

UNIVERSIDADE FEDERAL DO MARANHÃO
Programa de Pós-Graduação em Engenharia Elétrica

Ben-Hur Matthews Moreno Montel

**Filtragem Adaptativa de Kalman Nebulosa
Tipo-2 Intervalar Baseada em Critério de
Máxima Verossimilhança**

São Luís - MA

2024

Ben-Hur Matthews Moreno Montel

**Filtragem Adaptativa de Kalman Nebulosa Tipo-2
Intervalar Baseada em Critério de Máxima
Verossimilhança**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica, ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Maranhão.

Programa de Pós-Graduação em Engenharia Elétrica

Universidade Federal do Maranhão

Orientador: Prof. Dr. Ginalber Luiz de Oliveira Serra

São Luís - MA

2024

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

Montel, Ben-Hur Matthews Moreno.

Filtragem adaptativa de Kalman nebulosa tipo-2
intervalar baseada em critério de máxima verossimilhança /
Ben-Hur Matthews Moreno Montel. - 2024.

132 p.

Orientador(a): Ginalber Luiz de Oliveira Serra.

Dissertação (Mestrado) - Programa de Pós-graduação em
Engenharia Elétrica/ccet, Universidade Federal do
Maranhão, São Luís, 2024.

1. Agrupamento nebuloso tipo-2 intervalar. 2.
Critério da informação de Akaike. 3. Filtro de Kalman. 4.
Máxima Verossimilhança. 5. Séries temporais
multivariáveis. I. Serra, Ginalber Luiz de Oliveira. II.
Título.

Ben-Hur Matthews Moreno Montel

Filtragem Adaptativa de Kalman Nebulosa Tipo-2 Intervalar Baseada em Critério de Máxima Verossimilhança

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica, ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Maranhão.

Trabalho . São Luís - MA, 09 de Fevereiro de 2024:

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

Orientador
Instituto Federal do Maranhão

**Prof. Dr. Francisco Das Chagas de
Souza**

Examinador Interno
Universidade Federal do Maranhão

Prof. Dr. Mateus Giesbrecht

Examinador Externo
Universidade Estadual de Campinas

São Luís - MA
2024

Agradecimentos

Agradeço primeiramente a Deus pela dádiva da vida, saúde e disposição concedidas, o que foi essencial para a conclusão deste trabalho.

Agradeço a meus pais, Arnaldo e Ana Rosa, e a minha irmã Marianna, por todo apoio, incentivo e lições que levarei adiante. Agradeço também à meus tios e tias por todo o suporte nesta caminhada.

Agradeço ao meu orientador, professor Dr. Ginalber Serra, por todos os ensinamentos, conhecimento e momentos compartilhados nesta jornada.

Agradeço aos meus ex-orientadores, professora Dr. Danúbia Pires e professor Dr. Orlando Filho, pelo suporte durante a graduação, o incentivo e exemplo para iniciar o mestrado.

Agradeço ao grupo de pesquisa em Inteligência Computacional Aplicada à Tecnologia - ICAT, em especial aos amigos do laboratório, Daiana Gomes, Jefferson Georgy e Luis Miguel, pelas discussões que incentivaram o desenvolvimento deste trabalho.

Agradeço a todos os amigos feitos durante toda a minha jornada acadêmica (escola, graduação e mestrado), por todos os conselhos e momentos compartilhados.

Agradeço a Universidade Federal do Maranhão - UFMA, em especial ao Programa de Pós-Graduação em Engenharia Elétrica - PPGEE, e ao Instituto Federal de Educação - IFMA campus Monte Castelo pelo apoio estrutural fornecido durante o desenvolvimento deste trabalho.

Agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pelo suporte financeiro.

“Podemos considerar o presente estado do universo como resultado de seu passado e a causa do seu futuro. Se um intelecto em certo momento tiver conhecimento de todas as forças que colocam a natureza em movimento, e a posição de todos os itens dos quais a natureza é composta, e se esse intelecto for grandioso o bastante para submeter tais dados à análise, ele incluiria numa única fórmula os movimentos dos maiores corpos do universo e também os do átomo mais diminutos; para tal intelecto nada seria incerto e o futuro, assim como o passado, estaria ao alcance de seus olhos”

(Pierre Simon Laplace)

Resumo

Nesta dissertação, uma metodologia para projeto de filtros de Kalman nebulosos tipo-2 intervalares baseada em critério de máxima verossimilhança é proposta. A metodologia desenvolvida consiste em etapas *offline* e *online*. Na primeira etapa, um treinamento para uma base de dados inicial é realizado, enquanto que na etapa *online*, correções ao modelo são feitas a cada nova medição disponível. Para o treinamento *offline*, a base de dados é pré-processada pela versão em batelada do algoritmo de análise espectral singular. A base de dados resultante é então particionada pela versão nebulosa tipo-2 intervalar do algoritmo de agrupamento de estimação em máxima verossimilhança baseada na análise de componentes principais. Finalizando a etapa *offline*, um submodelo de filtro de Kalman é estimado para cada partição obtida. Esta estimação é realizada por uma versão do algoritmo *Observer/Kalman Filter Identification* desenvolvida de forma a propiciar que cada regra ou submodelo possua sua própria ordem. Esta é definida com base no critério da informação de Akaike. Durante a etapa *online*, o modelo é corrigido a cada nova medição. Para isso, foi desenvolvida uma adaptação para a versão recursiva da análise espectral singular, na qual operações redundantes são removidas a fim de reduzir o tempo necessário para processar as componentes espectrais mais significativas da medição mais recente. O erro entre a estimativa do modelo e a atualização das componentes espectrais é utilizado para corrigir os submodelos através da versão recursiva e refinada do algoritmo *Observer/Kalman Filter Identification*, desenvolvida de forma a permitir a variação temporal na ordem do modelo. Por fim, uma modelagem acerca da incerteza da base de dados é desenvolvida e utilizada para adaptar o expoente de ponderação nebulosa utilizado no algoritmo de agrupamento, permitindo assim adaptar o intervalo de incerteza de acordo com a base de dados. Para validar a metodologia proposta, no intuito de demonstrar sua eficiência em diferentes cenários, testes computacionais e experimentais foram realizados. Como resultado computacional, destaca-se a previsão intervalar multivariável referente à posição e velocidade relativas entre satélites dispostos em formação PRISMA. Como resultado experimental, destacam-se o rastreamento intervalar da elevação e azimute de um helicóptero 2DoH, e o rastreamento e previsão intervalares da trajetória de um foguete de treinamento intermediário (veículo suborbital VS-30) do centro de lançamento em Alcântara-MA.

Palavras-chave: Filtro de Kalman, Máxima Verossimilhança, agrupamento nebuloso tipo-2 intervalar, séries temporais multivariáveis, Previsão, Análise de Componentes Principais, Modelagem de incerteza, Critério da informação de Akaike.

Abstract

This dissertation proposes a methodology for an interval type-2 fuzzy Kalman Filter project based on maximum likelihood criteria. The developed methodology consists of offline and online stages. In the first stage, an initial data set is used for training, while in the online stage, a model update is performed for each new available measurement. In the offline stage, an initial data set is pre-processed by the batch version of the singular spectral analysis. The resultant data set is then partitioned by the interval type-2 fuzzy version of the maximum likelihood estimation clustering algorithm based on principal component analysis. By the end of the offline stage, a Kalman filter sub model is estimated for each obtained partition. This estimation is performed by a fuzzy version of the Observer/Kalman Filter Identification algorithm, developed in a way that allows each rule or submodule to have its own order. This operation is based on the Akaike information criterion. During the online operation, the model is updated with each new measurement. For this, an adaptation for the recursive singular spectral analysis update is created in such a way that redundant operations are removed, reducing the time required to update the most significant components of the new measurement. The difference between the model estimate and the updated spectral components is used to update the submodels via the recursive and refined Observer/Kalman Filter Identification algorithm, which was designed to allow time change in each submodule order. Finally, a modeling of the data set uncertainty is developed and used to adapt the fuzzy weighting exponent used in the clustering algorithm, allowing the estimated interval to adapt to the data set. In order to test the proposed methodology with the intention of demonstrating its efficacy in different scenarios, computational and experimental tests were performed. As computational results, the multivariable forecast about the position and velocity of the PRISMA spacecraft formation stands out. As experimental results, the interval tracking of the helicopter with 2 degrees of freedom, and the interval forecasting of an intermediary training rocket (suborbital vehicle VS-30) of the Alcântara launch center stands out.

Keywords: Kalman Filter, Maximum Likelihood, Interval Type-2 Fuzzy clustering, Multivariate Time Series, Tracking, Principal Component Analysis, Uncertain Modeling, Akaike information criterion.

Lista de ilustrações

Figura 1 – Probabilidade de um dado ser selecionado dentro do universo de discurso. a) Densidade de probabilidade b) Curvas de nível.	32
Figura 2 – Partição obtida pelo algoritmo de agrupamento FCM. O conjunto de dados foi gerado pelo autor.	34
Figura 3 – Partição obtida pelo algoritmo de agrupamento GK. O conjunto de dados foi gerado pelo autor.	34
Figura 4 – Partição obtida pelo algoritmo de agrupamento ANEMV. O conjunto de dados foi gerado pelo autor.	35
Figura 5 – Partição realizada pelo algoritmo ANEMV com diferentes valores para o expoente de ponderação nebulosa. As curvas de nível representam o grau de pertinência em cada ponto da base de dados. A cor amarela representa valores próximos a 1, enquanto que as linhas roxas representam graus de pertinência tendendo a 0.	36
Figura 6 – Conjunto nebuloso tipo-2 com formato Gaussiano. a) Representação tridimensional do conjunto, apresentando os graus de pertinência primário e secundário. b) Projeção do grau de pertinência secundário no espaço $[x, \mu_{ik}]$, a cor amarela representa um grau de pertinência secundário próximo a 1, enquanto a cor azul representa uma pertinência próxima a zero.	37
Figura 7 – Função de pertinência para conjunto nebuloso tipo-2 Intervalar. A região sombreada representa os pontos onde o grau de pertinência secundário é igual a um.	37
Figura 8 – Funções de pertinência geradas por um algoritmo de agrupamento nebuloso com diferentes valores para m	38
Figura 9 – Fluxograma contendo os passos computacionais desenvolvidos na metodologia proposta.	60
Figura 10 – Modelo de função interpoladora para selecionar a abertura do intervalo de \tilde{m} conforme a incerteza da série temporal. Nela, é demonstrado o comportamento da função de acordo com a seleção do valor mínimo do intervalo.	70
Figura 11 – Fluxograma do processo de seleção da quantidade máxima de regressores p_m	74
Figura 12 – Diagrama de blocos para a simulação do movimento dos satélites da formação PRISMA.	88

Figura 13 – Posição relativa entre os satélites em formação PRISMA. A variável \mathbf{y}_1 representa o eixo X do vetor posição, \mathbf{y}_2 representa o eixo Y deste vetor, e \mathbf{y}_3 representa o eixo Z	88
Figura 14 – Posição relativa entre os satélites em formação. A variável \mathbf{y}_1 representa o eixo X do vetor velocidade, \mathbf{y}_2 representa o eixo Y deste vetor, e \mathbf{y}_3 representa o eixo Z	89
Figura 15 – Distribuição do ruído de medição presente em cada uma das variáveis de saída.	89
Figura 16 – Partição nebulosa para o conjunto de dados da posição relativa com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.	90
Figura 17 – Funções de pertinência intervalar para os cinco conjuntos nebulosos estimados.	91
Figura 18 – Comparação entre o intervalo estimado e o valor real da distância no eixo X entre os satélites da formação PRISMA no referencial $LVLH$	91
Figura 19 – Comparação entre o intervalo estimado e o valor real da distância no eixo Y entre os satélites da formação PRISMA no referencial $LVLH$	92
Figura 20 – Comparação entre o intervalo estimado e o valor real da distância no eixo Z entre os satélites da formação PRISMA no referencial $LVLH$	92
Figura 21 – Distribuição do erro e acurácia obtida nas 100 realizações ruidosas.	93
Figura 22 – Limites superior e inferior para o expoente de ponderação nebulosa \tilde{m} adaptado segundo o nível de incerteza das medições de cada variável.	94
Figura 23 – Variação temporal na ordem de cada um dos cinco submodelos.	95
Figura 24 – Previsão intervalar para o eixo X realizada durante as situações de falta de medição na simulação mista.	96
Figura 25 – Previsão intervalar para o eixo Y realizada durante as situações de falta de medição na simulação mista.	97
Figura 26 – Previsão intervalar para o eixo Z realizada durante as situações de falta de medição na simulação mista.	97
Figura 27 – Decomposição da base de dados em função de suas duas componentes principais, classificadas de acordo com os cinco conjuntos nebulosos adotados	98
Figura 28 – Partição nebulosa para o conjunto de dados da velocidade relativa com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.	99
Figura 29 – Funções de pertinência intervalar para os cinco conjuntos nebulosos estimados.	99
Figura 30 – Comparação entre o intervalo estimado e o valor real da velocidade no eixo X entre os satélites da formação PRISMA no referencial $LVLH$	100

Figura 31 – Comparação entre o intervalo estimado e o valor real da velocidade no eixo Y entre os satélites da formação PRISMA no referencial $LVLH$	100
Figura 32 – Comparação entre o intervalo estimado e o valor real da velocidade no eixo Z entre os satélites da formação PRISMA no referencial $LVLH$	101
Figura 33 – Distribuição do erro e acurácia obtida nas 100 realizações ruidosas.	101
Figura 34 – Limites superior e inferior para o expoente de ponderação nebulosa \tilde{m} adaptado segundo o nível de incerteza das medições de cada variável.	102
Figura 35 – Variação temporal na ordem de cada um dos cinco submodelos.	103
Figura 36 – Simulação com ausência de medições para a velocidade relativa no eixo X . As áreas sombreadas representam as regiões onde a simulação livre ocorreu.	104
Figura 37 – Simulação com ausência de medições para a velocidade relativa no eixo Y . As áreas sombreadas representam as regiões onde a simulação livre ocorreu.	104
Figura 38 – Simulação com ausência de medições para a velocidade relativa no eixo Z . As áreas sombreadas representam as regiões onde a simulação livre ocorreu.	105
Figura 39 – Decomposição da base de dados em função de suas duas componentes principais, classificadas de acordo com os cinco conjuntos nebulosos adotados	105
Figura 40 – Helicóptero com dois graus de liberdade.	107
Figura 41 – Entradas de controle do Helicóptero.	107
Figura 42 – Variáveis de Saída do Helicóptero.	108
Figura 43 – Partição nebulosa para o conjunto de dados com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.	108
Figura 44 – Funções de pertinência para a etapa <i>offline</i> da metodologia.	109
Figura 45 – Comparação entre o intervalo estimado e o valor real para o ângulo de elevação e azimute.	109
Figura 46 – Limites superior e inferior para o expoente de ponderação nebulosa \tilde{m} adaptado segundo o nível de incerteza das medições de cada variável.	110
Figura 47 – Veículo suborbital VS-30 (COMAER). Fonte: Força aérea brasileira (FAB MIL).	111
Figura 48 – Partição nebulosa para o conjunto de dados com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.	112
Figura 49 – Funções de pertinência para a etapa <i>offline</i> da metodologia.	113
Figura 50 – Comparação entre o intervalo estimado e o valor real do ângulo de elevação do VS-30 em graus.	113

Figura 51 – Comparação entre o intervalo estimado e o valor real do ângulo de azimute do VS-30 em graus.	114
Figura 52 – Comparação entre o intervalo estimado e o valor real do distância percorrida pelo VS-30 em <i>km</i>	114
Figura 53 – Limites superior e inferior para o expoente de ponderação nebulosa \tilde{m} adaptado segundo o nível de incerteza das medições de cada variável. . .	115
Figura 54 – Variação temporal da ordem para os dois submodelos estimados.	116
Figura 55 – Simulação com ausência de medições para o ângulo de elevação. As áreas sombreadas representam as regiões onde a simulação livre ocorreu.	117
Figura 56 – Simulação com ausência de medições para o ângulo de azimute. As áreas sombreadas representam as regiões onde a simulação livre ocorreu.	117
Figura 57 – Simulação com ausência de medições para Distância percorrida. As áreas sombreadas representam as regiões onde a simulação livre ocorreu.	118

Lista de tabelas

Tabela 1 – Parâmetros iniciais para a estimação paramétrica da metodologia. . . .	89
Tabela 2 – Quantidade de componentes espectrais selecionadas.	90
Tabela 3 – Comparação com outras metodologias presentes na literatura.	92
Tabela 4 – Desempenho da simulação livre.	96
Tabela 5 – Parâmetros iniciais para a estimação paramétrica da metodologia. . . .	98
Tabela 6 – Quantidade de componentes espectrais selecionadas.	98
Tabela 7 – Comparação com outras metodologias presentes na literatura.	100
Tabela 8 – Regiões onde a simulação livre foi realizada.	103
Tabela 9 – Parâmetros iniciais para a estimação paramétrica da metodologia. . . .	106
Tabela 10 – Quantidade de componentes espectrais selecionadas.	108
Tabela 11 – Acurácia e erro da estimativa obtida.	110
Tabela 12 – Parâmetros iniciais para a estimação paramétrica da metodologia. . . .	112
Tabela 13 – Quantidade de componentes espectrais selecionadas.	112
Tabela 14 – Acurácia e erro da estimativa obtida.	115
Tabela 15 – Parâmetros iniciais para a estimação paramétrica da metodologia. . . .	117

Lista de abreviaturas e siglas

ARX	<i>Autoregressive with Exogenous input</i>
ARMA	<i>Autoregressive with Moving Average</i>
ARMAX	<i>Autoregressive with Moving Average and Exogenous input</i>
NARX	<i>Non-Linear Autoregressive with Exogenous input</i>
NARMAX	<i>Non-Linear Autoregressive with Moving Average and Exogenous input</i>
OKID	<i>Observer/Kalman Filter Identification</i>
ERA	<i>Eigensystem Realization Algorithm</i>
TS	<i>Takagi and Sugeno</i>
ANT-2IEMV	Agrupamento Nebuloso Tipo-2 Intervalar de Estimação em Máxima Verossimilhança
FOGTREIN	Foguete de Treinamento
CLA	Centro de Lançamento de Alcântara
H-2GL	Helicóptero com 2 Graus de Liberdade
AES	Análise Espectral Singular
PCA	<i>Principal Component Analysis</i>
FCM	<i>Fuzzy C-means</i>
GK	Gustafson-Kessel
ANEMV	Agrupamento Nebuloso de Estimação em Máxima Verossimilhança
NaN	<i>Not a Number</i>
COMAER	Comando da Força aérea brasileira
RMSE	Do inglês <i>Root Mean-Square Error</i>
LVLH	<i>Local-Vertical-Local-Horizontal</i>

Lista de símbolos

\mathbf{Z}	Matriz de dados
\mathbf{Z}_k	Vetor de dados
\mathbf{U}	Matriz de partição
μ_{ik}	Grau de pertinência
\mathbf{v}_i	Centro de i -ésimo conjunto
$\ \cdot \ _S$	Operador norma de distância
\mathbf{P}_i	Probabilidade a priori
c	Quantidade de conjunto nebulosos
ϵ	Critério de tolerância do algoritmo de agrupamento nebuloso
\mathbf{F}_i	Matriz de covariância do i -ésimo conjunto nebuloso
\mathcal{D}_k^i	Norma de distância do agrupamento
m	Expoente de ponderação nebulosa
$\tilde{(\cdot)}$	Variável ou parâmetro intervalar
$(\cdot)^+$	Limite superior do intervalo
$(\cdot)^-$	Limite inferior do intervalo
\mathbf{x}_k	Vetor de estados
\mathbf{y}_k	Vetor de variáveis de saída
\mathbf{u}	Vetor de variáveis de entrada
$\tilde{\mathbf{A}}$	Matriz de estados
$\tilde{\mathbf{B}}$	Matriz de entradas
$\tilde{\mathbf{C}}$	Matriz de saídas
$\tilde{\mathbf{D}}$	Matriz de transmissão direta
$\tilde{\mathbf{K}}$	Ganho de Kalman

G	Ganho do observador
ι	Ordem do sistema
w	Quantidade de variáveis de saída
n	Quantidade de variáveis de entrada
Y	Parâmetros de Markov
G	Ganho do observador
Y	Matriz de regressores
$\tilde{\mathcal{P}}$	Matriz de observabilidade
Q	Matriz de controlabilidade
H	Matriz de Hankel
R	Matriz de vetores singulares a esquerda de H
\mathcal{E}	Valores singulares de H
S	Matriz de vetores singulares a direita de H
p	Quantidade de regressores
M	Matriz trajetória
η	Número de linhas de M
S	matriz covariância de M
Φ	Matriz de vetores singulares a esquerda de M
Σ	Autovalores de S
\mathcal{D}	Matriz de vetores singulares a esquerda de M
$\tilde{\epsilon}_k$	Sequência de resíduo da i -ésima regra
o_j	Quantidade de componentes espectrais mais significativas
χ	Fator de seleção de componentes espectrais
Γ	Matriz trajetória decomposta
ψ_k	Vetor de regressores
Ω e Λ	Matrizes de perturbação

