUC Rio, 2008. ISBN 978-850-7193-190-9. Citado 6 vezes nas páginas 95, 187, 188, 189, 190 e 192.

ANGELOV, P. *Autonomous Learning Systems: From Data Streams to Knowledge in Real-time*. [S.l.]: IEEE Press Series on Computational Intelligence, Willey, 2013. ISBN 9781119951520. Citado 9 vezes nas páginas 27, 29, 33, 39, 47, 48, 55, 57 e 58.

ANGELOV, P.; FILEV, D.; KASABOV, N. *Evolving Intelligent Systems: methodology and applications*. [S.l.]: IEEE Press Series on Computational Intelligence, Willey–Blackwell, 2010. ISBN 9780470287194. Citado 16 vezes nas páginas 14, 29, 31, 33, 48, 54, 55, 81, 97, 98, 100, 101, 106, 144, 146 e 148.

ANGELOV, P.; ZHOU, X. Evolving fuzzy systems from data streams in real-time. In: *2006 International Symposium on Evolving Fuzzy Systems*. [S.l.: s.n.], 2006. p. 29–35. Citado 2 vezes nas páginas 148 e 155.

A.PAPOULIS; PILAI, U. *Probability, Random, Variables, and Stochastic Processes*. [S.l.]: McGraw-Hill, 2002. ISBN 9780070486584. Citado 6 vezes nas páginas 58, 95, 187, 188, 189 e 190.

ASSAWINCHAICHOTE, W.; NGUANG, S.; SHI, P. *Fuzzy Control and Filter Design for Uncertain Fuzzy Systems*. [S.l.]: Springer Berlin Heidelberg, 2007. (Lecture Notes in Control and Information Sciences). Citado na página 28.

BABUSKA B. SCHUTTER, Z. L. R.; GUERRA, T. M. *Stability Analysis and Nonlinear Observer Design Using Takagi-Sugeno Fuzzy Models*. [S.l.]: Springer Publishing Company, Incorporated, 2010. v. 262. (Studies in Fuzziness and Soft Computing, v. 262). Citado 10 vezes nas páginas 28, 31, 33, 36, 37, 86, 94, 101, 104 e 113.

BABUSKA, R. *Fuzzy Modeling Control*. [S.l.]: Kluwer Academic Publishers, 1998. Citado 27 vezes nas páginas 10, 15, 28, 29, 31, 32, 33, 36, 37, 40, 43, 44, 45, 55, 56, 57, 58, 81, 82, 83, 84, 86, 94, 101, 102, 104 e 156.

- BABUSKA, R.; VEEN, P. J. van der; KAYMAK, U. Improved covariance estimation for gustafson-kessel clustering. In: *Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on*. [S.l.: s.n.], 2002. v. 2, p. 1081–1085. Citado na página 104.
- BAHREMAND, A. Hess opinions: Advocating process modeling and de-emphasizing parameter estimation. *Hydrology and Earth System Sciences Discussions*, v. 12, n. 11, p. 12377–12393, 2015. Citado na página 25.
- BARUAH, R. D.; ANGELOV, P. Dynamically evolving clustering for data streams. In: *Evolving and Adaptive Intelligent Systems (EAIS), 2014 IEEE Conference on*. [S.l.: s.n.], 2014. p. 1–6. Citado 3 vezes nas páginas 29, 32 e 33.
- BENZAOUIA, A.; HAJJAJI, A. *Advanced Takagi Sugeno Fuzzy Systems: Delay and Saturation*. [S.l.]: Springer, 2014. Citado 6 vezes nas páginas 28, 29, 36, 37, 47 e 55.
- BORDIGNON, F.; GOMIDE, F. Uninorm based evolving neural networks and approximation capabilities. *Neurocomputing*, Elsevier, v. 127, n. 0, p. 13 – 20, 2014. Citado 2 vezes nas páginas 29 e 47.
- BOSCAINO, V. et al. Comparison of models of fuel cells based on experimental data for the design of power electronics systems. *Renewable Power Generation, IET*, v. 9, n. 6, p. 660–668, 2015. Citado na página 26.
- CAMINHAS, A. S. M.; LEMOS, A. P.; GOMIDE, F. Evolving neo-fuzzy neural network with adaptive feature selection. In: *Computational Intelligence and 11th Brazilian Congress on Computational Intelligence (BRICS-CCI CBIC), 2013 BRICS Congress on*. [S.l.: s.n.], 2013. p. 341–349. Citado 4 vezes nas páginas 29, 33, 47 e 101.
- CAPONETTO, R. et al. Ipmc frequency dependent multiphysics model considering electrodes high surface and fractional effects. In: *Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, 2014 IEEE International*. [S.l.: s.n.], 2014. p. 1529–1532. Citado na página 25.
- COSTA, B.; ANGELOV, P.; GUEDES, L. Fully unsupervised fault detection and identification based on recursive density estimation and self-evolving cloud-based classifier. *Neurocomputing*, v. 150, Part A, p. 289 – 303, 2015. Citado 2 vezes nas páginas 27 e 33.
- DADHE, K. et al. Neural networks as a tool for gray box modelling in reactive distillation. In: _____. *Fuzzy Days*. [S.l.]: Springer, 2001. (Lecture Notes in Computer Science, v. 2206), p. 576–588. Citado na página 26.
- DOVZAN, D.; LOGAR, V.; SKRJANC, I. Implementation of an evolving fuzzy model (efumo) in a monitoring system for a waste-water treatment process. *IEEE Transactions on Fuzzy Systems*, v. 23, n. 5, p. 1761–1776, 2015. Citado 3 vezes nas páginas 31, 33 e 47.
- DU, H.; ZHANG, N. Application of evolving takagi-sugeno fuzzy model to nonlinear system identification. *Appl. Soft Comput.*, v. 8, p. 676–686, 2008. Citado na página 28.
- GADEWADIKAR, J.; HORVAT, K.; KULJACA, O. A graphic user interface for h-infinity static output feedback controller design. Springer New York, v. 152, p. 527–536, 2013. Citado na página 156.