ξ	Sequência de ruído estimado
Ξ	Covariância do ruído estimado
\mathfrak{W}	Norma de distância não exponencial
M	Norma de distância exponencial
I_{adp}	Nível de incerteza da série temporal
$\tilde{\mathfrak{R}}$	Matriz de correlação de dados
\tilde{W}^i	Matriz de ponderação nebulosa
AIC	Critério da informação de Akaike
Y^S	Parâmetros de Markov do sistema
Y^O	Parâmetros de Markov do observador
v	vetor de regressores
\tilde{U}^i	Matriz de graus de pertinência adaptado
\mathcal{N}	Distribuição de probabilidade Gaussiana
e_{3D}	Erro 3D-RMS
ACC	Acurácia

Sumário

I	CONTEXTUALIZAÇÃO INICIAL	21
1	INTRODUÇÃO	22
1.1	Trabalhos Relacionados	23
1.2	Motivação e Justificativa	25
1.3	Objetivos	26
1.3.1	Objetivo Geral	26
1.3.2	Objetivos Específicos	26
1.4	Contribuições da Dissertação	26
1.5	Organização da Dissertação	27
1.6	Trabalhos Publicados Pelo Autor	27
1.6.1	Capítulos de Livro	27
1.6.2	Congressos	28
1.6.3	Prêmios	28
II	FUNDAMENTAÇÃO TEÓRICA	29
2	<i>AGRUPAMENTO DE DADOS</i>	30
2.1	Agrupamento Nebuloso Baseado em Máxima Verossimilhança	30
2.1.1	Formulação Matemática	31
2.1.2	Solução Numérica	33
2.1.3	Expoente de Ponderação Nebulosa	35
2.2	Agrupamentos Nebulosos Tipo-2 Intervalares	36
2.2.1	Incerteza no Parâmetro de Ponderação Nebulosa m	38
2.2.2	Extensão para os Conjuntos Nebulosos Tipo-2 Intervalares	39
2.2.3	Análise de Componentes Principais (PCA)	39
3	ALGORITMO ERA/OKID	41
3.1	<i>Observer/Kalman Filter Identification (OKID)</i>	41
3.1.1	Parâmetros de Markov do Observador	43
3.1.2	Parâmetros de Markov do Sistema	45
3.1.3	Parâmetros de Markov do Ganho do Observador	46
3.2	<i>Eigensystem Realization Algorithm (ERA)</i>	47
3.3	Relação Entre o Observador e o Filtro de Kalman	50
4	ANÁLISE ESPECTRAL SINGULAR	52
4.1	Estágio de Decomposição	52

4.1.1	Etapa de Incorporação	52
4.1.2	Etapa de Decomposição em Valores Singulares	53
4.2	Estágio de Reconstrução	54
4.2.1	Etapa de Agrupamento de Índices	54
4.2.2	Etapa de Média da Diagonal	55
4.3	Atualização Recursiva	55
III	REFERENCIAL METODOLÓGICO	57
5	METODOLOGIA PROPOSTA: PROJETO DO FILTRO DE KALMAN NEBULOSO TIPO-2 INTERVALAR	58
5.1	Modelo do Filtro de Kalman Nebuloso Tipo-2 Intervalar	58
5.2	Pré-Processamento de Dados	59
5.2.1	Etapa <i>Offline</i>	60
5.2.2	Etapa <i>online</i>	62
5.2.3	Estimação da Covariância do Ruído de Medição	64
5.3	Estimação Paramétrica do Antecedente	65
5.3.1	Agrupamento Nebuloso Tipo-2 Intervalar de Estimação em Máxima Verossimilhança	66
5.3.2	Modelagem do Parâmetro de Incerteza	68
5.3.3	Classificação da Nova Amostra de Dados	70
5.4	Estimação Paramétrica do Consequente	71
5.4.1	OKID Nebuloso Tipo-2 Intervalar	71
5.4.2	ERA com Seleção Automática de Ordem	76
5.4.3	Atualização Recursiva e Refinada	78
5.4.3.1	Mecanismo de Variação Temporal de Ordem	80
5.5	Simulação do Modelo	82
5.5.1	Simulação um Passo a Frente	82
5.5.2	Simulação Livre	82
IV	RESULTADOS	84
6	ANÁLISE COMPUTACIONAL	85
6.1	Previsão intervalar multivariável de posição e velocidade relativas entre satélites em formação PRISMA	85
6.1.1	Formulação matemática de satélites em formação PRISMA	85
6.1.2	Posição Relativa	89
6.1.2.1	Simulação um Passo a Frente	90
6.1.2.2	Simulação Mista	94
6.1.3	Velocidade Relativa	98

6.1.3.1	Simulação um Passo a Frente	99
6.1.3.2	Simulação Mista	102
7	ANÁLISE EXPERIMENTAL	106
7.1	Rastreamento intervalar da elevação e azimute de um helicóptero 2DoH	106
7.2	Rastreamento e previsão intervalares da trajetória de um veículo suborbital	111
7.2.1	Simulação um passo a frente	113
7.2.2	Simulação Mista	116
V	CONTEXTUALIZAÇÃO FINAL	119
8	CONSIDERAÇÕES FINAIS	120
8.1	Propostas Futuras	121
	REFERÊNCIAS	122

Parte I

CONTEXTUALIZAÇÃO INICIAL

1 Introdução

O conceito de filtragem está presente na sociedade humana há séculos, sendo geralmente associado à remoção de impurezas. Em aplicações na engenharia e controle, uma das possíveis interpretações esta relacionada ao processo de recuperação de um sinal original (ou aproximação deste) através de medições corrompidas (geralmente por ruído), em tempo contínuo ou discreto (ANDERSON; MOORE, 2012). Entre os filtros digitais, uma ferramenta matemática bastante utilizada é o filtro de Kalman (HOSSAIN; HAQUE; ARIF, 2022; CUI et al., 2022; ZHOU et al., 2020; XU et al., 2020). Esta técnica foi introduzida por Kalman (1960) e consiste em um estimador de estado recursivo, que realiza uma estimação estocástica baseada nas medições disponíveis e um modelo previamente determinado. A qualidade desta estimação depende de diversos fatores, entre eles a precisão do modelo disponível, criando assim a necessidade de modelar o sistema da forma mais precisa possível (AGUIRRE, 2007).

O desenvolvimento de modelos matemáticos é um tópico essencial para grande parte das ciências, pois através deste, é possível entender o comportamento de fenômenos e controlá-los. A modelagem matemática pode ser classificada em três grupos de acordo com a forma que o conhecimento ou informações são utilizados: caixa-branca, caixa-preta e caixa-cinza (BABUŠKA, 2012; AGUIRRE, 2007). Quando leis matemáticas ou físicas são utilizadas para descrever o comportamento de um sistema, a modelagem é chamada de caixa-branca. Esse tipo de modelo é interpretável, pois seus parâmetros possuem sentido físico (KEFER et al., 2022; WEI et al., 2021; MOHAN et al., 2010; VELEZ et al., 2021). Quando o conhecimento está na forma de dados ou medições, a modelagem é conhecida como caixa-preta. Neste caso, os modelos são mais difíceis de serem interpretados, pois os parâmetros são selecionados de forma a gerar um modelo que apresente uma resposta ótima (ÅSTRÖM; EYKHOFF, 1971; LJUNG, 1998; SJÖBERG et al., 1995; LOYOLA-GONZALEZ, 2019). Para casos onde tanto as leis físicas quanto os dados observados são utilizados para determinar os modelos, a modelagem é conhecida como caixa-cinza (PINTELAS; LIVIERIS; PINTELAS, 2020; YU et al., 2022; PITCHFORTH et al., 2021). Esta dissertação é focada na modelagem caixa-preta, também conhecida como identificação de sistemas. Este tipo de modelagem se estabelece por meio dos seguintes passos: experimento, seleção da estrutura do modelo, estimação paramétrica e validação do modelo (SERRA, 2005). Durante o experimento, é coletado um conjunto de dados sobre as variáveis de entrada e saída do sistema. Na seleção da estrutura do modelo, uma família de equações matemáticas mais apropriada para modelar o sistema é selecionada. Nesse contexto, existem algumas representações já conhecidas na literatura, tais como *Finite Impulse Response filter* (FIR), *Infinite Impulse Response filter* (IIR), *AutoRegressive with eXogenous input* (ARX),

AutoRegressive with Moving Average and eXogenous input (ARMAX), representações no espaço de estado, entre outras (AGUIRRE, 2007; SARIOĞLU; KURAL, 2015; AVVENTI; LINDQUIST; WAHLBERG, 2012; DING; CHEN, 2005; PEETERS et al., 2022; OBEID; AHMADI; JHA, 2020). Na estimação paramétrica, técnicas de otimização como mínimos quadrados (e suas variantes) (STREJC, 1980; SORENSON, 1970; KARIYA; KURATA, 2004), máxima verossimilhança (MYUNG, 2003; RICHARDS, 1961; ELIASON, 1993) e variável instrumental (SÖDERSTRÖM; STOICA, 2002; BOWDEN; TURKINGTON, 1990; NEWEY; POWELL, 2003) são aplicadas para determinar os valores dos parâmetros da estrutura de modelo selecionada, baseado em um critério de otimização adotado a uma função de custo. Por fim, na etapa de validação, o modelo estimado será testado contra um novo conjunto de dados.

Para a identificação de sistemas no espaço de estados, uma técnica bastante utilizada é o algoritmo *Eigen System Realization Algorithm Observer/Kalman Filter Identification* (OKID). Esta técnica é utilizada para identificar um filtro de Kalman para a modelagem de certos fenômenos (JUANG, 1994). Para lidar com as limitações deste algoritmo, a lógica nebulosa pode ser aplicada, de forma a dividir um problema complexo em vários problemas mais simples (PIRES; SERRA, 2018).

A metodologia proposta nesta dissertação está inserida no contexto do projeto de filtros de Kalman nebulosos tipo-2 intervalares. Para tanto, a metodologia adotada formula o problema através das etapas de pré-processamento, estimação paramétrica do antecedente, e estimação paramétrica do consequente. Na etapa de pré-processamento, a Análise Espectral Singular (AES) é aplicada para reduzir a dependência de ruído na base de dados. Na estimação paramétrica do antecedente, a base de dados tem sua dimensionalidade reduzida pela análise de componentes principais (do inglês *Principal Component Analysis* PCA), e então é classificada em conjuntos nebulosos através da versão tipo-2 intervalar do algoritmo de agrupamento nebuloso de estimação em máxima verossimilhança desenvolvida nesta dissertação. Na estimação paramétrica do consequente, o algoritmo ERA/OKID, baseado no critério de informação de Akaike, é usado para obter a melhor configuração possível ao modelo no espaço de estado (seleção da ordem do sistema). Esta sequência de eventos ocorre durante as etapas de processamento *offline* e *online*. Nesta última é utilizada uma formulação recursiva para estimar os parâmetros do consequente, bem como atualizar a ordem dos modelos locais, na base de regras do Filtro de Kalman nebuloso.

1.1 Trabalhos Relacionados

A aplicação da lógica nebulosa em filtros de Kalman tem sido objeto de estudo em diversos contextos. Em Chen, Xie e Shieh (1998), um sistema intervalar com conhecimento prévio da distribuição de dados é incorporado a um sistema de inferência nebulosa para

gerar uma estimativa escalar a cada iteração. Em [Sasiadek, Wang e Zeremba \(2000\)](#), um sistema de regras nebulosas é aplicado para adaptar as matrizes de covariância do ruído de medição e processo do filtro de Kalman. Em [Simon \(2003\)](#), a filtragem de Kalman é utilizada para projetar um estimador do estado em ambientes ruidosos para modelos nebulosos do tipo Takagi–Sugeno no tempo discreto. Para isso, um filtro de Kalman é alocado em cada modelo linear local, e o filtro global é obtido pela combinação linear dos filtros locais. Em [Madbouly, Abdalla e Banby \(2009\)](#), é desenvolvida uma fusão de dados de múltiplos sensores baseada em um filtro de Kalman adaptativo. Nesta aplicação, as medições de cada sensor alimentam um filtro de Kalman adaptativo nebuloso para estimar as covariâncias dos ruídos. O algoritmo prioriza os sensores mais estáveis e confiáveis. Em [Sung, Lee e You \(2010\)](#), um controlador híbrido que utiliza um observador filtro de Kalman adaptativo foi desenvolvido. No método, para minimizar as diferenças entre as medições e as estimativas, as matrizes de covariância são atualizadas por um observador filtro de Kalman e um controlador nebuloso. Em [Huo et al. \(2014\)](#), uma nova versão adaptativa do filtro de Kalman é proposta. O valor inicial da matriz de covariância do ruído de medição é determinado por um algoritmo evolucionário. Este valor é atualizado durante a operação *online* por um sistema de inferência nebulosa. Em [Páramo-Carranza et al. \(2017\)](#), o filtro de Kalman é combinado com a técnica de modelagem nebulosa tipo Takagi-Sugeno para estender a estimativa linear de estados do filtro de Kalman para situações não lineares. Em [Pires e Serra \(2018\)](#), [Pires e Serra \(2019\)](#), é apresentada uma combinação entre um algoritmo de agrupamento nebuloso evolutivo e o algoritmo OKID para ajustar os parâmetros e estrutura do filtro de Kalman nebuloso através de dados experimentais. Em [Gomes e Serra \(2021b\)](#), [Gomes e Serra \(2021a\)](#), um filtro de Kalman nebuloso tipo-2 intervalar evolutivo é designado para processar componentes espectrais não observáveis oriundas de dados experimentais. A metodologia adotada considera os seguintes passos: desenvolvimento de um modelo inicial para o filtro, o qual é identificado de maneira *offline* através de uma janela de dados experimentais; atualização da proposição do antecedente, a qual é feita pela formulação nebulosa tipo-2 intervalar do agrupamento *evolving* Takagi-Sugeno (eTS); e atualização da proposição do consequente, a qual é realizada pela formulação nebulosa tipo-2 intervalar do algoritmo OKID. Este algoritmo leva em consideração a versão recursiva e multivariável da análise espectral singular.

Diferentemente das abordagens supracitadas, nesta dissertação é proposta uma metodologia para identificação paramétrica de filtros de Kalman através de um algoritmo de agrupamento nebuloso tipo-2 intervalar baseado em critério de máxima verossimilhança junto a Análise de Componentes Principais (ACP). Além disso, a metodologia adota um mecanismo de seleção da quantidade de regressores ideal, baseado no critério da informação de Akaike. Esta adaptação permite que o algoritmo selecione uma ordem diferente para cada submodelo na regra nebulosa e a atualize em função tempo. Ainda, é proposta uma adaptação do expoente de ponderação nebulosa baseada no nível de incerteza dos dados.

Ademais, um esquema de simulação mista é proposto, onde a falta de dados durante um período de tempo, é simulado. Assim, uma simulação livre é realizada e, após esse período, o processo de atualização recursiva e estimação de um passo à frente ocorre normalmente.

1.2 Motivação e Justificativa

O filtro de Kalman tradicional possui duas limitações conhecidas. A primeira delas é com relação a bases de dados não-lineares, pois para este tipo de aplicação, as distribuições de probabilidade não podem ser consideradas como normais. A segunda é a dependência da filtragem ao modelo disponível para propagar as estimativas (AGUIRRE, 2007). Para a estimação de estado em problemas não lineares, pode-se destacar duas técnicas clássicas conhecidas na literatura, as quais são o filtro de Kalman Estendido e o filtro de Kalman *Unscented*. Na versão estendida do filtro de Kalman, a matriz de covariância do erro de estimação é propagada através de uma linearização para aproximar o modelo não-linear disponível. Porém, esta aproximação pode introduzir grandes erros nas covariâncias de estados posteriores (WAN; MERWE, 2000). Já o filtro de Kalman *Unscented* realiza estimativas sobre a média e covariância através de uma amostragem sobre certa janela de dados (UMAMAGESWARI; IGNATIOUS; VINODHA, 2012). Estas técnicas buscam resolver o problema relativo a não-linearidade inerente ao sistema dinâmico objeto de estudo, porém, não levam em consideração a estimação paramétrica do modelo. Assim, visando melhorar a qualidade do filtro de Kalman, algumas técnicas de inteligência computacional e aprendizagem de máquina foram aplicadas, visando uma estimação paramétrica ótima, inclusive no espaço de estados, como se pode observar nas referências (PIRES; SERRA, 2018; PIREs; SERRA, 2019; GOMES; SERRA, 2021b; GOMES; SERRA, 2021a). Tais referências são aplicações nebulosas do algoritmo OKID, e possuem margens factíveis para adaptações. O primeiro ponto é o fato de que em tais pesquisas, apenas o algoritmo de agrupamento nebuloso Gustafson-Kessel (tipo-1 e tipo-2 intervalar) foi aplicado, enquanto em Babuška (2012), Rocha Filho e de Oliveira Serra (2018) é expresso que o agrupamento nebuloso baseado em máxima verossimilhança, por não possuir restrição volumétrica, consegue se adaptar melhor a conjuntos com diferentes volumes e formatos. Além disso, essas pesquisas não exploraram técnicas de redução da dimensionalidade para melhorar a partição de dados. O segundo ponto, é o fato de que os modelos no espaço de estado são obtidos através de uma representação ARX, na qual a quantidade de termos regressores é definida pelo projetista. Porém, a influência deste parâmetro à resposta do modelo ainda não foi discutida. Além disso, espera-se que quanto maior o número de regressores, maior será a ordem do sistema. O terceiro ponto, é o fato de que em estimação recursiva, é importante garantir que os parâmetros não variem mais rápido do que os estados (BARSHALOM; LI; KIRUBARAJAN, 2001), sendo necessário o desenvolvimento de uma versão refinada para o algoritmo OKID. Por fim, as pesquisas utilizam um intervalo fixo para

o expoente de ponderação nebulosa, o que torna viável a busca por um mecanismo de adaptação deste intervalo de acordo com o grau de incerteza inerente aos dados.

1.3 Objetivos

1.3.1 Objetivo Geral

Desenvolver uma metodologia para estimação paramétrica recursiva de filtros de Kalman nebulosos tipo-2 intervalares, baseada em critério de máxima verossimilhança e componentes espectrais não-observáveis.

1.3.2 Objetivos Específicos

- Apresentar o embasamento teórico acerca da identificação paramétrica de filtros de Kalman baseado no algoritmo OKID e AES no contexto dos sistemas nebulosos tipo-2 intervalares.
- Formular matematicamente uma metodologia para estimação paramétrica recursiva e refinada de filtros de Kalman adaptativos nebulosos tipo-2 intervalar, baseado em critério de máxima verossimilhança.
- Validar a metodologia proposta com resultados computacionais com dados simulados acerca satélites da formação PRISMA, e dados experimentais obtidos de um helicóptero com dois graus de liberdade e um veículo suborbital VS-30.

1.4 Contribuições da Dissertação

As principais contribuições que demonstram a originalidade desta pesquisa, são:

- Desenvolver um mecanismo de adaptação para a AES recursiva de forma a reduzir o número de operações redundantes e, conseqüentemente, o custo computacional da aplicação.
- Formular matematicamente uma versão nebulosa tipo-2 intervalar do algoritmo de agrupamento nebuloso de máxima verossimilhança, que supere limitações relacionadas à inicialização e às indeterminações matemáticas.
- Formular um mecanismo para seleção e atualização da ordem de cada submodelo no espaço de estados, na base de regras do filtro de Kalman nebuloso tipo-2 intervalar.
- Formular matematicamente um mecanismo para atualização do expoente de ponderação nebulosa baseada na estimação da covariância do ruído via AES.

- A exploração de um esquema de simulação mista, onde num mesmo conjunto de teste, tanto a estimativa livre como a um passo a frente ocorram.

1.5 Organização da Dissertação

Esta dissertação está dividida em 5 partes: Contextualização Inicial, Fundamentação Teórica, Referencial metodológico, Resultados e Contextualização Final.

A parte **I**, CONTEXTUALIZAÇÃO INICIAL, contém apenas o capítulo **1**. Neste são apresentados os tópicos de interesse desta dissertação, bem como a motivação, justificativa, objetivos e contribuições presentes na metodologia proposta. Por fim, os trabalhos publicados pelo autor foram apresentados.

A parte **II**, FUNDAMENTAÇÃO TEÓRICA, contém os capítulos **2**, **3** e **4**. No capítulo **2**, conceitos acerca de algoritmos de agrupamentos nebulosos são apresentados. No capítulo **3**, os conceitos acerca do algoritmo ERA/OKID são apresentados. No capítulo **4**, os conceitos e formulações da análise espectral singular são apresentados. Tais conceitos constituem a base para a metodologia desenvolvida nesta dissertação.

A parte **III**, REFERENCIAL METODOLÓGICO, contém o capítulo **5**. Neste, A metodologia para projeto de filtro de Kalman nebuloso tipo-2 intervalar proposta nesta dissertação é apresentada. Neste capítulo, são abordadas as etapas de pré-processamento de dados e identificação paramétrica do antecedente e consequente.

A parte **IV**, RESULTADOS, contém os capítulos **6** e **7**. No capítulo **6**, a metodologia desenvolvida foi testada em ambientes computacionais referente à posição e velocidade relativa entre os satélites da formação PRISMA. No capítulo **7**, a metodologia desenvolvida é testada com dados experimentais referentes a elevação e azimute de um helicóptero com dois graus de liberdade, e da trajetória do veículo suborbital VS-30.

A parte **V**, CONTEXTUALIZAÇÃO FINAL, contém o capítulo **8**. Este, apresenta uma seção de comentários finais acerca da dissertação, bem como uma seção contendo propostas para futuras implementações.

1.6 Trabalhos Publicados Pelo Autor

1.6.1 Capítulos de Livro

- MONTEL, Ben-Hur Matthews Moreno; DE OLIVEIRA SERRA, Ginalber Luiz. A Data-Driven Interval Type-2 Fuzzy Kalman Filter of Minimum Realization for Forecasting Spacecraft Formation on Low Earth Orbit. In: Global Perspectives on Robotics and Autonomous Systems: Development and Applications. IGI Global,

2023. p. 93-129.

1.6.2 Congressos

- MONTEL, B. H. M. M.; SERRA, G. L. O. An Interval Type-2 Maximum Likelihood Fuzzy Clustering Algorithm. In: Anais do XIX Encontro Nacional de Inteligência Artificial e Computacional (pp. 473-484). SBC, 2022. Campinas - SP.
- MONTEL, B. H. M. M.; SERRA, G. L. O. Projeto de Filtro de Kalman Nebuloso Tipo-2 Intervalar Baseado em Máxima Verossimilhança para a Previsão de Séries Temporais Multivariáveis. In: proceedings of XVI Brazilian Conference on Computational Intelligence, CBIC, 2023. Salvador - BA.

1.6.3 Prêmios

- O artigo intitulado “Projeto de Filtro de Kalman Nebuloso Tipo-2 Intervalar Baseado em Máxima Verossimilhança para a Previsão de Séries Temporais Multivariáveis” recebeu o prêmio de melhor artigo da área *Fuzzy and Stochastic Systems* do XVI Congresso Brasileiro de Inteligência Computacional (CBIC-2023).

-

Parte II

FUNDAMENTAÇÃO TEÓRICA

2 AGRUPAMENTO DE DADOS

Algoritmos de agrupamento são técnicas de classificação de dados que os agrupam de acordo com certos critérios de similaridade. Assim, os dados que pertencem ao mesmo conjunto são mais similares do que aqueles pertencentes a conjuntos distintos (ROKACH; MAIMON, 2005; MADHULATHA, 2012; KRIEGEL et al., 2011). Os algoritmos de agrupamento estão entre as técnicas de classificação não supervisionada de dados, pois não precisam de nenhuma classe informada a priori (GHAHRAMANI, 2003; DAYAN; SAHANI; DEBACK, 1999). Entre estes algoritmos, existem as técnicas baseadas na lógica nebulosa, como os algoritmos Gustafson-Kessel (GK) e o agrupamento nebuloso de máxima verossimilhança. Tais algoritmos são baseados em uma métrica de distância adaptada pela covariância dos dados (WUNSCH; XU, 2008; BABUŠKA, 2012). Com a introdução dos sistemas de lógica nebulosa tipo-2 intervalares (LIANG; MENDEL, 2000), foram propostas várias técnicas de classificação de dados para essa nova versão dos conjuntos nebulosos (HWANG; RHEE, 2007; QIU et al., 2013; WU; LIU, 2016; YIN; SHENG; QIN, 2022). Neste capítulo, a teoria dos agrupamentos nebulosos baseados em máxima verossimilhança, e dos agrupamentos nebulosos tipo-2 intervalares, os quais são fundamentais ao entendimento desta pesquisa, serão apresentados e discutidos.

2.1 Agrupamento Nebuloso Baseado em Máxima Verossimilhança

Os algoritmos de agrupamento nebulosos são aplicados à classificação de conjuntos de dados amostrados, sejam eles quantitativos, qualitativos, ou uma mistura de ambos (OLIVEIRA; PEDRYCZ, 2007). Usualmente, estes dados estão armazenados em uma matriz de dados (\mathbf{Z}), onde cada linha representa uma variável:

$$\mathbf{Z} = [\mathbf{z}_1 \quad \mathbf{z}_2 \quad \cdots \quad \mathbf{z}_N], \quad \mathbf{Z} \in \mathbb{R}^{n \times N}, \quad (2.1)$$

onde n é o número de variáveis, N é a quantidade de observações, e $\mathbf{z}_N \in \mathbb{R}^n$ é o vetor de dados que contém a N -ésima medição. O objetivo do algoritmo é particionar a base de dados em c conjuntos nebulosos distintos, resultando na seguinte matriz de partição:

$$\mathbf{U} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1N} \\ \cdots & \cdots & \ddots & \cdots \\ \mu_{c1} & \mu_{c2} & \cdots & \mu_{cN} \end{bmatrix}, \quad (2.2)$$

onde $\mathbf{U} \in \mathbb{R}^{c \times N}$ é matriz de partição nebulosa, e μ_{cN} é o grau de pertinência da N -ésima medição em relação ao conjunto c . Por ser uma partição nebulosa, a matriz \mathbf{U} deve possuir as seguintes propriedades:

$$\mu_{ik} \in [0, 1], \quad 1 \leq i \leq c, \quad 1 \leq k \leq N, \quad (2.3)$$

$$\sum_{i=1}^c \mu_{ik} = 1, \quad 1 \leq k \leq N, \quad (2.4)$$

$$0 \leq \sum_{k=1}^N \mu_{ik} < N, \quad 1 \leq i \leq c. \quad (2.5)$$

A equação (2.3) determina que os graus de pertinência são normalizados. A equação (2.4) mostra que os conjuntos são complementares. Já a equação (2.5) define que entre os conjuntos estimados, não existe um conjunto vazio, nem conjunto universo (BABUŠKA, 2012; PEDRYCZ; GOMIDE, 2007).

2.1.1 Formulação Matemática

Na literatura, os dois algoritmos de agrupamento nebuloso mais conhecidos são o *Fuzzy C-means* (FCM) e o Gustafson-Kessel (NAYAK; NAIK; BEHERA, 2015; ASKARI, 2021; LU et al., 2013; BAS; EGRIOGLU, 2022; DOVŽAN; ŠKRJANC, 2011; BABUKA; VEEN; KAYMAK, 2002). Ambos, são algoritmos de agrupamento baseados na seguinte função objetivo:

$$J(\mathbf{Z}; \mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m \| \mathbf{z}_k - \mathbf{v}_i \|_S^2 \quad (2.6)$$

onde $\mathbf{v}_i \in \mathbb{R}^n$ é o centro do i -ésimo conjunto, $\mathbf{V} \in \mathbb{R}^{n \times c}$ é a matriz que contém os centros de todos os conjuntos, e $\| \cdot \|_S$ é um operador que representa uma norma de distância. O algoritmo FCM possui uma norma de distância não adaptativa, mantendo o formato de sua partição constante para todos os conjuntos (GREKOUSIS; THOMAS, 2012). Já o algoritmo GK, possui uma restrição volumétrica em sua formulação, assim, suas partições não se adaptam a certos cenários (BABUŠKA, 2012; GUSTAFSON; KESSEL, 1979). Nesse contexto, em Trauwaert, Kaufman e Rousseeuw (1991), foi proposto um algoritmo de agrupamento nebuloso baseado em máxima verossimilhança, que é capaz de detectar conjuntos de vários formatos, tamanhos e densidades.

O algoritmo de Agrupamento Nebuloso de Estimacão de Máxima Verossimilhança (ANEMV) tem como objetivo determinar suas partições de forma que o conjunto de dados \mathbf{Z} seja, no universo de discurso, os pontos mais prováveis a serem observados (BEZDEK; DUNN, 1975). A figura 1 apresenta duas imagens que representam esse processo. Na primeira imagem, é apresentada a distribuição de probabilidade dos dados pertencentes ao universo de discurso serem selecionados por certo conjunto. Já na segunda imagem, as curvas de nível desta distribuição são apresentadas, junto com a base de dados e os centros de cada conjunto.

Para formular matematicamente o processo de otimização, o primeiro ponto é definir a função de custo. Para o ANEMV, é utilizada a seguinte função de verossimilhança:

$$L(\mathbf{Z}, \mathbf{v}_i, \mathbf{S}_i) = \sum_{k=1}^N \sum_{i=1}^c P_i g(\mathbf{z}_k | \mathbf{v}_i; \mathbf{S}_i), \quad (2.7)$$

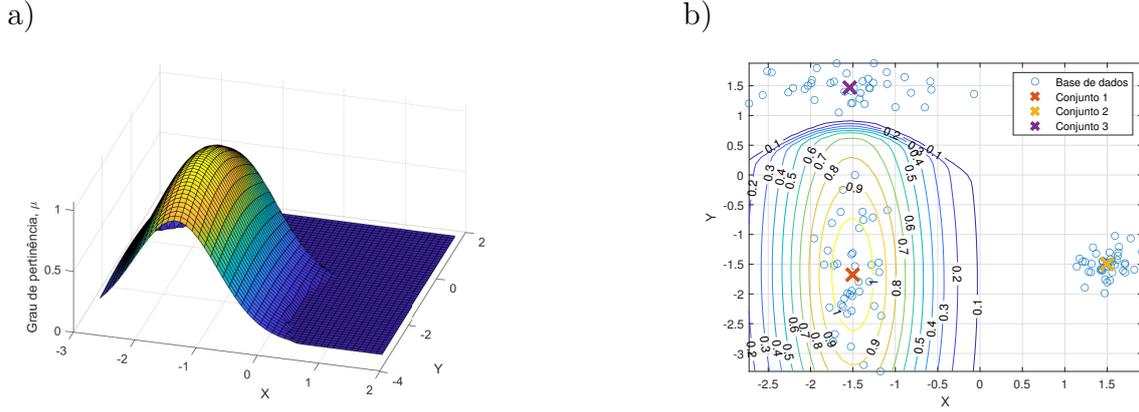


Figura 1 – Probabilidade de um dado ser selecionado dentro do universo de discurso. a) Densidade de probabilidade b) Curvas de nível.

onde P_i é a probabilidade a priori do conjunto i ser selecionado, e $g(\mathbf{z}_k|\mathbf{v}_i; \mathbf{S}_i)$ é a probabilidade da medição \mathbf{z}_k ser escolhida, dado a média \mathbf{v}_i e a covariância do conjunto $\mathbf{S}_i \in \mathbb{R}^{n \times n}$. Tal probabilidade condicional é expressa por:

$$g(\mathbf{Z}|\mathbf{v}_i; \mathbf{S}_i) = \frac{[\det \mathbf{S}_i]^{-1/2}}{2\pi^{n/2}} \exp[-0.5(\mathbf{Z} - \mathbf{v}_i)^T \mathbf{S}_i^{-1}(\mathbf{Z} - \mathbf{v}_i)]. \quad (2.8)$$

Por ser uma função exponencial, é mais simples otimizá-la através do seu logaritmo natural (BEZDEK; DUNN, 1975). Assim, o processo de maximizar a função de verossimilhança é expresso por:

$$\max_{\mathbf{v}_i, \mathbf{S}_i} \ln L(\mathbf{Z}, \mathbf{v}_i, \mathbf{S}_i). \quad (2.9)$$

Esta equação é uma mistura gaussiana com todos os parâmetros desconhecidos (P_i , \mathbf{v}_i e \mathbf{S}_i). Nesse sentido, em Duda, Hart et al. (1973), foi demonstrado que a estimativa de tais parâmetros para um ponto de máximo local da função $L(\mathbf{Z}, \mathbf{v}_i, \mathbf{S}_i)$ tem o seguinte formato:

$$P_i = \frac{1}{N} \sum_{k=1}^N p_i(i|\mathbf{z}_k), \quad (2.10)$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^N p_i(i|\mathbf{z}_k) \mathbf{z}_k}{\sum_{k=1}^N p_i(i|\mathbf{z}_k)}, \quad (2.11)$$

$$\mathbf{S}_i = \frac{\sum_{k=1}^N p_i(i|\mathbf{z}_k) (\mathbf{z}_k - \mathbf{v}_i)(\mathbf{z}_k - \mathbf{v}_i)^T}{\sum_{k=1}^N \mu_{ik}}, \quad (2.12)$$

$$p_i(i|\mathbf{z}_k) = \frac{\det \mathbf{S}_i^{-0.5} \exp[0.5(\mathbf{z}_k - \mathbf{v}_i)^T \mathbf{S}_i^{-1}(\mathbf{z}_k - \mathbf{v}_i)] P_i}{\sum_{j=1}^c \det \mathbf{S}_j^{-0.5} \exp[0.5(\mathbf{z}_k - \mathbf{v}_j)^T \mathbf{S}_j^{-1}(\mathbf{z}_k - \mathbf{v}_j)] P_j}, \quad (2.13)$$

onde $p_i(i|\mathbf{z}_k)$ é a probabilidade do conjunto i ser selecionado dado a medição \mathbf{z}_k . Essa probabilidade pode ser interpretada como o grau de pertinência da medição \mathbf{z}_k em relação ao conjunto i (μ_{ik}).

2.1.2 Solução Numérica

A solução analítica para a equação (2.9) é complicada de se obter, pois envolve a determinação de funções de densidade não triviais. As iterações de Picard são um método computacional popular para otimizar essa equação (ROBIN, 2010; LIU et al., 2008). O algoritmo de ANEMV pode ser implementado através dos passos apresentados na sequência.

Para inicializar o método, deve-se definir a base de dados utilizada (\mathbf{Z}), o número de conjuntos desejados ($1 < c < N$), um valor para o expoente de ponderação nebulosa ($m > 1$), o critério de tolerância (ϵ) e uma matriz de partição nebulosa inicial ($\mathbf{U}^{(0)}$). Os próximos passos serão repetidos até que o critério de convergência seja ativado.

O primeiro passo consiste em utilizar a matriz de partição disponível para calcular a estimativa para o centro dos conjuntos (\mathbf{v}_i^l), o que é demonstrado por:

$$\mathbf{v}_i^l = \frac{\sum_{k=1}^N \mu_{ik} \mathbf{z}_k}{\sum_{k=1}^N \mu_{ik}}, \quad 1 \leq i \leq c. \quad (2.14)$$

O segundo passo consiste em utilizar estas informações para calcular a matriz de covariância do agrupamento (\mathbf{F}^i) e sua probabilidade a priori (P_i):

$$\mathbf{F}^i = \frac{\sum_{k=1}^N \mu_{ik} (\mathbf{z}_k - \mathbf{v}_i) (\mathbf{z}_k - \mathbf{v}_i)^T}{\sum_{k=1}^N \mu_{ik}}, \quad 1 \leq i \leq c, \quad (2.15)$$

$$P_i = \frac{1}{N} \sum_{k=1}^N \mu_{ik}, \quad 1 \leq i \leq c. \quad (2.16)$$

O terceiro passo consiste em calcular as normas de distância de cada dado para o centro de cada conjunto:

$$\mathcal{D}_k^i = \frac{[\det \mathbf{F}^i]^{0.5}}{P_i} \exp[0.5(\mathbf{z}_k - \mathbf{v}_i^l)^T \mathbf{F}^{i-1} (\mathbf{z}_k - \mathbf{v}_i^l)], \quad 1 \leq i \leq c, \quad 1 \leq k \leq N. \quad (2.17)$$

Pode-se observar que a equação 2.17 é similar a uma função densidade de probabilidade. Assim, o algoritmo usa como métrica de similaridade uma função probabilística, o que está de acordo com o seu objetivo. O quarto passo consiste em atualizar a matriz de partição:

$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^c \left(\frac{\mathcal{D}_k^i}{\mathcal{D}_k^j}\right)^{2/(m-1)}}. \quad (2.18)$$

O algoritmo de ANEMV necessita de uma boa inicialização, pois ele tende a encontrar um ótimo local próximo a matriz de partição inicial. O algoritmo GK é tido como uma boa inicialização para este método (BABUŠKA, 2012; Rocha Filho; de Oliveira Serra, 2018).

Diferentemente do algoritmo GK, o ANEMV não possui restrição volumétrica. Neste algoritmo, a probabilidade a priori funciona como uma métrica de densidade do

conjunto. Logo, suas partições se adaptam a conjuntos de diferentes volumes dentro da mesma base de dados. Nesse sentido, as figuras 2, 3 e 4 ilustram a diferença entre os formatos das partições obtidas pelos algoritmos de agrupamento nebulosos FCM, GK e ANEMV.

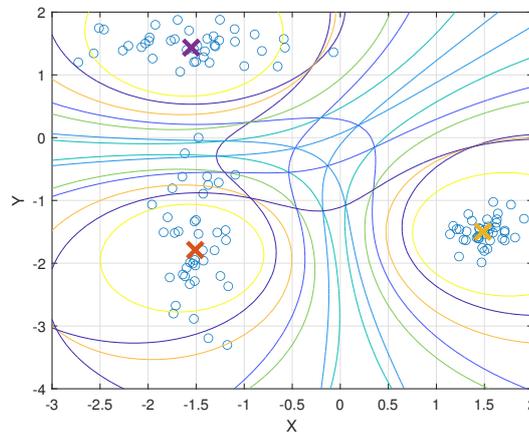


Figura 2 – Partição obtida pelo algoritmo de agrupamento FCM. O conjunto de dados foi gerado pelo autor.

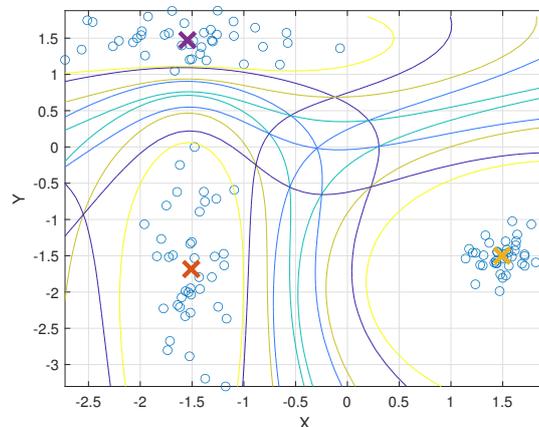


Figura 3 – Partição obtida pelo algoritmo de agrupamento GK. O conjunto de dados foi gerado pelo autor.

Na figura 2, pode-se observar que o resultado obtido pelo agrupamento FCM mantém um padrão comum para todos os conjuntos estimados, isso ocorre pela não adaptação da norma por uma matriz de covariância. Na figura 3, pode-se observar que o formato das partições variou em cada conjunto, porém, devido a restrição volumétrica, o formato do conjunto menos disperso ficou mais abrangente que o ideal. Já na figura 4, pode-se observar que o formato das partições se adaptou melhor ao conjunto de dados disponível, tanto em formato, quanto em volume.

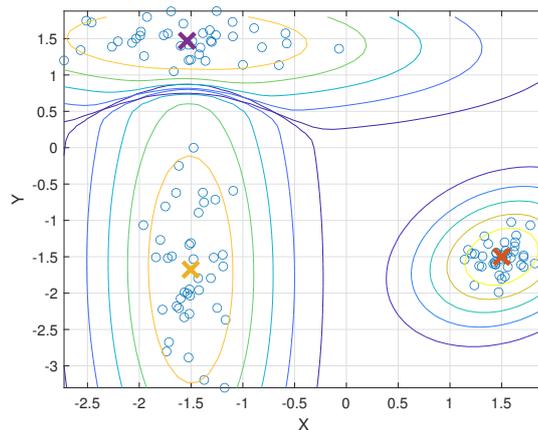


Figura 4 – Partição obtida pelo algoritmo de agrupamento ANEMV. O conjunto de dados foi gerado pelo autor.

2.1.3 Expoente de Ponderação Nebulosa

Entre os parâmetros presente nos algoritmos de agrupamento nebulosos, o expoente de ponderação (m) é um dos mais importantes. Ele determina o quão nebulosa ou incerta é a partição resultante (PEI et al., 2001). Quando m se aproxima de 1, a partição tende a um formato determinístico (*Hard*). Quando m tende a infinito, a solução converge para o centro de massa da base de dados (YU; CHENG; HUANG, 2004; BABUŠKA, 2012; BEZDEK; EHRLICH; FULL, 1984).