- GATH, I.; GEVA, A. B. Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 11, n. 7, p. 773–780, Jul 1989. ISSN 0162-8828. Citado 4 vezes nas páginas 45, 82, 83 e 94.
- GEORGIEVA, O.; FILEV, D. Gustafson-kessel algorithm for evolving data stream clustering. In: *Proceedings of the International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*. [S.l.]: ACM, 2009. p. 62:1–62:6. Citado 13 vezes nas páginas 14, 29, 31, 32, 54, 80, 81, 83, 84, 98, 144, 146 e 148.
- GHAHRAMANI, S. *Fundamentals of Probability with Stochastic Processes*. [S.l.]: PEARSON, 2005. ISBN 0-13-145340-8. Citado na página 95.
- GOODWIN, G.; PAYNE, R. *Dynamic System Identification: Experiment Design and Data Analysis*. [S.l.]: Academic Press, 1977. ISBN 9780122897504. Citado 4 vezes nas páginas 59, 61, 65 e 66.
- GUO, F. et al. T-s fuzzy model identification of mimo nonlinear systems based on data-driven. In: *Electronics, Communications and Control (ICECC), 2011 International Conference on*. [S.l.: s.n.], 2011. p. 1186–1189. Citado na página 28.
- HASSANEIN, O. et al. Auto-generating fuzzy system modelling of physical systems. In: *Control Applications (CCA), 2015 IEEE Conference on*. [S.l.: s.n.], 2015. p. 1142–1147. Citado na página 26.
- ISERMANN, R.; NCHHOF, M. *Identification of Dynamic Systems*. [S.l.]: Springer, 2011. Citado 8 vezes nas páginas 25, 26, 27, 31, 32, 61, 193 e 194.
- JEEVA, A.; PONNUSAMY, L.; SANGAPILLAI, S. T-s fuzzy modeling for siso and mimo processes. In: *Advanced Computing (ICoAC), 2013 Fifth International Conference on*. [S.l.: s.n.], 2013. p. 227–235. Citado 3 vezes nas páginas 28, 55 e 86.
- JUANG, J.-N. *Applied System Identification*. [S.l.]: Prentice Hall, 1994. Citado 2 vezes nas páginas 25 e 27.
- KASABOV, N. *Evolving Connectionist Systems*. [S.l.]: Springer, 2007. Citado 3 vezes nas páginas 29, 31 e 48.
- KASABOV, N.; FILEV, D. Evolving intelligent systems: Methods, learning, applications. In: *Evolving Fuzzy Systems, 2006 International Symposium on*. [S.l.: s.n.], 2006. p. 8–18. Citado 3 vezes nas páginas 29, 31 e 48.
- KASABOV, N. e. Q. S. Denfis: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *Fuzzy Systems, IEEE Transactions on*, v. 10, n. 2, p. 144–154, 2002. Citado 6 vezes nas páginas 48, 54, 80, 81, 83 e 84.
- KEESMAN, K. J. *System Identification: an introduction*. [S.l.]: Springer, 2011. Citado 3 vezes nas páginas 25, 27 e 40.
- KOHONEN, T. The self-organizing map. *Neurocomputing*, Elsevier, v. 21, n. 1, p. 1–6, 1998. Citado na página 98.
- LEITE, D. *Evolving Granular Systems*. Tese (Doutorado) — UNICAMP/SP, 2012. Citado 2 vezes nas páginas 31 e 47.