Em geral, a seleção do expoente m é feita pelo projetista. Geralmente, é recomendado o uso de $m = 2$ para partições nebulosas (BABUŠKA, 2012; BEZDEK, 1993). No entanto, alguns estudos acerca da seleção do expoente de ponderação nebulosa para o algoritmo de agrupamento FCM foram desenvolvidos (YU; CHENG; HUANG, 2004; JING; DENG; YU, 2014; PAL; BEZDEK, 1995; PEI et al., 2001). Os trabalhos citados apresentam intervalos para a seleção do parâmetro m . Em Pal e Bezdek (1995), é sugerido que a melhor escolha para m está provavelmente dentro do intervalo $[1.5, 2.5]$. Em Pei et al. (2001), é apresentado que o valor de m está dentro do intervalo $[1.5, 3.5]$, com $m = 2.5$ sendo considerado o valor provável para aplicações práticas.

A Figura 5 mostra as partições obtidas pelo algoritmo ANEMV para diferentes valores do expoente de ponderação nebulosa. Para $m = 2$, a partição está bem próxima de um formato determinístico. Para $m = 2.1$ ou $m = 2.5$, pode-se observar partições nebulosas que se adaptam ao formato esperado. Para $m = 10$, os centros dos conjuntos tendem ao centro de massa da base de dados.

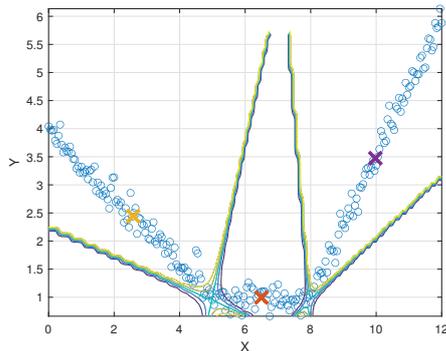
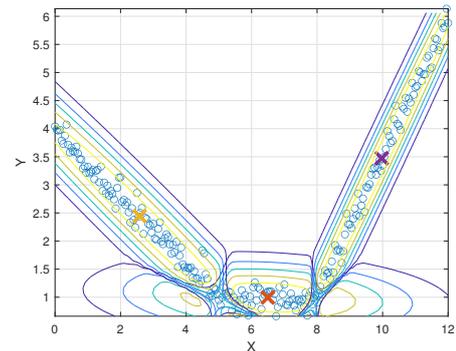
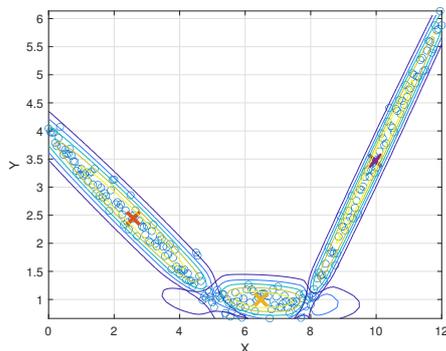
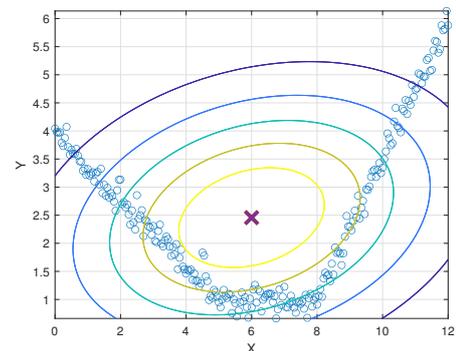
a) ANEMV com $m = 2$ b) ANEMV com $m = 2.1$ c) ANEMV com $m = 2.5$ d) ANEMV com $m = 10$ 

Figura 5 – Partição realizada pelo algoritmo ANEMV com diferentes valores para o expoente de ponderação nebulosa. As curvas de nível representam o grau de pertinência em cada ponto da base de dados. A cor amarela representa valores próximos a 1, enquanto que as linhas roxas representam graus de pertinência tendendo a 0.

2.2 Agrupamentos Nebulosos Tipo-2 Intervalares

Os conjunto nebulosos tipo-2 são uma extensão do conceito tradicional de conjuntos nebulosos, sendo caracterizados por uma incerteza adicionada ao formato das funções de pertinência. Um conjunto nebuloso tipo-2 é definido por um grau de pertinência primário μ_{ik} , que define o quanto este dado está de acordo com as características do conjunto, e um grau de pertinência secundário $\mu_i(k, \mu_{ik})$, que indica a incerteza sobre as características do conjunto (LIANG; MENDEL, 2000; KARNIK; MENDEL, 1998). Nesse contexto, a figura 6 mostra um conjunto nebuloso tipo-2 e a projeção de seu grau de pertinência secundário. Nesta, pode-se observar que a função de pertinência não possui um formato fixo, mas sim uma faixa de valores com diferentes pertinências associadas, criando assim uma segunda ordem de incerteza.

A partir da figura 6, pode-se constatar que o conjunto nebuloso tipo-2 mapeia todo o espaço formado pelo universo de discurso x e o grau de pertinência primário. Logo, realizar operações lógicas com estes conjuntos é computacionalmente custoso (LIANG; MENDEL, 2000). Por tal motivo, foi introduzido o conceito dos conjuntos nebulosos

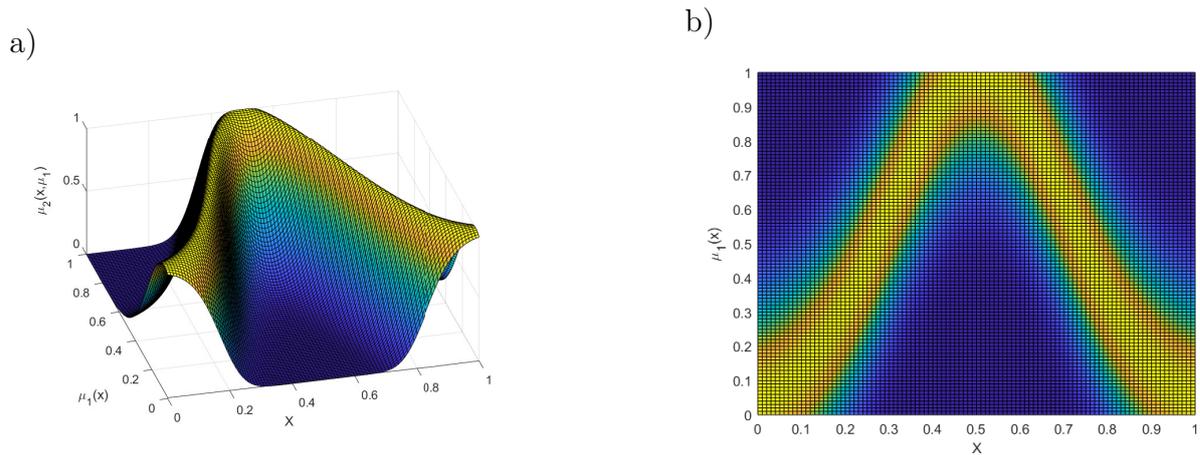


Figura 6 – Conjunto nebuloso tipo-2 com formato Gaussiano. a) Representação tridimensional do conjunto, apresentando os graus de pertinência primário e secundário. b) Projeção do grau de pertinência secundário no espaço $[x, \mu_{ik}]$, a cor amarela representa um grau de pertinência secundário próximo a 1, enquanto a cor azul representa uma pertinência próxima a zero.

tipo-2 intervalares, onde seus graus de pertinência secundários podem assumir apenas dois valores, zero ou um, criando assim uma região de incerteza. Tal adaptação reduz o custo computacional, pois agora, as operações nebulosas podem ser realizadas através da aritmética intervalar. Assim, os conjuntos passam a ser definidos pelos limites superior e inferior das funções de pertinência, como apresentado na figura 7 (LIANG; MENDEL, 2000; MENDEL; JOHN; LIU, 2006; WU, 2012). Com o surgimento deste tipo de conjuntos, pode-se observar na literatura o desenvolvimento de alguns algoritmos de agrupamento nebuloso tipo-2 intervalares e suas aplicações (HWANG; RHEE, 2007; QIU et al., 2013; YIN; SHENG; QIN, 2022; XU et al., 2019). Para gerar a região de incerteza, tais trabalhos utilizam o expoente de ponderação nebulosa m .

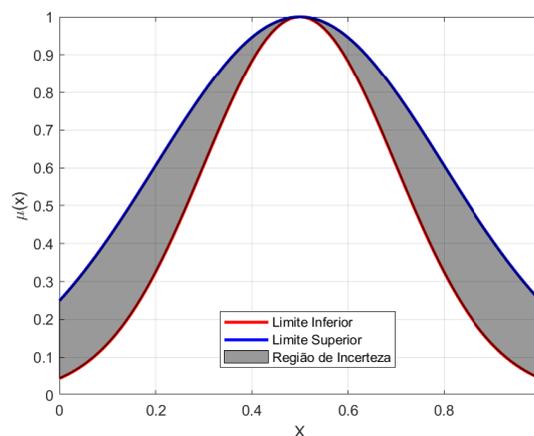


Figura 7 – Função de pertinência para conjunto nebuloso tipo-2 Intervalar. A região sombreada representa os pontos onde o grau de pertinência secundário é igual a um.

2.2.1 Incerteza no Parâmetro de Ponderação Nebulosa m

Como demonstrado na seção 2.1.3, o expoente de ponderação nebulosa m interfere no formato das partições. E, uma vez que tais formatos são uma representação dos graus de pertinência, pode-se inferir que o parâmetro m influencia no formato das funções de pertinência. Isso ocorre, pois em geral, os agrupamentos nebulosos baseados na função objetivo *c-means*, possuem os parâmetros apresentados pelas equações (2.14)-(2.18) dependentes (direta ou indiretamente) do expoente m . Assim, ao escolher diferentes valores para m , pode-se gerar funções de pertinência com transições mais ríspidas ou mais suaves, como ilustrado na figura 8 (HWANG; RHEE, 2007).

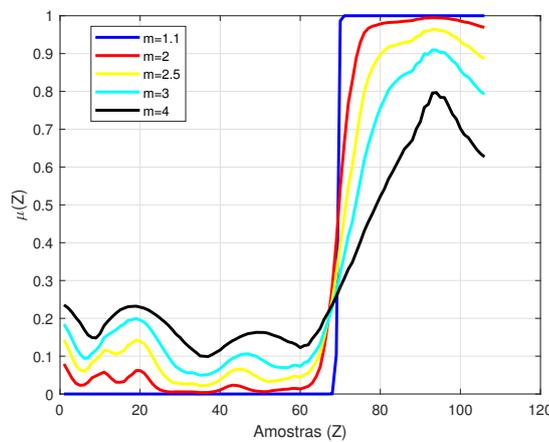


Figura 8 – Funções de pertinência geradas por um algoritmo de agrupamento nebuloso com diferentes valores para m .

Alguns estudos foram realizados com o objetivo de encontrar um valor ideal para m . Em Pei et al. (2001), constatou-se que, para o algoritmo FCM, existe um valor ótimo de m aos conjuntos de dados específicos. Porém, este valor varia de acordo com a separabilidade do conjunto de dados. Os testes realizados mostraram que o valor ótimo está dentro do intervalo $m \in [1.5, 3.5]$. Nesse sentido, é mais interessante realizar o agrupamento nebuloso num intervalo do expoente m , do que procurar por um valor ótimo para este expoente (LINDA; MANIC, 2012). Nesse contexto, o expoente de ponderação nebulosa passa a ser representado por um intervalo, como demonstrado a seguir:

$$\tilde{m} = [m^-, m^+], \quad (2.19)$$

onde \tilde{m} é o intervalo do expoente de ponderação nebulosa, m^- é o limite inferior deste intervalo e m^+ é o limite superior. Dessa forma, \tilde{m} passa a caracterizar duas ordens de incerteza.

2.2.2 Extensão para os Conjuntos Nebulosos Tipo-2 Intervalares

A incerteza de segunda ordem proposta pela lógica nebulosa tipo-2 intervalar pode ser modelada através do expoente de ponderação nebulosa intervalar (\tilde{m}) (RUBIO; CASTILLO, 2013; NGO; MAI; PEDRYCZ, 2015; KALHORI; ZARANDI, 2015; CHOI; RHEE, 2009). Assim, o objetivo dos algoritmos de agrupamento nebulosos tipo-2 intervalares é determinar uma matriz de partição intervalar, dada por:

$$\tilde{\mathbf{U}} = [\mathbf{U}^-, \mathbf{U}^+] = \begin{bmatrix} \tilde{\mu}_{11} & \tilde{\mu}_{12} & \cdots & \tilde{\mu}_{1N} \\ \cdots & \cdots & \ddots & \cdots \\ \tilde{\mu}_{c1} & \tilde{\mu}_{c2} & \cdots & \tilde{\mu}_{cN} \end{bmatrix}, \quad (2.20)$$

$$\tilde{\mu}_{ik} = [\mu_{ik}^-, \mu_{ik}^+], \quad (2.21)$$

onde $\tilde{\mu}_{ik}$ são os graus de pertinência intervalares presentes na matriz de partição $\tilde{\mathbf{U}}$. Desse modo, para obter valores ótimos para estes parâmetros, deve-se otimizar a versão intervalar da função objetivo *c-means*, expressa por:

$$\tilde{J}(\tilde{\mathbf{U}}, \tilde{\mathbf{v}}) = [\mathbf{U}^-, \mathbf{U}^+] = \sum_{i=1}^c \sum_{k=1}^N \tilde{\mu}_{ik} \| \mathbf{z}_k - \tilde{\mathbf{v}}_i \|^2, \quad (2.22)$$

onde $\tilde{\mathbf{v}}_i$ é o centro intervalar do conjunto i . A otimização desta função deve ser restrita a gerar uma matriz de partição de tipo reduzido que siga as propriedades apresentadas nas equações (2.3)-(2.5). Este processo é realizado através dos multiplicadores de Lagrange, encontrando o seguinte formato para os parâmetros desejados:

$$\tilde{\mathbf{v}}_i = \frac{\sum_{k=1}^N \tilde{\mu}_{ik}^{\tilde{m}} \mathbf{z}_k}{\sum_{k=1}^N \tilde{\mu}_{ik}^{\tilde{m}}}, \quad 1 \leq i \leq c, \quad (2.23)$$

$$\mu_{ik}^{(l)-} = \min \left[\frac{1}{\sum_{j=1}^c \left(\frac{D_j^-(z_k)}{D_j^+(z_k)} \right)^{2/(m^- - 1)}}, \frac{1}{\sum_{j=1}^c \left(\frac{D_j^+(z_k)}{D_j^-(z_k)} \right)^{2/(m^+ - 1)}} \right], \quad (2.24)$$

$$\mu_{ik}^{(l)+} = \max \left[\frac{1}{\sum_{j=1}^c \left(\frac{D_j^-(z_k)}{D_j^+(z_k)} \right)^{2/(m^- - 1)}}, \frac{1}{\sum_{j=1}^c \left(\frac{D_j^+(z_k)}{D_j^-(z_k)} \right)^{2/(m^+ - 1)}} \right]. \quad (2.25)$$

Para a solução numérica deste processo, basta utilizar as iterações de Picard, de modo que, de posse de uma matriz de partição inicial ($\tilde{\mathbf{U}}^{(0)}$), é possível calcular os centros intervalares dos conjuntos e utilizá-los para atualizar a matriz de partição. Este processo deve ser repetido até que não haja variação significativa nos termos da matriz de partição, seguindo o mesmo molde apresentado nas equações (2.14)-(2.18), porém levando em consideração que as variáveis são intervalares.

2.2.3 Análise de Componentes Principais (PCA)

A análise de componentes principais é uma técnica que visa reduzir a dimensionalidade de uma base de dados (JOLLIFFE, 2002; BRO; SMILDE, 2014; GREENACRE et al., 2022). Para realizar esta tarefa, deve-se implementar os passos apresentados na sequência.

Passo 1- Definir a matriz de dados, a qual é expressa por:

$$\mathbf{Z} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nN} \\ y_{11}^a & y_{12}^p & \cdots & y_{1N}^a \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1}^a & y_{m2}^a & \cdots & y_{mN}^a \end{bmatrix} = [\mathbf{Z}_1 \quad \mathbf{Z}_2 \quad \cdots \quad \mathbf{Z}_N] \in \mathbb{R}^{n+m \times N}, \quad (2.26)$$

onde u_{jk} é a k -ésima medição da j -ésima variável de entrada, e y_{jk}^a é k -ésimo valor processado da j -ésima variável de saída (valor obtido pela AES).

Passo 2- Calcular o ponto médio da base de dados:

$$\bar{\mathbf{Z}} = \frac{1}{N} \sum_{j=1}^N \mathbf{Z}_j, \quad (2.27)$$

onde $\bar{\mathbf{Z}} \in \mathbb{R}^{w+m}$ é o vetor que contém as coordenadas do ponto médio de \mathbf{Z} no \mathbb{R}^{w+n} .

Passo 3- Calcular a variância amostral do conjunto de dados:

$$\mathbf{C}_Z = \sum_{j=1}^N [\mathbf{Z}_j - \bar{\mathbf{Z}}][\mathbf{Z}_j - \bar{\mathbf{Z}}]^T. \quad (2.28)$$

Diferente da formulação tradicional da variância amostral, a equação (2.28) não é multiplicada pelo termo $\frac{1}{N-1}$. Isso ocorre, pois os autovetores não são alterados por essa relação. Sendo assim, torna-se conveniente trabalhar com esta representação (JOLLIFFE, 2002).

Passo 4- Obter os autovalores e autovetores de \mathbf{C}_Z :

$$[\mathbf{U}, \mathbf{V}] = eig(\mathbf{C}_Z), \quad (2.29)$$

onde \mathbf{U} são os autovetores de \mathbf{C}_Z , e \mathbf{V} são seus respectivos autovalores. Estes devem ser ordenados em ordem decrescente de significância.

Passo 5- Selecionar as componentes que juntas retêm pelo menos 99% da variação do conjunto original. Esta quantidade é determinada pelo menor índice q , no qual a seguinte expressão é verdadeira:

$$\frac{\sum_{r=1}^q \mathbf{v}_r}{\sum_{j=1}^{m+n} \mathbf{v}_j} \times 100\% \geq 99\%, \quad (2.30)$$

onde \mathbf{v}_i é o i -ésimo autovalor mais significativo de \mathbf{C}_Z .

Passo 6- Calcular a matriz de componentes principais \mathbf{Z}^{cp} :

$$\mathbf{Z}^{cp} = \mathbf{Z}^T \mathbf{U}^q, \quad \in \mathbb{R}^q \quad (2.31)$$

onde \mathbf{U}^q é a matriz que contém os primeiros q autovetores de \mathbf{C}_Z . Esta mesma equação pode ser usada para reduzir a dimensão das novas amostras de dados durante a etapa *online*. Para tanto, basta substituir a matriz de dados pelo vetor \mathbf{Z}_k .

3 Algoritmo ERA/OKID

A filtragem de Kalman é um método para estimar o estado de um sistema dinâmico a partir de medições ruidosas. O método funciona combinando uma estimativa a priori do estado com as medições disponíveis. A estimativa a priori é obtida a partir das equações de movimento do sistema, das medições anteriores e de uma estimativa inicial do estado. A qualidade do modelo é um fator importante que afeta a precisão da filtragem de Kalman. Se o modelo for incorreto, a estimação também será incorreta (AGUIRRE, 2007). Neste sentido, algumas técnicas para a identificação de sistemas no espaço de estado foram desenvolvidas (MOOR; OVERSCHEE; FAVOREEL, 1999; YU; LJUNG; VERHAEGEN, 2018; JOHANSSON et al., 2000; VERHAEGEN, 1994; VALASEK; CHEN, 2003; SOARES; BOTTURA, 2011). Entre estas técnicas, pode-se destacar o algoritmo ERA/OKID, o qual realiza uma identificação de modelo no espaço de estado junto ao seu ganho de Kalman (JUANG, 1994). Neste capítulo, o embasamento teórico, assim como, os passos do algoritmo ERA/OKID serão apresentados, conforme descrito por (JUANG, 1994).

3.1 Observer/Kalman Filter Identification (OKID)

O algoritmo OKID é utilizado para a identificação paramétrica dos parâmetros de Markov de um sistema linear discreto no espaço de estado. Tais sistemas são geralmente descritos por:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad (3.1)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k, \quad (3.2)$$

onde $\mathbf{x}_k \in \mathbb{R}^\iota$ é o vetor de estados, ι é a ordem do sistema, $\mathbf{y}_k \in \mathbb{R}^w$ é o vetor de saídas do sistema, w é o número de saídas, $\mathbf{u}_k \in \mathbb{R}^n$ é o vetor de entradas do sistema, n é o número de entradas, e $\mathbf{A} \in \mathbb{R}^{\iota \times \iota}$, $\mathbf{B} \in \mathbb{R}^{\iota \times n}$, $\mathbf{C} \in \mathbb{R}^{w \times \iota}$ e $\mathbf{D} \in \mathbb{R}^{w \times n}$ são as matrizes do sistema.

Para realizar uma identificação paramétrica no espaço de estado, deve-se determinar as matrizes do sistema de acordo com algum critério de otimalidade relacionado ao erro de estimação. Porém, as únicas medições disponíveis são a respeito das variáveis de entrada e saída, e para determinar todos os parâmetros fazem-se necessárias as informações sobre os estados. Tal limitação pode ser solucionada se for considerado que os sistemas partem de uma condição inicial nula $x(0) = 0$, assim é possível escrever uma equação para os estados em função das entradas passadas:

$$\begin{aligned}
\mathbf{x}_0 &= 0, \\
\mathbf{x}_1 &= \mathbf{B}\mathbf{u}_0, \\
\mathbf{x}_2 &= \mathbf{A}\mathbf{B}\mathbf{u}_0 + \mathbf{B}\mathbf{u}_1, \\
\mathbf{x}_3 &= \mathbf{A}^2\mathbf{B}\mathbf{u}_0 + \mathbf{A}\mathbf{B}\mathbf{u}_1 + \mathbf{B}\mathbf{u}_2, \\
&\vdots \\
\mathbf{x}_{N-1} &= \sum_{i=1}^{N-1} \mathbf{A}^{(i-1)}\mathbf{B}\mathbf{u}_{N-i-1},
\end{aligned} \tag{3.3}$$

onde N é o número de medições realizadas. Como a equação (3.2) é uma função dos estados e das variáveis de entrada, é possível reescrevê-la em função apenas das variáveis de entrada:

$$\mathbf{y}_{N-1} = \sum_{i=1}^{N-1} \mathbf{C}\mathbf{A}^{(i-1)}\mathbf{B}\mathbf{u}_{N-i-1} + \mathbf{D}\mathbf{u}_{N-1}. \tag{3.4}$$

Assim, é possível agrupar esta equação para o seguinte formato matricial:

$$\mathbf{y} = \mathbf{Y}\mathbf{L}, \tag{3.5}$$

onde

$$\mathbf{y} = [\mathbf{y}_0 \ \mathbf{y}_1 \ \mathbf{y}_2 \ \cdots \ \mathbf{y}_{N-1}], \tag{3.6}$$

$$\mathbf{Y} = [\mathbf{D} \ \mathbf{C}\mathbf{B} \ \mathbf{C}\mathbf{A}\mathbf{B} \ \cdots \ \mathbf{C}\mathbf{A}^{(N-2)}\mathbf{B}], \tag{3.7}$$

$$\mathbf{L} = \begin{bmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_{N-1} \\ 0 & \mathbf{u}_0 & \mathbf{u}_1 & \cdots & \mathbf{u}_{N-2} \\ 0 & 0 & \mathbf{u}_0 & \cdots & \mathbf{u}_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{u}_0 \end{bmatrix}, \tag{3.8}$$

onde $\mathbf{Y} \in \mathbb{R}^{w \times nN}$ são os parâmetros de Markov do sistema e $\mathbf{L} \in \mathbb{R}^{nN \times N}$ é a matriz que contém os dados de entrada do sistema. Através das equações (3.5) e (3.6), pode-se constatar que existem $w \times nN$ parâmetros desconhecidos e $w \times N$ observações. Logo, a solução da equação (3.5) resultaria num sistema subdeterminado (múltiplas soluções).

Ao considerar que \mathbf{A} é assintoticamente estável, ou seja, $\mathbf{A}^k \approx 0$, para $k \geq p$, a quantidade de termos desconhecidos pode ser reduzida, como segue:

$$\begin{aligned}
\mathbf{Y} &= [\mathbf{D} \ \mathbf{C}\mathbf{B} \ \mathbf{C}\mathbf{A}\mathbf{B} \ \cdots \ \mathbf{C}\mathbf{A}^{(p-1)}\mathbf{B} \ 0 \ \cdots \ 0] \\
\mathbf{Y}_T &= [\mathbf{D} \ \mathbf{C}\mathbf{B} \ \mathbf{C}\mathbf{A}\mathbf{B} \ \cdots \ \mathbf{C}\mathbf{A}^{(p-1)}\mathbf{B}] \in \mathbb{R}^{w \times np}.
\end{aligned} \tag{3.9}$$

Assim, a equação (3.5) pode ser reescrita como:

$$\mathbf{y} \approx \mathbf{Y}_T\mathbf{L}_T, \tag{3.10}$$

onde \mathbf{L}_T é a versão reduzida da matriz de entrada, expressa por:

$$\mathbf{L}_T = \begin{bmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_p & \cdots & \mathbf{u}_{N-1} \\ 0 & \mathbf{u}_0 & \mathbf{u}_1 & \cdots & \mathbf{u}_{p-1} & \cdots & \mathbf{u}_{N-2} \\ 0 & 0 & \mathbf{u}_0 & \cdots & \mathbf{u}_{p-2} & \cdots & \mathbf{u}_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{u}_0 & \cdots & \mathbf{u}_{N-p-1} \end{bmatrix}. \quad (3.11)$$

Pode-se observar que na equação (3.10) existem $m \times r(p+1)$ parâmetros desconhecidos e $m \times N$ observações. Assim, ao escolher um valor para a quantidade de regressores $p < (N/r) - 1$, é possível obter uma estimativa para \mathbf{Y}_T através da pseudo inversa de \mathbf{L}_T , de forma que o erro de aproximação decaia à medida que p aumenta. Porém, altos valores de p tornam impraticável a solução numérica para \mathbf{Y}_T .

3.1.1 Parâmetros de Markov do Observador

À medida que o número de regressores p aumenta, o erro de aproximação do sistema identificado tende a decair. Porém, a dimensão da matriz $[\mathbf{L}\mathbf{L}^T]$, necessária para a obtenção da solução para \mathbf{Y}_T aumenta, tornando o processo cada vez mais complexo. Uma forma de aumentar a taxa de amortecimento de forma a propiciar a escolha de um número baixo de regressores é através de uma realimentação de sistemas obtida por uma observação de estados. Isso ocorre através da adição de termos $\mathbf{G}\mathbf{y}_k$ na equação (3.1), obtendo:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{G}\mathbf{y}_k - \mathbf{G}\mathbf{y}_k \\ &= (\mathbf{A} + \mathbf{G}\mathbf{C})\mathbf{x}_k + (\mathbf{B} + \mathbf{G}\mathbf{D})\mathbf{u}_k - \mathbf{G}\mathbf{y}_k \\ &= \bar{\mathbf{A}}\mathbf{x}_k + \bar{\mathbf{B}}\mathbf{v}_k, \end{aligned} \quad (3.12)$$

onde $\mathbf{G} \in \mathbb{R}^{l \times w}$ é o ganho do observador, e $\bar{\mathbf{A}}$, $\bar{\mathbf{B}}$ e \mathbf{v}_k são definidos como:

$$\begin{aligned} \bar{\mathbf{A}} &= \mathbf{A} + \mathbf{G}\mathbf{C}, \\ \bar{\mathbf{B}} &= [\mathbf{B} + \mathbf{G}\mathbf{D}, -\mathbf{G}], \\ \mathbf{v}_k &= [\mathbf{u}_k, \mathbf{y}_k]^T. \end{aligned} \quad (3.13)$$

Pode-se observar que a equação (3.10) foi modelada assumindo condição inicial nula, o que pode não ocorrer para todos os casos. Assim, pode-se obter uma equação neste formato para o observador de estado descrito em (3.12), através de:

$$\begin{aligned} \mathbf{x}_{k+1} &= \bar{\mathbf{A}}\mathbf{x}_k + \bar{\mathbf{B}}\mathbf{v}_k, \\ \mathbf{x}_{k+2} &= \bar{\mathbf{A}}\mathbf{x}_{k+1} + \bar{\mathbf{B}}\mathbf{v}_k \\ &= \bar{\mathbf{A}}^2\mathbf{x}_k + \bar{\mathbf{A}}\bar{\mathbf{B}}\mathbf{v}_k + \bar{\mathbf{B}}\mathbf{v}_{k+1}, \\ \mathbf{x}_{k+3} &= \bar{\mathbf{A}}\mathbf{x}_{k+2} + \bar{\mathbf{B}}\mathbf{v}_{k+2} \\ &= \bar{\mathbf{A}}^3\mathbf{x}_k + \bar{\mathbf{A}}^2\bar{\mathbf{B}}\mathbf{v}_k + \bar{\mathbf{A}}\bar{\mathbf{B}}\mathbf{v}_{k+1} + \bar{\mathbf{B}}\mathbf{v}_{k+2}, \\ \mathbf{x}_{k+p} &= \bar{\mathbf{A}}^p\mathbf{x}_k + \sum_{j=1}^p \bar{\mathbf{A}}^{p-j}\bar{\mathbf{B}}\mathbf{v}_{k-j+1}. \end{aligned} \quad (3.14)$$

Aplicando-se este resultado na equação de medições 3.2, têm-se:

$$\begin{aligned} \mathbf{y}_{k+p} &= \mathbf{C}\mathbf{x}_{k+p} + \mathbf{D}\mathbf{u}_{k+p} \\ &= \mathbf{C}\bar{\mathbf{A}}^p\mathbf{x}_k + \sum_{j=1}^p \mathbf{C}\bar{\mathbf{A}}^{p-j}\mathbf{v}_{k-j+1} + \mathbf{D}\mathbf{u}_{k+p}. \end{aligned} \quad (3.15)$$

Esta equação pode ser escrita na forma matricial, da seguinte forma:

$$\bar{\mathbf{y}} = \mathbf{C}\bar{\mathbf{A}}^p\mathbf{x} + \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}}. \quad (3.16)$$

Na seção anterior, foi considerado que \mathbf{A} é assintoticamente estável, assim como para $\bar{\mathbf{A}}$. Logo, o termo $\mathbf{C}\bar{\mathbf{A}}\mathbf{x}$ pode ser negligenciado, resultando em:

$$\bar{\mathbf{y}} = \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}}, \quad (3.17)$$

onde:

$$\bar{\mathbf{y}} = [\mathbf{y}_p \quad \mathbf{y}_{p+1} \quad \cdots \quad \mathbf{y}_{l-1}], \quad (3.18)$$

$$\bar{\mathbf{y}} = [\mathbf{x}_0 \quad \mathbf{x}_1 \quad \cdots \quad \mathbf{x}_{l-p-2}], \quad (3.19)$$

$$\bar{\mathbf{Y}}_T = [\mathbf{D} \quad \mathbf{C}\bar{\mathbf{B}} \quad \mathbf{C}\bar{\mathbf{A}}\bar{\mathbf{B}} \quad \cdots \quad \mathbf{C}\bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}}], \quad (3.20)$$

$$\bar{\mathbf{\Gamma}} = \begin{bmatrix} \mathbf{u}_p & \mathbf{u}_{p+1} & \cdots & \mathbf{u}_{l-1} \\ \mathbf{v}_{p-1} & \mathbf{v}_p & \cdots & \mathbf{v}_{l-2} \\ \mathbf{v}_{p-2} & \mathbf{v}_{p-1} & \cdots & \mathbf{v}_{l-3} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_0 & \mathbf{v}_1 & \cdots & \mathbf{v}_{l-p-1} \end{bmatrix}. \quad (3.21)$$

Observa-se que na equação (3.17), o objetivo é obter os parâmetros de Markov do observador ($\bar{\mathbf{Y}}_T$). Para isso, pode-se utilizar a solução em batelada através do método de mínimos quadrados, como apresentado a seguir:

$$\begin{aligned} \min f_{ob} &= [\mathbf{y} - \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}}][\mathbf{y} - \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}}]^T \\ \frac{\partial [\mathbf{y} - \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}}][\mathbf{y} - \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}}]^T}{\partial \bar{\mathbf{Y}}_T} &= 0 \\ 2(\mathbf{y} - \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}})\bar{\mathbf{V}}^T &= 0 \\ \mathbf{y}\bar{\mathbf{\Gamma}}^T - \bar{\mathbf{\Gamma}}_T\bar{\mathbf{\Gamma}}\bar{\mathbf{\Gamma}}^T &= 0 \\ \bar{\mathbf{Y}}_T\bar{\mathbf{\Gamma}}\bar{\mathbf{\Gamma}}^T &= \mathbf{y}\bar{\mathbf{\Gamma}}^T \\ \bar{\mathbf{Y}}_T &= \mathbf{y}\bar{\mathbf{\Gamma}}^T[\bar{\mathbf{\Gamma}}\bar{\mathbf{\Gamma}}^T]^{-1}. \end{aligned} \quad (3.22)$$

A solução para $\bar{\mathbf{Y}}_T$ obtida em (3.22) é válida para os parâmetros de Markov do observador, ou seja, tal matriz, possui elementos tanto das matrizes do sistema como do ganho do observador adicionado em (3.12). Sendo assim necessário a separação destes termos.

3.1.2 Parâmetros de Markov do Sistema

Como apresentado na equação (3.20), os parâmetros de Markov do observador são calculados em função das matrizes $\bar{\mathbf{A}}$ e $\bar{\mathbf{B}}$. Estes, como apresentado na equação (3.13), possuem influência da realimentação. Portanto, para extrair as matrizes de estado do sistema, deve-se separar os parâmetros de Markov do observador em parâmetros do sistema (\mathbf{Y}_T) e do ganho do observador (\mathbf{Y}_O). Para isso, deve-se agrupar os parâmetros de Markov, da seguinte forma:

$$\bar{\mathbf{Y}}_T = [\bar{\mathbf{Y}}_0 \quad \bar{\mathbf{Y}}_1 \quad \bar{\mathbf{Y}}_2 \quad \cdots \quad \bar{\mathbf{Y}}_p], \quad (3.23)$$

onde

$$\begin{aligned} \bar{\mathbf{Y}}_0 &= \mathbf{D} \\ \bar{\mathbf{Y}}_k &= \mathbf{C}\bar{\mathbf{A}}^{k-1}\bar{\mathbf{B}}, \\ &= [\mathbf{C}(\mathbf{A} + \mathbf{G}\mathbf{C})^{k-1}(\mathbf{B} + \mathbf{G}\mathbf{D}) \quad -\mathbf{C}(\mathbf{A} + \mathbf{G}\mathbf{C})^{k-1}\mathbf{G}] \\ &\triangleq [\bar{\mathbf{Y}}_k^{(1)} \quad -\bar{\mathbf{Y}}_k^{(2)}], \quad k = 1, 2, 3, \dots, p. \end{aligned} \quad (3.24)$$

Para obter os parâmetros do sistema a partir desta representação, deve-se recuperar a sequência $\mathbf{C}\bar{\mathbf{A}}^{k-1}\bar{\mathbf{B}}$ a partir da expressão de $\bar{\mathbf{Y}}_k^{(1)}$. Como esta expressão é dependente de uma potência, deve-se observar seu formato para diferentes valores de k . Assim, para $k = 1$, tem-se:

$$\begin{aligned} \bar{\mathbf{Y}}_1^{(1)} &= \mathbf{C}(\mathbf{A} + \mathbf{G}\mathbf{C})^0(\mathbf{B} + \mathbf{G}\mathbf{D}) \\ &= \mathbf{C}\mathbf{B} + \mathbf{C}\mathbf{G}\mathbf{D}. \end{aligned} \quad (3.25)$$

Porém, sabe-se que $\mathbf{Y}_T^{(1)} = \mathbf{C}\mathbf{B}$ e $\bar{\mathbf{Y}}_1^{(2)} = \mathbf{C}\mathbf{G}$. Portanto, aplicando-se estas definições na equação acima, tem-se:

$$\begin{aligned} \bar{\mathbf{Y}}_1^{(1)} &= \mathbf{Y}_T^{(1)} + \bar{\mathbf{Y}}_1^{(2)}\mathbf{D} \\ \mathbf{Y}_T^{(1)} &= \bar{\mathbf{Y}}_1^{(1)} - \bar{\mathbf{Y}}_1^{(2)}\mathbf{D}. \end{aligned} \quad (3.26)$$

Aplicando-se o mesmo procedimento para $k = 2, 3$, obtêm-se as seguintes expressões:

$$\begin{aligned} \mathbf{Y}_T^{(2)} &= \bar{\mathbf{Y}}_2^{(1)} - \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_T^{(1)} - \bar{\mathbf{Y}}_2^{(2)}\mathbf{D} \\ \mathbf{Y}_T^{(3)} &= \bar{\mathbf{Y}}_3^{(1)} - \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_T^{(2)} - \bar{\mathbf{Y}}_2^{(2)}\mathbf{Y}_T^{(1)} - \bar{\mathbf{Y}}_3^{(2)}\mathbf{D}. \end{aligned} \quad (3.27)$$

Nesse sentido, é possível formular uma equação para $\mathbf{Y}_T^{(k)}$ através do seguinte somatório:

$$\mathbf{Y}_T^{(k)} = \bar{\mathbf{Y}}_k^{(1)} - \sum_{i=1}^k \bar{\mathbf{Y}}_i^{(2)}\mathbf{Y}_T^{(k-i)}, \quad k = 1, \dots, p. \quad (3.28)$$

Porém, para obter as matrizes de estado, o algoritmo ERA monta uma matriz de Hankel com os parâmetros de Markov do sistema. Portanto, é necessário a obtenção de termos além dos p parâmetros gerados pela equação (3.28). Para obter estes termos adicionais,

poderíamos gerar os termos necessários através de $\bar{\mathbf{Y}}_k = \mathbf{C}\bar{\mathbf{A}}^{k-1}\bar{\mathbf{B}}$. Sabe-se que \mathbf{A} é assintoticamente estável, e $\bar{\mathbf{Y}}_k = 0$ para $k > p$. Assim, o somatório na equação (3.28) passa a gerar parcelas nulas a partir de um certo ponto. A expressão para os termos dos parâmetros de Markov do sistema com $k > p$ pode ser modelada da seguinte forma:

$$\begin{aligned} \mathbf{Y}_T^{(k)} &= \mathbf{C}\bar{\mathbf{A}}^k\bar{\mathbf{B}} - \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_T^{(k)} - \dots - \bar{\mathbf{Y}}_p^{(2)}\mathbf{Y}_T^{(k-p)} - \mathbf{C}\bar{\mathbf{A}}^p\bar{\mathbf{B}}\mathbf{Y}_T^{(k-p)} - \dots - \mathbf{C}\bar{\mathbf{A}}^k\bar{\mathbf{B}}\mathbf{Y}_T^{(0)} \\ \mathbf{Y}_T^{(k)} &= 0 - \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_T^{(k)} - \dots - \bar{\mathbf{Y}}_p^{(2)}\mathbf{Y}_T^{(k-p)} - 0 - \dots - 0 \\ \mathbf{Y}_T^{(k)} &= \sum_{i=1}^p \bar{\mathbf{Y}}_i^{(2)}\mathbf{Y}_T^{k-i}, \quad k = p+1, \dots \end{aligned} \quad (3.29)$$

A quantidade necessária de parâmetros de Markov do sistema depende da dimensão escolhida para a matriz de Hankel no algoritmo ERA.