- LEITE, D.; COSTA, P.; GOMIDE, F. Evolving granular neural networks from fuzzy data streams. *Neural Networks*, Elsevier, v. 38, p. 1–16, 2013. Citado 2 vezes nas páginas 31 e 47.
- LEITE, D. et al. Fuzzy granular evolving modeling for time series prediction. In: *Fuzzy Systems (FUZZ), 2011 IEEE International Conference on*. [S.l.: s.n.], 2011. p. 2794–2801. Citado na página 33.
- LEMONS, A.; CAMINHAS, W.; GOMIDE, F. Multivariable gaussian evolving fuzzy modeling system. *IEEE Transactions on Fuzzy Systems*, v. 19, n. 1, p. 91–104, 2011. Citado 4 vezes nas páginas 113, 114, 148 e 155.
- LI, Z. *Fuzzy Chaotic Systems - Modeling, Control and Applications*. [S.l.]: Springer - Verlag, 2006. Citado 2 vezes nas páginas 28 e 36.
- LIMA, E. et al. Evolving fuzzy modeling using participatory learning. *Evolving intelligent systems: methodology and applications*, Wiley-Interscience/IEEE Press, p. 67–86, 2010. Citado 3 vezes nas páginas 33, 100 e 101.
- LIN, Z. *Understanding and Simulating Software Evolution*. [S.l.]: IEEE Press, 2013. 1411-1414 p. (ICSE 2013). Citado na página 25.
- LIU, Y.; ZHANG, W.; ZHANG, Y. Data driven modeling of human welder intelligence: A neuro-fuzzy approach. In: *Automation Science and Engineering (CASE), 2013 IEEE International Conference on*. [S.l.: s.n.], 2013. p. 663–668. Citado na página 33.
- LJUNG, L. *System Identification: Theory for the User*. [S.l.]: Prentice Hall, 1999. ISBN 0136566952. Citado 6 vezes nas páginas 25, 27, 31, 59, 65 e 66.
- LUGHOFER, E. *Evolving Fuzzy Systems: Methodologies, Advanced Concepts and Applications*. [S.l.]: Springer, 2011. ISBN 9783642180866. Citado 8 vezes nas páginas 27, 28, 46, 47, 48, 49, 55 e 81.
- LUGHOFER, E. et al. Generalized smart evolving fuzzy systems. *Evolving Systems*, v. 6, n. 4, p. 269–292, 2015. Citado 2 vezes nas páginas 27 e 33.
- LUGHOFER, E. D. Flexfis: A robust incremental learning approach for evolving takagi; sugeno fuzzy. *IEEE Transactions on Fuzzy Systems*, v. 16, n. 6, p. 1393–1410, Dec 2008. ISSN 1063-6706. Citado 4 vezes nas páginas 31, 47, 148 e 155.
- MAAN, N.; AHMAD, T. Fuzzy number in determining the optimal solution of multiple input multiple output (mimo) system. In: *Computational Intelligence for Modelling Control Automation, 2008 International Conference on*. [S.l.: s.n.], 2008. p. 1182–1187. Citado na página 29.
- MACIEL, L.; GOMIDE, F.; BALLINI, R. Mimo evolving participatory learning fuzzy modeling. In: *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*. [S.l.: s.n.], 2012. p. 1–8. Citado 2 vezes nas páginas 33 e 101.
- MAGERL, M.; CEPERIC, V.; BARIC, A. Echo state networks for black-box modelling of integrated circuits. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, PP, n. 99, p. 1–1, 2015. Citado na página 26.