3.1.3 Parâmetros de Markov do Ganho do Observador

Para obter os parâmetros de Markov, deve-se primeiro observar que a expressão para $\bar{\mathbf{Y}}_k^{(2)}$ pode ser escrita no seguinte formato:

$$\bar{\mathbf{Y}}_k^{(2)} = \mathbf{C}[\mathbf{A} + \mathbf{GC}]^{k-1}\mathbf{G} = \mathbf{C}\bar{\mathbf{A}}\mathbf{G}. \quad (3.30)$$

Logo, para obter os parâmetros do ganho do observador, deve-se extrair a sequência $\mathbf{Y}^O = \mathbf{C}\bar{\mathbf{A}}^{k-1}\mathbf{G}$ da expressão apresentada na equação (3.30). Para isso, deve-se observar as expressões obtidas para diferentes valores de k . Nesse sentido, para $k = 1$, tem-se o caso mais simples, apresentado por:

$$\bar{\mathbf{Y}}_1^{(2)} = \mathbf{C}\mathbf{G} = \mathbf{C}\bar{\mathbf{A}}^0\mathbf{G} = \mathbf{Y}_1^O. \quad (3.31)$$

O próximo termo da sequência ($k = 2$), é obtido através das seguintes associações:

$$\begin{aligned} \bar{\mathbf{Y}}_2^{(2)} &= \mathbf{C}[\mathbf{A} + \mathbf{GC}]\mathbf{G} \\ \bar{\mathbf{Y}}_2^{(2)} &= \mathbf{C}\bar{\mathbf{A}}\mathbf{G} + \mathbf{C}\mathbf{G}\mathbf{C}\mathbf{G} \\ \bar{\mathbf{Y}}_2^{(2)} &= \mathbf{Y}_2^O + \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_1^O \\ \mathbf{Y}_2^O &= \bar{\mathbf{Y}}_2^{(2)} - \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_1^O. \end{aligned} \quad (3.32)$$

De forma similar, o termo para $k = 3$ é definido como:

$$\begin{aligned} \bar{\mathbf{Y}}_3^{(2)} &= \mathbf{C}[\mathbf{A} + \mathbf{GC}][\mathbf{A} + \mathbf{GC}]\mathbf{G} \\ \bar{\mathbf{Y}}_3^{(2)} &= \mathbf{C}[\mathbf{A}^2 + \mathbf{C}\bar{\mathbf{A}}\mathbf{G}\mathbf{C} + \mathbf{C}\mathbf{G}\mathbf{C}\bar{\mathbf{A}}\mathbf{G} + \mathbf{C}\mathbf{G}\mathbf{C}\mathbf{G}\mathbf{C}\mathbf{G}]\mathbf{G}, \\ \bar{\mathbf{Y}}_3^{(2)} &= \mathbf{C}\bar{\mathbf{A}}^2\mathbf{G} + \mathbf{C}\mathbf{G}\mathbf{C}\bar{\mathbf{A}}\mathbf{G} + \mathbf{C}\bar{\mathbf{A}}\mathbf{G}\mathbf{C}\mathbf{G} + \mathbf{C}\mathbf{G}\mathbf{C}\mathbf{G}\mathbf{C}\mathbf{G} \\ \bar{\mathbf{Y}}_3^{(2)} &= \mathbf{C}\bar{\mathbf{A}}^2\mathbf{G} + \mathbf{C}\mathbf{G}\mathbf{C}\bar{\mathbf{A}}\mathbf{G} + \mathbf{C}[\mathbf{A} + \mathbf{GC}]\mathbf{G}\mathbf{C}\mathbf{G} \\ \bar{\mathbf{Y}}_3^{(2)} &= \mathbf{Y}_3^O + \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_2^O + \bar{\mathbf{Y}}_2^{(2)}\mathbf{Y}_1^O \\ \mathbf{Y}_3^O &= \bar{\mathbf{Y}}_3^{(2)} - \bar{\mathbf{Y}}_1^{(2)}\mathbf{Y}_2^O - \bar{\mathbf{Y}}_2^{(2)}\mathbf{Y}_1^O. \end{aligned} \quad (3.33)$$

De posse destes termos, é possível obter uma expressão para \mathbf{Y}_k^O de forma análoga à obtenção da expressão para os parâmetros do sistema. Assim, a expressão para os parâmetros do ganho do observador é dada por:

$$\begin{aligned}\mathbf{Y}_1^O &= \mathbf{C}\mathbf{G} = \bar{\mathbf{Y}}_1^{(2)} \\ \mathbf{Y}_k^O &= \bar{\mathbf{Y}}_k^{(2)} - \sum_{i=1}^{k-1} \bar{\mathbf{Y}}_i^{(2)} \mathbf{Y}_{k-i}^O, \quad k = 2, \dots, p \\ \mathbf{Y}_k^O &= - \sum_{i=1}^p \bar{\mathbf{Y}}_i^{(2)} \mathbf{Y}_{k-i}^O, \quad k = p+1, \dots\end{aligned}\tag{3.34}$$

A partir desta representação, é possível obter o ganho do observador, para isso, deve-se observar que a sequência obtida na equação acima pode ser escrita em função da matriz de observabilidade:

$$\mathbf{Y}^O = \begin{bmatrix} \mathbf{C}\mathbf{G} \\ \mathbf{C}\mathbf{A}\mathbf{G} \\ \vdots \\ \mathbf{C}\mathbf{A}^k\mathbf{G} \end{bmatrix} = \tilde{\mathbf{P}}\mathbf{G},\tag{3.35}$$

onde $\tilde{\mathbf{P}}$ é a matriz de observabilidade. Assim, ao aplicar o método de mínimos quadrados, obtêm-se:

$$\mathbf{G} = [\tilde{\mathbf{P}}^T \tilde{\mathbf{P}}]^{-1} \tilde{\mathbf{P}}^T \mathbf{Y}^O,\tag{3.36}$$

A matriz de observabilidade é dependente de k termos. Porém, esta possui posto igual a ordem do sistema identificado.

3.2 Eigensystem Realization Algorithm (ERA)

O algoritmo ERA é uma técnica de estimação paramétrica no espaço de estados. Nele, o objetivo é obter as matrizes A, B e C do modelo através dos parâmetros de Markov do sistema. Para isso, a seguinte matriz de Hankel é analisada:

$$\mathbf{H}(k-1) = \begin{bmatrix} \mathbf{Y}_T^{(k)} & \mathbf{Y}_T^{(k+1)} & \dots & \mathbf{Y}_T^{(k+\beta-1)} \\ \mathbf{Y}_T^{(k+1)} & \mathbf{Y}_T^{(k+2)} & \dots & \mathbf{Y}_T^{(k+\beta)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Y}_T^{(k+\alpha-1)} & \mathbf{Y}_T^{(k+\alpha)} & \dots & \mathbf{Y}_T^{(k+\beta+\alpha-2)} \end{bmatrix},\tag{3.37}$$

onde α e β são inteiros maiores do que a ordem do sistema (n). Sabe-se que os parâmetros de Markov do sistema foram obtidos de forma a estimar a seguinte sequência:

$$\mathbf{Y}_T^{(k)} = \mathbf{C}\mathbf{A}^{k-1}\mathbf{B}.\tag{3.38}$$

Logo, a expressão da matriz de Hankel para $k = 1$ pode ser escrita como:

$$\mathbf{H}(0) = \begin{bmatrix} \mathbf{CB} & \mathbf{CAB} & \cdots & \mathbf{CA}^\beta \mathbf{B} \\ \mathbf{CAB} & \mathbf{CA}^2 \mathbf{B} & \cdots & \mathbf{CA}^{\beta+1} \mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{(\alpha)} \mathbf{B} & \mathbf{CA}^{(\alpha+1)} \mathbf{B} & \cdots & \mathbf{CA}^{(\beta+\alpha-1)} \mathbf{B} \end{bmatrix}. \quad (3.39)$$

É possível modelar a matriz $\mathbf{H}(0)$ em função das matrizes de controlabilidade (\mathbf{Q}) e observabilidade ($\tilde{\mathbf{P}}$), da seguinte forma:

$$\mathbf{H}(0) = \tilde{\mathbf{P}}\mathbf{Q}, \quad (3.40)$$

onde:

$$\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^\alpha \end{bmatrix}, \quad (3.41)$$

$$\mathbf{Q} = [\mathbf{B} \quad \mathbf{AB} \quad \cdots \quad \mathbf{A}^\beta \mathbf{B}], \quad (3.42)$$

e é possível estimar as matrizes \mathbf{B} e \mathbf{C} através de uma relação entre a matriz de Hankel e as matrizes de controlabilidade e observabilidade. Tal relação pode ser obtida através da decomposição do valor singular de $\mathbf{H}(0)$, expressa por:

$$\mathbf{H}(0) = \mathbf{R}\boldsymbol{\mathcal{E}}\mathbf{S}^T, \quad (3.43)$$

onde $\mathbf{R} \in \mathbb{R}^{\alpha \times \alpha}$ e $\mathbf{S} \in \mathbb{R}^{\beta \times \beta}$ são matrizes ortogonais e $\boldsymbol{\mathcal{E}}$ é uma matriz diagonal contendo os autovalores de $\mathbf{H}(0)$. A partir das propriedades dessas matrizes, pode-se escrever uma pseudo-inversa para $\mathbf{H}(0)$, da seguinte forma:

$$\mathbf{H}^\dagger = \mathbf{S}\boldsymbol{\mathcal{E}}^{-1}\mathbf{R}^T. \quad (3.44)$$

Os autovalores da decomposição podem ser utilizados para determinar a ordem do sistema. Este procedimento é feito através da seguinte análise acerca da significância dos autovalores:

$$\boldsymbol{\epsilon}_1 > \boldsymbol{\epsilon}_2 > \cdots > \boldsymbol{\epsilon}_\iota > \epsilon > \boldsymbol{\epsilon}_\iota \cdots > \boldsymbol{\epsilon}_{\min(\alpha, \beta)}, \quad (3.45)$$

onde $\boldsymbol{\epsilon}$ são os autovalores de $\mathbf{H}(0)$ e ι é a ordem do sistema. Esta é utilizada para truncar a matriz $\mathbf{H}(0)$, da seguinte forma:

$$\mathbf{H}_\iota(0) = \mathbf{R}_\iota \boldsymbol{\mathcal{E}}_\iota \mathbf{S}_\iota^T, \quad (3.46)$$

onde o subscrito ι representa que as dimensões desta matriz foram reduzidas para a realização da ordem mínima. Observa-se que a versão truncada das matrizes de

observabilidade e controlabilidade, resultantes de $\mathbf{H}_l(0)$ se reduzem as matrizes \mathbf{C} e \mathbf{B} , respectivamente.

Através da comparação entre as equações (3.40) e (3.46), pode-se observar que $\tilde{\mathcal{P}}_l$ está relacionado à \mathbf{R}_l , enquanto que \mathbf{Q}_l está relacionado à \mathcal{S}_l^T . Todavia, a matriz \mathcal{E}_l não permite criar uma relação direta entre tais termos. Porém, ao reescrever a equação (3.46) da seguinte forma:

$$\mathbf{H}_l(0) = \mathbf{R}_l \mathcal{E}_l^{\frac{1}{2}} \mathcal{E}_l^{\frac{1}{2}} \mathcal{S}_l^T, \quad (3.47)$$

a seguinte relação pode ser criada:

$$\hat{\mathbf{C}} = \tilde{\mathcal{P}}_l = \mathbf{R}_l \mathcal{E}_l^{\frac{1}{2}}, \quad (3.48)$$

$$\hat{\mathbf{B}} = \mathbf{Q}_l = \mathcal{E}_l^{\frac{1}{2}} \mathcal{S}_l^T, \quad (3.49)$$

e esta escolha faz com que as matrizes de observabilidade e controlabilidade sejam balanceadas (JUANG, 1994).

Para estimar a matriz \mathbf{A} , deve-se observar que a matriz de Hankel, apresentada na equação (3.37), pode ser escrita em função das matrizes de observabilidade e controlabilidade, da seguinte forma:

$$\mathbf{H}(k-1) = \mathbf{P}\mathbf{A}^{k-1}\mathbf{Q}. \quad (3.50)$$

Portanto, para se obter uma estimativa de \mathbf{A} , deve-se buscar uma forma de separar estes termos. Para isto, deve-se desenvolver a relação através da equação dos parâmetros de Markov do sistema:

$$\begin{aligned} \mathbf{Y}_T^k &= \mathbf{C}\mathbf{A}^{k-1}\mathbf{B} \\ \mathbf{C}\mathbf{A}^{k-1}\mathbf{B} &= \tilde{\mathcal{P}}_l \mathbf{A}^{k-1} \mathbf{Q}_l \\ &= \mathbf{H}_l(0) \mathbf{H}_l^\dagger \tilde{\mathcal{P}}_l \mathbf{A}^{k-1} \mathbf{Q}_l \mathbf{H}_l^\dagger \mathbf{H}_l(0) \\ &= \mathbf{R}_l \mathcal{E}_l \mathcal{S}_l^T \mathcal{S}_l \mathcal{E}_l^{-1} \mathbf{R}_l^T \tilde{\mathcal{P}}_l \mathbf{A}^{k-1} \mathbf{Q}_l \mathcal{S}_l \mathcal{E}_l^{-1} \mathbf{R}_l^T \mathbf{R}_l \mathcal{E}_l \mathcal{S}_l^T \\ &= \mathbf{R}_l \mathcal{E}_l^{\frac{1}{2}} [\mathcal{E}_l^{-\frac{1}{2}} \mathbf{R}_l^T \tilde{\mathcal{P}}_l \mathbf{A}^{k-1} \mathbf{Q}_l \mathcal{S}_l \mathcal{E}_l^{-\frac{1}{2}}] \mathcal{E}_l^{\frac{1}{2}} \mathcal{S}_l, \\ &= \mathbf{C} [\mathcal{E}_l^{-\frac{1}{2}} \mathbf{R}_l^T \tilde{\mathcal{P}}_l \mathbf{A}^{k-1} \mathbf{Q}_l \mathcal{S}_l \mathcal{E}_l^{-\frac{1}{2}}] \mathbf{B} \\ &= \mathbf{C} [\mathcal{E}_l^{-\frac{1}{2}} \mathbf{R}_l^T \tilde{\mathcal{P}}_l \mathbf{A} \mathbf{Q}_l \mathcal{S}_l \mathcal{E}_l^{-\frac{1}{2}}]^{k-1} \mathbf{B} \\ &= \mathbf{C} [\mathcal{E}_l^{-\frac{1}{2}} \mathbf{R}_l^T \mathbf{H}(1) \mathcal{S}_l \mathcal{E}_l^{-\frac{1}{2}}]^{k-1} \mathbf{B}. \end{aligned} \quad (3.51)$$

Neste ponto, observa-se que como $\tilde{\mathcal{P}}_l = \mathbf{R}_l \mathcal{E}_l^{\frac{1}{2}}$ e $\mathbf{Q}_l = \mathcal{E}_l^{\frac{1}{2}} \mathcal{S}_l^T$, as seguintes relações podem ser descritas:

$$\mathcal{E}_l^{-\frac{1}{2}} \mathbf{R}_l^T \tilde{\mathcal{P}}_l = \mathbf{I}_l, \quad (3.52)$$

$$\mathbf{Q}_l \mathcal{S}_l \mathcal{E}_l^{-\frac{1}{2}} = \mathbf{I}_l. \quad (3.53)$$

Diante disso, o expoente $k-1$ pode ser transferido para fora dos colchetes. Com base na expressão fornecida por (3.51), a matriz \mathbf{A} pode ser estimada da seguinte forma:

$$\hat{\mathbf{A}} = \mathcal{E}_l^{-\frac{1}{2}} \mathbf{R}_l^T \mathbf{H}(1) \mathcal{S}_l \mathcal{E}_l^{-\frac{1}{2}}. \quad (3.54)$$

Observa-se que as expressões obtidas nesta seção são referentes a uma realização mínima baseada na significância dos autovalores. Além disso, deve-se notar que a matriz \mathbf{D} não é obtida pelo ERA, pois essa é estimada diretamente através dos parâmetros de Markov do observador, como apresentado na equação (3.24). Assim, o método descrito estima o seguinte modelo para o observador:

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{A}}\hat{\mathbf{x}}_k + \hat{\mathbf{B}}\mathbf{u}_k + \hat{\mathbf{G}}(\hat{\mathbf{y}}_k - \mathbf{y}_k), \quad (3.55)$$

$$\hat{\mathbf{y}}_k = \hat{\mathbf{C}}\hat{\mathbf{x}}_k + \hat{\mathbf{D}}\mathbf{u}_k, \quad (3.56)$$

3.3 Relação Entre o Observador e o Filtro de Kalman

No processo tradicional de filtragem de Kalman, o ganho do filtro é utilizado para obter uma estimativa ótima baseada nas medições e num valor estimado de forma a priori. Para determinar este ganho, é necessário conhecimento do modelo do sistema e das covariâncias dos ruídos de medição e do processo. Nas seções anteriores, foi apresentado um método para estimar um modelo matemático do sistema no espaço de estado. A covariância do ruído de medição pode ser estimada através de informações sobre os sensores utilizados, porém, a covariância do ruído de processo não pode ser obtida diretamente. Nesse sentido, esta seção visa criar uma relação entre o ganho do observador identificado e o ganho do filtro de Kalman. A formulação matemática desta seção foi baseada em (JUANG, 1994).

Para formular a relação, deve-se partir do seguinte modelo no espaço de estado:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \boldsymbol{\varpi}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{e}_k, \end{aligned} \quad (3.57)$$

onde $\boldsymbol{\varpi}_k$ é o ruído do processo, assumido ser um ruído Gaussiano de média zero e com matriz de covariância \mathbf{Q} , \mathbf{e}_k é o ruído de medição, o qual assume-se possuir as mesmas características de $\boldsymbol{\varpi}_k$, porém com matriz de covariância \mathbf{R} . O processo de filtragem de Kalman consiste em uma etapa de propagação e outra de assimilação. Na etapa de propagação, a estimativa a posteriori é propagada utilizando o modelo do sistema, da seguinte forma:

$$\hat{\mathbf{x}}_{k+1}^- = \mathbf{A}\hat{\mathbf{x}}_k^+ + \mathbf{B}\mathbf{u}_k, \quad (3.58)$$

onde o sobrescrito $(.)^+$ indica uma estimativa a posteriori e o sobrescrito $(.)^-$ indica uma estimativa a priori. Na etapa de assimilação do filtro, a medição mais recente é utilizada para corrigir a estimativa a priori:

$$\hat{\mathbf{x}}_{k+1}^+ = \hat{\mathbf{x}}_{k+1}^- + \mathbf{K}[\mathbf{y}_k - \hat{\mathbf{y}}_k], \quad (3.59)$$

onde \mathbf{K} é o ganho do filtro de Kalman. Estas duas etapas podem ser agregadas numa única equação, através de:

$$\hat{\mathbf{x}}_{k+1}^+ = \mathbf{A}\hat{\mathbf{x}}_k^+ + \mathbf{B}\mathbf{u}_k + \mathbf{K}[\mathbf{y}_k - \hat{\mathbf{y}}_k], \quad (3.60)$$

onde $\hat{\mathbf{y}}$ é a estimativa da variável de saída, obtida através do modelo do sistema. Observando-se esta equação, pode-se analisar que, em caso de $\mathbf{G} = -\mathbf{K}$, a equação 3.60 torna-se equivalente à equação (3.55). No entanto, para afirmar que o observador de estados obtido nas equações (3.48), (3.49) e (3.54) é um filtro de Kalman, o erro de aproximação deve ter as mesmas características que uma sequência de resíduo de um filtro de Kalman. Para analisar esta relação, deve-se definir o erro de aproximação dos parâmetros do observador como:

$$\boldsymbol{\epsilon} = \bar{\mathbf{y}} - \bar{\mathbf{Y}}_T \bar{\boldsymbol{\Upsilon}}, \quad (3.61)$$

e então manipular algebricamente a expressão da seguinte forma:

$$\begin{aligned} \boldsymbol{\epsilon} \bar{\boldsymbol{\Upsilon}}^T &= \bar{\mathbf{y}} \bar{\boldsymbol{\Upsilon}}^T - \bar{\mathbf{Y}}_T \bar{\boldsymbol{\Upsilon}} \bar{\boldsymbol{\Upsilon}}^T \\ \boldsymbol{\epsilon} \bar{\boldsymbol{\Upsilon}}^T &= \bar{\mathbf{y}} \bar{\boldsymbol{\Upsilon}}^T - \bar{\mathbf{y}} \bar{\boldsymbol{\Upsilon}}^T [\bar{\boldsymbol{\Upsilon}} \bar{\boldsymbol{\Upsilon}}^T]^{-1} [\bar{\boldsymbol{\Upsilon}} \bar{\boldsymbol{\Upsilon}}^T] \\ \boldsymbol{\epsilon} \bar{\boldsymbol{\Upsilon}}^T &= 0. \end{aligned} \quad (3.62)$$

Este termo pode ser analisado elemento a elemento através da seguinte expressão:

$$\boldsymbol{\epsilon} \bar{\mathbf{v}}_i^T = \sum_{k=p}^{l-1} \epsilon_k \mathbf{v}_{k-p+i}^T, \quad i = 0, 1, \dots, p, \quad (3.63)$$

onde $\bar{\mathbf{v}}_i$ representa a i -ésima linha da matriz de regressores $\bar{\boldsymbol{\Upsilon}}$. Dado que o sistema estimado é estacionário, e que a quantidade de dados é grande o suficiente, é possível utilizar a propriedade ergódica para estimar a média da variável, da seguinte maneira:

$$E[\epsilon_k \mathbf{v}_{k-p+i}^T] = \lim_{l \rightarrow \infty} \frac{1}{l-p} \sum_{k=p}^{l-1} \epsilon_k \mathbf{v}_{k-p+i}^T = 0, \quad k > p. \quad (3.64)$$

Devido a definição de \mathbf{v} , a expressão acima pode ser reescrita em função das variáveis de entrada e saída:

$$E[\epsilon_k \mathbf{u}_{k-i}^T] = 0; \quad i = 0, \dots, p, \quad (3.65)$$

$$E[\epsilon_k \mathbf{y}_{k-j}^T] = 0; \quad j = 0, \dots, p. \quad (3.66)$$

Desse modo, se a quantidade de dados e a ordem do sistema forem suficientemente grandes, o observador estimado é um filtro de Kalman e seu ganho pode ser calculado da seguinte forma:

$$\mathbf{K} = -\mathbf{G}. \quad (3.67)$$

Porém, alguns fatores como não linearidades e ruídos não brancos podem resultar na perda dessa característica. Neste caso, o modelo estimado se resume a um observador (JUANG, 1994).