- MANGOUROVA, V.; RINGWOOD, J. Gray box modelling of arterial vasoaction. In: *Irish Signals and Systems Conference, 2006. IET*. [S.l.: s.n.], 2006. p. 315–320. Citado na página 26.
- MORETTIN, C. M. C. T. P. A. *Análise de Séries Temporais*. [S.l.]: Blucher, 2006. ISBN 9788521203896. Citado 7 vezes nas páginas 17, 113, 114, 162, 164, 191 e 192.
- PAL, N.; SARKAR, K. What and when can we gain from the kernel versions of c-means algorithm? *Fuzzy Systems, IEEE Transactions on*, v. 22, n. 2, p. 363–379, 2014. Citado na página 37.
- PEDRYCZ, J. V. de O. W. *Advances in fuzzy clustering and its applications*. [S.l.]: Wiley Online Library, 2007. Citado na página 40.
- PEDRYCZ, W.; GOMIDE, F. *Fuzzy Systems Engineering: Toward Human-Centric Computing*. [S.l.]: Wiley-IEEE Press, 2007. ISBN 0471788570. Citado 7 vezes nas páginas 28, 29, 46, 47, 48, 55 e 100.
- PETELIN, D.; KOCIJAN, J. Evolving gaussian process models for predicting chaotic time-series. In: *Evolving and Adaptive Intelligent Systems (EAIS), 2014 IEEE Conference on*. [S.l.: s.n.], 2014. p. 1–8. Citado na página 32.
- PIRES, D. S.; SERRA, G. L. de O. Fuzzy digital pid controller design based on robust stability criteria. In: *2014 12th IEEE International Conference on Industrial Informatics (INDIN)*. [S.l.: s.n.], 2014. p. 654–659. ISSN 1935-4576. Citado na página 143.
- PRATAMA, M.; ANAVATTI, S.; LUGHOFER, E. Evolving fuzzy rule-based classifier based on genefis. In: *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*. [S.l.: s.n.], 2013. p. 1–8. Citado 2 vezes nas páginas 27 e 33.
- QUANSER. *2-DOF Helicopter Experiment Manual*. [S.l.: s.n.], 2014. Citado na página 156.
- RAHMAN, R. Z. A.; SHOUMY, N. J. Neural network controller for two-degree-freedom helicopter control system. In: *Control, Automation and Systems (ICCAS), 2012 12th International Conference on*. [S.l.: s.n.], 2012. p. 101–106. Citado na página 157.
- RUBIO, J. d. J. Sofmls: Online self-organizing fuzzy modified least-squares network. *IEEE Transactions on Fuzzy Systems*, v. 17, n. 6, p. 1296–1309, Dec 2009. ISSN 1063-6706. Citado 4 vezes nas páginas 31, 47, 148 e 155.
- SERRA, G. *Frontiers in Advanced Control Systems*. [S.l.]: InTech, 2012. ISBN 9789535106777. Citado 7 vezes nas páginas 10, 31, 32, 55, 63, 66 e 67.
- SERRA, G.; BOTTURA, C. Fuzzy instrumental variable concept and identification algorithm. In: *Fuzzy Systems, 2005. The 14th IEEE International Conference on*. [S.l.: s.n.], 2005. p. 1062–1067. Citado na página 75.
- SERRA, G.; BOTTURA, C. An iv-qr algorithm for neuro-fuzzy multivariable online identification. *Fuzzy Systems, IEEE Transactions on*, v. 15, n. 2, p. 200–210, 2007. Citado 7 vezes nas páginas 10, 32, 55, 66, 67, 75 e 114.

- SERRA, G.; BOTTURA, C. *Fuzzy Instrumental Variable approach for nonlinear discrete-time systems identification in a noisy environment*. [S.l.: s.n.], 2009. v. 160. 500-520 p. Citado 3 vezes nas páginas 56, 75 e 160.
- SILVA, A. et al. Real-time nonlinear modeling of a twin rotor mimo system using evolving neuro-fuzzy network. In: *Computational Intelligence in Control and Automation (CICA), 2014 IEEE Symposium on*. [S.l.: s.n.], 2014. p. 1–8. Citado na página 32.
- SKRJANC, I. Evolving fuzzy-model-based design of experiments with supervised hierarchical clustering. *IEEE Transactions on Fuzzy Systems*, v. 23, n. 4, p. 861–871, 2015. Citado na página 33.
- SKRJANC, I.; DOVZAN, D.; GOMIDE, F. Evolving fuzzy-madel-based on c-regression clustering. In: *Evolving and Adaptive Intelligent Systems (EAIS), 2014 IEEE Conference on*. [S.l.: s.n.], 2014. p. 1–7. Citado 3 vezes nas páginas 32, 47 e 113.
- SMITH, S. W. *The Scientist and Engineers Guide to Digital Signal Processing*. San Diego, CA, USA: California Technical Publishing, 2002. ISBN 075067444X. Citado 2 vezes nas páginas 70 e 71.
- SODERSTROM, T.; STOICA, P. *Instrumental Variable Methods for System Identification*. [S.l.]: Springer-Verlag, 1983. ISBN 038712814X. Citado 5 vezes nas páginas 32, 60, 63, 64 e 70.
- SODERSTROM, T.; STOICA, P. *System Identification*. [S.l.]: Prentice Hall, 1989. ISBN 9780131276062. Citado 10 vezes nas páginas 32, 59, 60, 61, 62, 63, 64, 65, 66 e 70.
- SOLEIMANI-B, H.; C.; BABAK, A. Recursive gath-geva clustering as a basis for evolving neuro-fuzzy modeling. In: *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*. [S.l.: s.n.], 2010. p. 1–7. Citado 2 vezes nas páginas 45 e 95.
- SUN, Z. et al. Knowledge-based evolving clustering algorithm for data stream. In: *Service Systems and Service Management (ICSSSM), 2014 11th International Conference on*. [S.l.: s.n.], 2014. p. 1–6. Citado 2 vezes nas páginas 32 e 36.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-15, n. 1, p. 116–132, 1985. Citado na página 55.
- THORDARSON, F.; DANMARKS. *Gray Box Modelling of Hydrological Systems: With Focus on Uncertainties*. [S.l.]: DTU Informatics, 2012. Citado na página 26.
- TSUMURA, K. Optimal quantization of signals for system identification. *IEEE Transactions on Automatic Control*, v. 54, n. 12, p. 2909–2915, Dec 2009. ISSN 0018-9286. Citado na página 82.
- TSUMURA, K.; MACIEJOWSKI, J. Optimal quantization of signals for system identification. In: *2003 European Control Conference (ECC)*. [S.l.: s.n.], 2003. p. 785–790. Citado na página 82.
- TSUMURA, K.; OISHI, Y. Optimal length of data for identification of time varying system. In: *Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No.99CH36304)*. [S.l.: s.n.], 1999. v. 4, p. 3224–3229 vol.4. ISSN 0191-2216. Citado na página 82.