4 Análise Espectral Singular

A Análise Espectral Singular (AES) é uma técnica utilizada para a análise de séries temporais que visa decompô-la em uma soma de componentes espectrais não observáveis (GOLYANDINA; ZHIGLJAVSKY, 2013; GOLYANDINA; SHLEMOV, 2013; HASSANI; ZHIGLJAVSKY, 2009). Tal técnica pode ser utilizada para a suavização da série, redução de ruído, extração de tendências, entre outras. Este algoritmo é bastante utilizado na literatura (VAUTARD; GHIL, 1989; LIAO; LING; HO, 2022; MASYITOH; ZUKHRONAH; SLAMET, 2022), pois ele não necessita de um modelo, ou premissas acerca da estacionariedade da série temporal (HASSANI; THOMAKOS, 2010).

O algoritmo tradicional da AES é aplicado a séries temporais monovariáveis (GOLYANDINA; ZHIGLJAVSKY, 2013), com o objetivo de separar a série em uma soma de componentes interpretáveis, como apresentado a seguir:

$$x_k = \tau_k + \delta_k + \beta_k + \dots + \epsilon_k, \quad (4.1)$$

onde x_k é a k -ésima observação da série temporal x , τ_k é a tendência da série, δ_k é o padrão cíclico, β_k é a sazonalidade, e ϵ_k é a componente ruidosa (JÚNIOR; SERRA, 2020). A quantidade de componentes depende da complexidade da série temporal e do objetivo em questão.

Para uma fácil compreensão, o método da AES é dividido em estágios de decomposição e reconstrução. Estes serão apresentados na sequência.

4.1 Estágio de Decomposição

No estágio de decomposição, o objetivo é obter uma representação multidimensional da série original, e então, procurar padrões que contenham maiores proporções de variância (ALLEN; SMITH, 1996). Para isso, este estágio é dividido em etapas de incorporação e decomposição em valores singulares.

4.1.1 Etapa de Incorporação

Na etapa de incorporação, uma representação multivariável da série original, conhecida como matriz trajetória, é obtida. Esta é construída através de vetores formados por uma janela móvel de η observações acerca da série temporal original. Esta é expressa

a seguir:

$$\mathbf{M} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_\rho \\ x_2 & x_3 & x_4 & \cdots & x_{\rho+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_\eta & x_{\eta+1} & x_{\eta+2} & \cdots & x_N \end{bmatrix} \in \mathbb{R}^{\eta \times \rho}, \quad (4.2)$$

onde \mathbf{M} é a matriz trajetória, η é a dimensão da janela de observações, e ρ é a quantidade de vetores atrasados. Observa-se que cada coluna da matriz trajetória corresponde a um deslocamento na janela de observações da série original. A dimensão η é definida pelo projetista e deve ser menor que a metade do número de observações. Já a quantidade de vetores atrasados ρ é definida pela seguinte relação:

$$\rho = N - \eta + 1, \quad (4.3)$$

onde N número de observações.

4.1.2 Etapa de Decomposição em Valores Singulares

Nesta etapa, a matriz trajetória deve ser decomposta em função de seus valores singulares (*Singular Value Decomposition*- SVD), como demonstrado por:

$$\mathbf{M} = \mathbf{\Phi} \mathbf{\mathfrak{P}} \mathbf{\mathfrak{D}}^T, \quad (4.4)$$

onde $\mathbf{\Phi} \in \mathbb{R}^{\eta \times \eta}$ é a matriz que contém os vetores singulares a esquerda de \mathbf{M} , $\mathbf{\mathfrak{D}} \in \rho \times \rho$ é a matriz que contém os vetores singulares a direita de \mathbf{M} , e $\mathbf{\mathfrak{P}} \in \mathbb{R}^{\eta \times \rho}$ é uma matriz diagonal que contém os valores singulares de \mathbf{M} . Para realizar esta decomposição, deve-se partir da seguinte matriz de covariância:

$$\mathbf{S} = \mathbf{M} \mathbf{M}^T, \quad \in \mathbb{R}^{\eta \times \eta} \quad (4.5)$$

e então calcular seus autovalores e autovetores:

$$[\mathbf{\Phi}_S, \mathbf{\Sigma}] = \text{eig}(\mathbf{S}), \quad (4.6)$$

onde $\mathbf{\Phi}_S \in \mathbb{R}^{\eta \times \eta}$ é a matriz que contém os autovetores de \mathbf{S} , e $\mathbf{\Sigma} \in \mathbb{R}^\eta$ é um vetor que contém os autovalores da matriz de covariância. Estes estão armazenados em ordem de significância ($\sigma_1 > \sigma_2 > \cdots > \sigma_\eta$). Por propriedade, os autovetores de \mathbf{S} são iguais aos vetores singulares à direita de \mathbf{M} :

$$\mathbf{\Phi} = \mathbf{\Phi}_S, \quad (4.7)$$

porém, os valores singulares de \mathbf{M} estão relacionados aos autovalores de \mathbf{S} da seguinte forma:

$$\mathfrak{p}_k = \sqrt{\sigma_k}, \quad (4.8)$$

onde \mathfrak{p}_k é o k -ésimo valor singular de \mathbf{M} . Para obter os vetores singulares a direita de \mathbf{M} , deve-se manipular algebricamente a equação (4.4), levando em conta que $\mathbf{\Phi}$ e $\mathbf{\mathfrak{D}}$

são matrizes ortogonais, e \mathfrak{P} é uma matriz diagonal. Logo, suas inversas são de simples obtenção. Este processo é descrito a seguir:

$$\begin{aligned}
\mathbf{M} &= \mathbf{\Phi} \mathfrak{P} \mathfrak{D}^T \\
\mathfrak{P}^{-1} \mathbf{\Phi}^T \mathbf{M} &= \mathbf{\Phi}^T \mathfrak{P} \mathfrak{P}^{-1} \mathfrak{D}^T \\
\mathfrak{P}^{-1} \mathbf{\Phi}^T \mathbf{M} &= \mathbf{I} \mathfrak{D}^T \\
\mathfrak{D} &= \mathfrak{P}^{-1} \mathbf{M}^T \mathbf{\Phi} \\
\mathfrak{D} &= \begin{bmatrix} \frac{1}{\sqrt{\sigma_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{\sigma_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{\sigma_\delta}} \end{bmatrix} \mathbf{M}^T \mathbf{\Phi} \\
\mathfrak{D} &= \left[\frac{\mathbf{M}^T \phi_1}{\sqrt{\sigma_1}} \quad \frac{\mathbf{M}^T \phi_2}{\sqrt{\sigma_2}} \quad \cdots \quad \frac{\mathbf{M}^T \phi_\delta}{\sqrt{\sigma_\delta}} \right].
\end{aligned} \tag{4.9}$$

Assim, cada vetor singular contido em \mathfrak{D} pode ser expresso por:

$$\mathfrak{D}_k = \frac{\mathbf{M}^T \phi_k}{\sqrt{\sigma_k}}. \tag{4.10}$$

Dessa forma, a matriz trajetória pode ser decomposta em função de cada vetor e valor singular:

$$\mathbf{M} = \sum_{i=1}^{\delta} \mathbf{M}_i = \sum_{i=1}^{\delta} \sqrt{\sigma_i} \phi_i \mathfrak{D}_i^T. \tag{4.11}$$

4.2 Estágio de Reconstrução

No estágio de reconstrução, a versão decomposta da matriz trajetória deve retornar aos padrões de dimensionalidade da série original. Este estágio está dividido em etapas de agrupamento de índices e média da diagonal.

4.2.1 Etapa de Agrupamento de Índices

Para evitar problemas com custo computacional relacionado a extração de η componentes espectrais da série temporal, as componentes decompostas de \mathbf{M}_i serão agrupadas conforme a significância de seus valores singulares, criando assim, um conjunto de n componentes espectrais não-observáveis. Dependendo da aplicação, a forma como esses índices são agrupados pode variar. Em geral, o especialista informa quantas componentes espectrais serão extraídas (GOLYANDINA; ZHIGLJAVSKY, 2013). Assim, a equação (4.11) pode ser reescrita em função da soma de n subconjuntos:

$$\mathbf{M} = \mathbf{M}_{I_1} + \cdots + \mathbf{M}_{I_n}, \tag{4.12}$$

onde \mathbf{M}_{I_1} contém as r componentes mais significativas da série temporal. Esta é expressa por:

$$\mathbf{M}_{I_1} = \mathbf{M}_1 + \mathbf{M}_2 + \cdots + \mathbf{M}_r. \tag{4.13}$$

4.2.2 Etapa de Média da Diagonal

Nesta etapa, cada subconjunto \mathbf{M}_{I_n} é transformado em uma componente não-observável da série temporal. Para isso, é realizada uma média sobre os elementos da matriz de cada subconjunto. Este procedimento é expresso por:

$$a_k^j = \begin{cases} \frac{1}{k} \sum_{i=1}^k m_{i,k-i+1}^j, & 1 \leq k \leq \eta \\ \frac{1}{\delta} \sum_{i=1}^{\delta} m_{i,k-i+1}^j, & \eta \leq k \leq \rho \\ \frac{1}{N-k+1} \sum_{i=k-\rho+1}^{N-\rho+1} m_{i,k-i+1}^j, & \rho \leq k \leq N \end{cases}, \quad (4.14)$$

onde $m_{i,k}^j$ é o elemento da i -ésima linha e k -ésima coluna do subconjunto \mathbf{M}_{I_j} , e a_k^j é o k -ésimo elemento da j -ésima componente espectral. Como resultado, a série temporal original passa a ser aproximada como uma soma de suas componentes não-observáveis:

$$x_k = a_k^1 + a_k^2 + \cdots + a_k^n. \quad (4.15)$$

Dependendo da aplicação, apenas algumas componentes espectrais serão necessárias. Quando o objetivo é remover ruído da base de dados, apenas as componentes mais significativas são necessárias (a_k^1). Nestes casos, existem dois parâmetros que influenciam na qualidade da remoção das componentes ruidosas, são eles: a quantidade de componentes mais significativas e a dimensão da matriz trajetória.

4.3 Atualização Recursiva

Com a chegada de novas medições na base dados, se torna necessário atualizar as componentes espectrais. Nesse sentido, em Selmo 2022, foi desenvolvida uma atualização recursiva para as componentes espectrais. Para realizar essa atualização, é importante notar que com a chegada de uma nova medição, a dimensão da matriz trajetória aumenta em uma unidade, conforme:

$$\rho^l = \rho^{(l-1)} + 1, \quad (4.16)$$

onde $\rho^{(l-1)}$ corresponde ao valor anterior para a quantidade de colunas de \mathbf{M} . Esta nova coluna é definida por:

$$\boldsymbol{\psi}_k = [y_\rho, y_{\rho+1}, \cdots, y_k]^T. \quad (4.17)$$

Este novo vetor deve ser utilizado para atualizar recursivamente a matriz de covariância \mathbf{S} , através da seguinte expressão:

$$\mathbf{S}^l = \mathbf{S}^{l-1} + \boldsymbol{\psi}_k \boldsymbol{\psi}_k^T. \quad (4.18)$$

Após a atualização da matriz de covariância, pode-se atualizar os vetores e valores singulares obtidos nas equações (4.6) e (4.10). Para isso, utiliza-se a ideia de matrizes de

pertubação, conforme apresentado em [Peddaneni et al. \(2004\)](#). Para atualizar os autovalores e autovetores de \mathbf{S} , é necessário calcular o seguinte vetor:

$$\mathbf{F}_j^{(l)} = \Phi_{j,k}^{(l-1)T} \boldsymbol{\psi}_j^{(l)}, \quad \in \mathbb{R}^{\eta \times 1}. \quad (4.19)$$

Os elementos deste vetor são utilizados para calcular as seguintes matrizes de perturbação:

$$\boldsymbol{\Omega}_j^{(l)} = \begin{bmatrix} \omega_{j,1} & 0 & \cdots & 0 \\ 0 & \omega_{j,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_{j,\eta} \end{bmatrix}, \quad (4.20)$$

$$\boldsymbol{\Lambda}_j = \begin{bmatrix} 0 & \lambda_j^{1,2} & \cdots & \lambda_j^{1,\eta} \\ -\lambda_j^{1,2} & 0 & \cdots & \lambda_j^{2,\eta} \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_j^{1,\eta} & -\lambda_j^{2,\eta} & \cdots & 0 \end{bmatrix}, \quad (4.21)$$

onde:

$$\omega_{j,r} = \frac{(f_j^k)^2 - \sigma_{j-1}^k}{N}, \quad (4.22)$$

$$\lambda_j^{k,r} = \begin{cases} 0, & \text{quando } k = r, \\ \frac{f_k^j f_k^r}{k[\lambda_j^r - \lambda_j^j + \sigma_{j-1}^r - \sigma_{j-1}^k]}, & \text{quando } k \neq r, \end{cases} \quad (4.23)$$

As matrizes de perturbação calculadas nas equações anteriores são utilizadas para atualizar recursivamente os autovetores e autovalores de \mathbf{S} , através de:

$$\Phi_j^{(l)} = \Phi_j^{(l-1)}(\mathbf{I} + \boldsymbol{\Lambda}_j), \quad (4.24)$$

$$\boldsymbol{\Sigma}_j^{(l)} = \boldsymbol{\Sigma}_j^{(l-1)} + \boldsymbol{\Omega}_j. \quad (4.25)$$

Com tais elementos atualizados, deve-se calcular os vetores singulares a direita de \mathbf{M} , através da equação (4.10), e então aplicar a etapa da média da diagonal, para atualizar as componentes espectrais.

Parte III

REFERENCIAL METODOLÓGICO

5 METODOLOGIA PROPOSTA: PROJETO DO FILTRO DE KALMAN NEBULOSO TIPO-2 INTERVALAR

Neste capítulo, apresenta-se a metodologia para estimação paramétrica recursiva de filtros adaptativos de Kalman nebulosos tipo-2 intervalares, com base no critério de máxima verossimilhança. Esta metodologia é dividida em etapas *offline* e *online*. Na etapa *offline*, um conjunto de dados abrangente é utilizado para gerar um modelo inicial para o sistema. Na etapa *online*, este modelo é atualizado de forma recursiva. Em ambas as etapas, a ordem de cada submodelo é determinada de acordo com a quantidade de parâmetros de Markov, a qual é determinada através do critério da informação de Akaike.

5.1 Modelo do Filtro de Kalman Nebuloso Tipo-2 Intervalar

Esta dissertação está inserida no contexto da modelagem nebulosa tipo-2 intervalar, onde um conjunto de dados acerca das variáveis de entrada e saída do sistema é coletado e aplicado à estimação paramétrica deste. Para modelar este sistema, o seguinte formato de regra nebulosa tipo-2 intervalar é utilizado:

$$R^{(i)} : \text{ Se } \mathbf{Z}_k \text{ é } \tilde{\mathbf{U}}^i$$

$$\text{Então } \begin{cases} \tilde{\mathbf{x}}_k^i = \tilde{\mathbf{A}}_k^i \tilde{\mathbf{x}}_{k-1}^i + \tilde{\mathbf{B}}_k^i \mathbf{u}_{k-1} + \tilde{\mathbf{K}}_k^i \tilde{\boldsymbol{\epsilon}}_{k-1}^i, \\ \tilde{\mathbf{y}}_k^i = \tilde{\mathbf{C}}_k^i \tilde{\mathbf{x}}_k^i + \tilde{\mathbf{D}}_k^i \mathbf{u}_k \end{cases} \quad (5.1)$$

onde $\mathbf{Z}_k \in \mathbb{R}^{w+n}$ é a variável linguística do antecedente, $\tilde{\mathbf{U}}^i$ é o i -ésimo conjunto nebuloso tipo-2 intervalar, $\tilde{\mathbf{A}}^i \in \mathbb{R}^{\iota \times \iota}$, $\tilde{\mathbf{B}}^i \in \mathbb{R}^{\iota \times n}$, $\tilde{\mathbf{C}}^i \in \mathbb{R}^{w \times \iota}$, $\tilde{\mathbf{D}}^i \in \mathbb{R}^{w \times n}$ e $\tilde{\mathbf{K}}^i \in \mathbb{R}^{w \times \iota}$ são respectivamente as matrizes de estados, entrada, saída, transmissão direta e ganho de Kalman; $\tilde{\mathbf{x}}_k^i \in \mathbb{R}^{\iota \times w}$ é o vetor de estados estimados, $\tilde{\mathbf{y}}_k^i \in \mathbb{R}^w$ é o vetor de saídas estimadas, $\mathbf{u} \in \mathbb{R}^n$ é o vetor de variáveis de entrada do sistema, ι é a ordem do sistema, w é a quantidade de variáveis de saída, n é o número de variáveis de entrada, e $\tilde{\boldsymbol{\epsilon}}^i$ é a sequência de resíduo da i -ésima regra nebulosa definida por:

$$\tilde{\boldsymbol{\epsilon}}_{k-1}^i = \mathbf{y}_{k-1} - \tilde{\mathbf{y}}_{k-1}^i, \quad (5.2)$$

onde \mathbf{y}_{k-1} é a saída processada do sistema. Observe que o sobrescrito $(\tilde{\cdot})$ informa que esta é uma variável intervalar, enquanto que o sobrescrito $(\cdot)^i$ indica que a variável é relativa a i -ésima regra do modelo. Assim, a estimativa global para a variável de saída do modelo é obtida por:

$$\tilde{\mathbf{y}}_k = \sum_{i=1}^c \tilde{\mu}_{\mathbf{Z}_k}^i \tilde{\mathbf{y}}_k^i, \quad (5.3)$$

onde $\tilde{\mu}_{\mathbf{Z}_k}^i$ é o grau de pertinência intervalar da variável \mathbf{Z}_k em relação ao conjunto i .

A metodologia proposta nesta dissertação consiste em etapas de pré-processamento de dados, estimação paramétrica do antecedente e estimação paramétrica do conseqüente. No pré-processamento de dados, a AES é aplicada para remover os ruídos presentes na série temporal. Esta componente filtrada é utilizada para estimar os parâmetros do sistema e a covariância do ruído de medição. Na estimação paramétrica do antecedente, a matriz de partição $\tilde{\mathbf{U}}$ é obtida utilizando uma versão tipo-2 intervalar do algoritmo de agrupamento nebuloso de máxima verossimilhança. A partição obtida não é atualizada durante a etapa *online*, porém, o expoente de ponderação nebulosa é adaptado conforme o nível de incerteza da série temporal, o qual é informado pela estimação da covariância do ruído. Por fim, na estimação paramétrica do conseqüente, a representação no espaço de estado do filtro de Kalman é obtida. Para melhor estimação paramétrica possível, esta dissertação propõe utilizar o critério da informação de Akaike para selecionar a quantidade ideal de regressores para cada variável estimada, o que resulta em submodelos com ordens diferentes. Nesse sentido, a figura 9 apresenta os passos computacionais desenvolvidos na metodologia proposta.

Para a simulação dos modelos obtidos, consideram-se duas abordagens:

- **Previsão de um passo à frente:** os parâmetros do modelo são atualizados recursivamente e usados para estimar o valor da variável de saída para o próximo instante (AGUIRRE, 2007).
- **Simulação mista:** a atualização recursiva ocorre naturalmente até que uma situação onde os medidores param de enviar informações é simulada. Nesta situação, é realizada uma simulação livre até que novas medições sejam disponibilizadas. Após esse período, a atualização recursiva volta a ser realizada.

5.2 Pré-Processamento de Dados

Na modelagem matemática caixa-preta, os parâmetros do modelo escolhido são estimados mediante um conjunto de dados. Estes são corrompidos por ruído no momento de sua obtenção, o que aumenta a complexidade da estimação paramétrica (WINDMANN; HAEB-UMBACH, 2009; HENDRICKS; JANNERUP; SØRENSEN, 2008). Nesse sentido, para reduzir a dependência de ruído na base de dados, um conjunto de medições é pré-processado pela AES com o objetivo de remover componentes ruidosas. Este processo é realizado nas etapas *online* e *offline* da metodologia.