- WANG, L.; GARNIER, H. *System Identification, Environmental Modelling, and Control System Design*. [S.l.]: Springer Publishing Company, Incorporated, 2012. ISBN 978-0-85729-973-4. Citado 10 vezes nas páginas 25, 27, 31, 59, 61, 62, 63, 64, 65 e 66.
- WANG, L.-X. *A Course in Fuzzy Systems and Control*. [S.l.]: Prentice-Hall, Inc., 1997. Citado 2 vezes nas páginas 28 e 55.
- WATTS, M. A decade of kasabov's evolving connectionist systems: A review. *Systems, Man, e Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, v. 39, n. 3, p. 253–269, 2009. Citado 2 vezes nas páginas 29 e 31.
- WILLIAMSON, P. P. Mathematical statistics with applications. asha seth kapadia, wenyaw chan, and lemuel moye. *Journal of the American Statistical Association*, v. 101, p. 394–395, 2006. Citado na página 79.
- XIAO, C. et al. Using spearman's correlation coefficients for exploratory data analysis on big dataset. *Concurr. Comput. : Pract. Exper.*, John Wiley and Sons Ltd., Chichester, UK, v. 28, n. 14, p. 3866–3878, set. 2016. ISSN 1532-0626. Disponível em: <<https://doi.org/10.1002/cpe.3745>>. Citado 3 vezes nas páginas 114, 115 e 116.
- YOUNG, P. *Recursive Estimation and Time-series Analysis: An Introduction*. [S.l.]: Springer-Verlag New York, Inc., 1984. ISBN 0387136770. Citado 9 vezes nas páginas 10, 32, 59, 62, 65, 66, 67, 73 e 194.
- YOUNG, P. *Recursive Estimation and Time-Series Analysis. An Introduction for the student and Practitioner*. [S.l.]: Springer, 2011. ISBN 9783642219818. Citado 13 vezes nas páginas 10, 32, 59, 60, 62, 63, 64, 65, 66, 67, 73, 74 e 194.
- YOUNG, P. Refined instrumental variable estimation: maximum likelihood optimization of a unified box-jenkins model. *Automatica*, v. 52, p. 35–46, 2015. Citado 12 vezes nas páginas 10, 32, 59, 60, 62, 63, 64, 65, 66, 67, 116 e 194.
- ZAERI, A. H. et al. Disturbance rejection for a 2-dof nonlinear helicopter model by using mimo fuzzy sliding mode control with boundary layer. In: *Intelligent Systems, Modelling and Simulation (ISMS), 2012 Third International Conference on*. [S.l.: s.n.], 2012. p. 411–416. Citado na página 157.
- ZHU, Y. A box-jenkins method that is asymptotically globally convergent for open loop data. *IFAC Proceedings Volumes*, Elsevier, v. 44, n. 1, p. 9047–9051, 2011. Citado 2 vezes nas páginas 45 e 116.