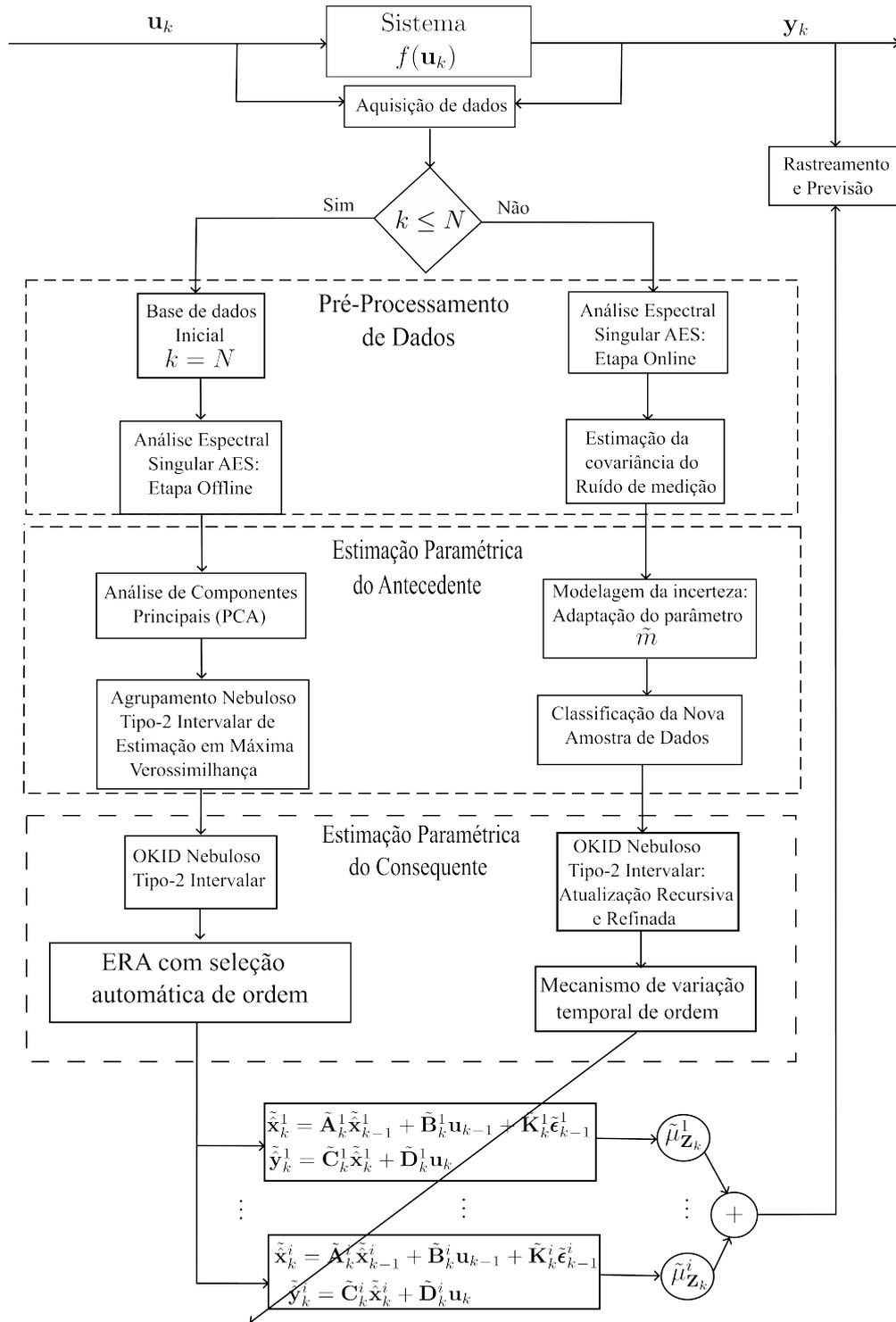


Figura 9 – Fluxograma contendo os passos computacionais desenvolvidos na metodologia proposta.

5.2.1 Etapa Offline

Na etapa *offline* da metodologia, uma versão em batelada da AES é aplicada. Nesta, algumas adaptações são propostas para facilitar o desenvolvimento do restante da metodologia. O conjunto de medições a ser processado é expresso pela seguinte série

temporal multivariável:

$$\mathbf{y} = \begin{bmatrix} y_1^{(1)} & y_2^{(1)} & \cdots & y_N^{(1)} \\ y_1^{(2)} & y_2^{(2)} & \cdots & y_N^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(w)} & y_2^{(w)} & \cdots & y_N^{(w)} \end{bmatrix}, \quad \mathbf{y} \in \mathbb{R}^{w \times N} \quad (5.4)$$

onde $y_k^{(j)}$ é k -ésima medição da j -ésima variável de saída, e N é o número de medições realizadas. O método tradicional da AES é aplicado a séries temporais monovariáveis. Portanto, para aplicá-la em um contexto multivariável, deve-se dividir \mathbf{y} em w séries temporais e processá-las paralelamente (JÚNIOR; SERRA, 2022). Assim, os procedimentos apresentados a seguir, devem ser realizados para cada variável da série temporal (w).

Passo 1 - Definir a dimensão da matriz trajetória. Para isso, o projetista deve definir a quantidade de linhas da matriz, respeitando a seguinte restrição:

$$\eta < \frac{N}{2}. \quad (5.5)$$

Assim, a quantidade de colunas é definida por:

$$\rho = N - \eta + 1. \quad (5.6)$$

Neste ponto, espera-se que quanto maior a quantidade de linhas da matriz, melhor será a qualidade da remoção do ruído, porém, será necessário um maior esforço computacional.

Passo 2- Calcular a matriz trajetória para cada série temporal, e suas matrizes de covariância:

$$\mathbf{M}_j = \begin{bmatrix} y_1^j & y_2^j & y_3^j & \cdots & y_\rho^j \\ y_2^j & y_3^j & y_4^j & \cdots & y_{\rho+1}^j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_\eta^j & y_{\eta+1}^j & y_{\eta+2}^j & \cdots & y_N^j \end{bmatrix} \in \mathbb{R}^{\eta \times \rho}, \quad (5.7)$$

$$\mathbf{S}_j = \mathbf{M}_j \mathbf{M}_j^T, \quad (5.8)$$

onde \mathbf{M}^j é a matriz trajetória da j -ésima série temporal, e $\mathbf{S}_j \in \mathbb{R}^{\eta \times \eta}$ é sua matriz de covariância.

Passo 3- Obter os autovetores e autovalores de \mathbf{S}_j em ordem decrescente de significância:

$$[\Phi_j \quad \Sigma_j] = eig(\mathbf{S}_j), \quad (5.9)$$

onde $\Phi_j \in \mathbb{R}^{\eta \times \eta}$ é a matriz cujas colunas são os autovetores de \mathbf{S}_j , e $\Sigma_j \in \mathbb{R}^\eta$ é o vetor contendo seus respectivos autovalores.

Passo 4- Selecionar a quantidade componentes espectrais mais significativas. Este é igual ao maior índice (k) no qual a seguinte expressão é verdadeira:

$$o_j = \arg \max_k \left(\frac{\sigma_{j,k}}{\sigma_{j,1}} \geq \chi \right), \quad (5.10)$$

onde $\sigma_{j,k}$ é o k -ésimo autovalor mais significativo da matriz \mathbf{S}_j , o_j é a quantidade de componentes espectrais selecionadas, e χ é um fator utilizado para selecionar a quantidade ideal de componentes mais significativas. Este valor pode variar de acordo com a aplicação.

Passo 5- Agrupar as componentes mais significativas da matriz trajetória:

$$\mathbf{\Gamma}_j = \sum_{k=1}^{o_j} \boldsymbol{\phi}_{j,k} \boldsymbol{\phi}_{j,k}^T \mathbf{M}_j, \quad (5.11)$$

onde $\mathbf{\Gamma}_j \in \mathbb{R}^{\eta \times \rho}$ é a matriz trajetória para a aproximação da componente real da série temporal \mathbf{y}_j , e $\boldsymbol{\phi}_{j,k}$ é o k -ésimo autovetor de \mathbf{S}_j . Observa-se que a equação (5.11) é uma simplificação do método apresentado na seção 4.1.2. Onde as componentes mais significativas da matriz trajetória são agrupadas sem a necessidade de calcular todos os seus autovetores a direita. Esta simplificação foi obtida ao manipular algebricamente as equações (4.10) e (4.11), da seguinte forma:

$$\begin{aligned} \mathbf{\Gamma}_j &= \sum_{k=1}^{o_j} \sqrt{\sigma_{j,k}} \boldsymbol{\phi}_{j,k} \boldsymbol{\Phi}^T \\ &= \sum_{k=1}^{o_j} \sqrt{\sigma_{j,k}} \boldsymbol{\phi}_{j,k} \frac{1}{\sqrt{\sigma_{j,k}}} [\mathbf{M}_j^T \boldsymbol{\phi}_{j,k}]^T \\ &= \sum_{k=1}^{o_j} \boldsymbol{\phi}_{j,k} \boldsymbol{\phi}_{j,k}^T \mathbf{M}_j. \end{aligned} \quad (5.12)$$

Onde o somatório seleciona apenas a o_j componentes mais significativas da matriz trajetória.

Passo 6- Obter uma aproximação para a componente real da série temporal:

$$\mathbf{y}_j^a = \begin{cases} \frac{1}{k} \sum_{i=1}^k \gamma_{i,k-i+1}, & 1 \leq k \leq \eta \\ \frac{1}{\eta} \sum_{i=1}^{\eta} \gamma_{i,k-i+1}, & \eta \leq k \leq \rho \\ \frac{1}{N-k+1} \sum_{i=k-\rho+1}^{N-\rho+1} \gamma_{i,k-i+1}, & \rho \leq k \leq N \end{cases}, \quad (5.13)$$

onde $\gamma_{i,k}$ é o elemento elemento presente na linha i e coluna k de $\mathbf{\Gamma}_j$, e \mathbf{y}_j^a é a uma aproximação para a componente real da j -ésima variável da série temporal \mathbf{y} . Este conjunto de dados é utilizado para estimar o modelo inicial do filtro de Kalman.

5.2.2 Etapa *online*

Para manter o modelo atualizado e lidar com possíveis mudanças de padrões, as componentes espectrais devem ser atualizadas. Porém, realizar o procedimento descrito acima a cada nova medição é bastante custoso em termos computacionais. Nesse contexto, esta dissertação propõe uma adaptação a atualização recursiva da AES proposta em Júnior e Serra (2022), de forma atualizar apenas as componentes mais significativas, reduzindo o custo computacional ao evitar operações redundantes.

Para iniciar o método, deve-se considerar que com a chegada de uma nova medição na série temporal, sua matriz trajetória passa a ter uma nova coluna, a qual é expressa por:

$$\boldsymbol{\psi}_k = [y_\rho, y_{\rho+1}, \dots, y_k]^T, \quad (5.14)$$

onde $\boldsymbol{\psi}_k$ é o vetor de regressores. Como o número de colunas aumentou, a variável ρ e a quantidade de observações N devem ser incrementadas em uma unidade.

A atualização das componetes espectrais é dependente da decomposição SVD. Portanto, os autovalores e autovetores de \mathbf{S}_j devem ser atualizados. Este processo é feito conforme o método apresentado em Júnior e Serra (2022). Dessa forma, os passos a seguir devem ser realizados a cada nova medição.

Passo 1- Calcular o vetor \mathbf{F}_j :

$$\mathbf{F}_j^{(l)} = \boldsymbol{\Phi}_{j,k}^{(l-1)T} \boldsymbol{\psi}_j^{(l)}, \quad (5.15)$$

onde o sobrescrito $(.)^{(l)}$ representa uma atualização da variável, enquanto que $(.)^{(l-1)}$ indica seu valor anterior.

Passo 2- Calcular as matrizes de perturbação $\boldsymbol{\Omega}_j^{(l)}$ e $\boldsymbol{\Lambda}_j^{(l)}$:

$$\boldsymbol{\Omega}_j^{(l)} = \begin{bmatrix} \omega_{j,1} & 0 & \dots & 0 \\ 0 & \omega_{j,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_{j,\eta} \end{bmatrix}, \quad (5.16)$$

$$\boldsymbol{\Lambda}_j = \begin{bmatrix} 0 & \lambda_j^{1,2} & \dots & \lambda_j^{1,\eta} \\ -\lambda_j^{1,2} & 0 & \dots & \lambda_j^{2,\eta} \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_j^{1,\eta} & -\lambda_j^{2,\eta} & \dots & 0 \end{bmatrix}, \quad (5.17)$$

onde os elementos $\lambda_j^{k,r}$ e $\omega_{j,r}$ são definidos por:

$$\omega_{j,r} = \frac{(f_j^k)^2 - \sigma_{j-1}^k}{N}, \quad (5.18)$$

$$\lambda_j^{k,r} = \begin{cases} 0, & \text{quando } k = r, \\ \frac{f_j^j f_k^r}{k[\lambda_j^r - \lambda_j^j + \sigma_{j-1}^r - \sigma_{j-1}^k]}, & \text{quando } k \neq r, \end{cases} \quad (5.19)$$

onde $f_{j,k}$ é o k -ésimo elemento do vetor $\mathbf{F}_j^{(l)}$.

Passo 3- Atualizar recursivamente os autovetores e autovalores:

$$\boldsymbol{\Phi}_j^{(l)} = \boldsymbol{\Phi}_j^{(l-1)} (\mathbf{I} + \boldsymbol{\Lambda}_j), \quad (5.20)$$

$$\boldsymbol{\Sigma}_j^{(l)} = \boldsymbol{\Sigma}_j^{(l-1)} + \boldsymbol{\Omega}_j. \quad (5.21)$$

Passo 4- Atualizar as componentes espectrais mais significativas através de:

$$y_j^a(N) = \sum_{j=1}^{o_j} \phi_k^j(N) \phi_k^j \psi_k, \quad (5.22)$$

onde $\phi_k^j(N)$ é o último elemento do autovetor ϕ_k^j . Com isso, uma aproximação para a componente filtrada da série temporal é obtida.

5.2.3 Estimação da Covariância do Ruído de Medição

Durante as etapas de estimação paramétrica, alguns parâmetros são adaptados de acordo com o nível de incerteza da saída do sistema. Assim, uma estimativa para a covariância de ruído de medição deve ser obtida. Para que tais parâmetros sejam adaptados conforme a necessidade local da série temporal, tal estimativa deve ser obtida em uma janela de observações. Nesse sentido, para a estimativa inicial, tanto a série temporal original quanto a processada devem ser truncadas, conforme:

$$\bar{\mathbf{y}}_j = \begin{bmatrix} y_{N-N_a}^{(j)} & y_{N-N_a+1}^{(j)} & \cdots & y_N^{(j)} \end{bmatrix}, \quad (5.23)$$

$$\bar{\mathbf{y}}_j^a = \begin{bmatrix} y_{N-N_a}^{a(j)} & y_{N-N_a+1}^{a(j)} & \cdots & y_N^{a(j)} \end{bmatrix}, \quad (5.24)$$

onde o sobrescrito $(\bar{\cdot})$ indica que esta série temporal contém apenas os últimos N_a elementos de sua versão inicial, e $y_k^{a(j)}$ é o k -ésimo termo da série processada. A janela de dados N_a deve ser grande o suficiente para capturar a covariância local. Assim, a sequência de ruído pode ser estimada da seguinte forma:

$$\boldsymbol{\xi}_j = \bar{\mathbf{y}}_j - \bar{\mathbf{y}}_j^a, \quad (5.25)$$

onde $\boldsymbol{\xi}_j$ é a janela de estimação do ruído de medição. Assim, a covariância desta sequência de ruído é obtida através de:

$$\Xi^{(0)} = \frac{1}{N_a} \boldsymbol{\xi}_j \boldsymbol{\xi}_j^T, \quad (5.26)$$

onde $\Xi^{(0)}$ é a estimativa inicial da covariância do ruído de medição. Este valor é obtido após a etapa *offline* da AES. Esta variável é atualizada a cada nova medição, através da seguinte relação recursiva:

$$\Xi^{(l)} = \Xi^{(l-1)} + \frac{1}{N_a} [(y_N^{(j)} - y_N^{a(j)})^2 - (y_{N-N_a}^{(j)} - y_{N-N_a}^{a(j)})^2], \quad (5.27)$$

lembrando que N é incrementado a cada nova medição disponível.

A etapa de pré-processamento de dados é implementada conforme descrito no algoritmo 1.

Algoritmo 1: Análise Espectral Singular Recursiva

```

1 Entrada:  $\mathbf{y}$ ,  $\eta$ ,  $\chi$ 
2 Saída:  $\mathbf{y}^a$ 
3 % Etapa Offline
4 for  $j = 1 : 3$  do
5     Passo 1: Calcular a matriz trajetória  $\mathbf{M}_j$ , Eq. (5.7);
6     Passo 2: Calcular a matriz covariância  $\mathbf{S}_j$ , Eq. (5.8);
7     Passo 3: Calcular os autovalores e autovetores de  $\mathbf{S}_j$ , Eq. (5.9);
8     Passo 4: Obter o número de componentes mais significativas  $o_j$ , Eq (5.10);
9     Passo 5: Agrupar as componentes mais significativas  $\mathbf{\Gamma}_j$ , Eq. (5.11);
10    Passo 6: Obter a componente processada da série temporal  $\mathbf{y}^a$ , Eq. (5.13);
11    Passo 7: Calcular a covariância de Ruído  $\Xi^{(0)}$ , Eq. (5.25)-(5.26).
12 end
13 % Etapa Online
14 while  $k > N + 1$  do
15     for  $j = 1 : 3$  do
16         Passo 8: Incrementar  $\eta$  and  $N$  em uma unidade;
17         Passo 9: Montar o vetor de regressores, Eq. (5.14);
18         Passo 10: Calcular o vetor  $\mathbf{F}_j$ , Eq. (5.15);
19         Passo 11: Calcular as matrizes de perturbação  $\mathbf{\Omega}_j$  e  $\mathbf{\Lambda}_j$ , Eq. (5.16)-(5.19);
20         Passo 11: Atualizar os autovetores e autovalores, Eq. (5.20)-(5.21);
21         Passo 12: Atualizar a componente  $\mathbf{y}_j^a$ , Eq. (5.22);
22         Passo 13: Atualizar a estimação da covariância  $\Xi^{(l)}$ , Eq. (5.27).
23     end
24 end

```

5.3 Estimação Paramétrica do Antecedente

Nesta dissertação, a estimação paramétrica do antecedente é dividida em etapas *offline* e *online*. Na primeira, a dimensão da base de dados é reduzida utilizando a análise de componentes principais. Este novo conjunto é então classificado utilizando o algoritmo de Agrupamento Nebuloso Tipo-2 Intervalar de Estimação em Máxima Verossimilhança (ANT-2IEMV). Já na etapa *online*, as novas amostras de dados são classificadas de acordo com os padrões obtidos na etapa *offline*. Esta classificação servirá como peso nebuloso tanto para a estimação paramétrica do consequente, como na estimação da saída do modelo. Nesta última, devido à ausência de um redutor de tipo, a abertura do intervalo deve ser controlada. Assim, o expoente de ponderação nebulosa é adaptado de acordo com a covariância do ruído nas variáveis de saída.

5.3.1 Agrupamento Nebuloso Tipo-2 Intervalar de Estimação em Máxima Verossimilhança

Após a redução da dimensionalidade, o novo conjunto de dados (\mathbf{Z}^{cp}) deve ser particionado em c conjuntos nebulosos. A versão tradicional deste algoritmo possui dois problemas para implementação, os quais são a necessidade de uma boa inicialização e a aparição de indeterminações matemáticas. Ambos os problemas são decorrentes de sua norma de distância exponencial. Porém, este algoritmo tende a se adaptar melhor ao formato e volume dos conjuntos (BABUŠKA, 2012). Visando os pontos positivos deste algoritmo, o ANT-2IEMV foi formulado com mecanismos para lidar com os dois problemas citados. Esta formulação é apresentada na sequência.

O ANT-2IEMV deve ser inicializado por alguns parâmetros, são eles: uma matriz de partição pseudoaleatória ($\tilde{\mathbf{U}}^0$), o critério de parada (tol), a quantidade de conjuntos (c) e o expoente de ponderação nebulosa (\tilde{m}). A maioria destes é definida pelo projetista e influenciam na qualidade do modelo.

Como apresentado em Pei et al. (2001), o expoente \tilde{m} possui um valor ótimo para o algoritmo c -means dentro do intervalo $[1, 5 \ 3, 5]$. Assim, em vez de procurar por um valor ótimo, é mais interessante do ponto de vista computacional analisá-lo como uma grandeza intervalar. Porém, devido à norma exponencial, os agrupamentos baseados em critério de máxima verossimilhança possuem formatos bem determinísticos para valores menores que dois. Nesse sentido, esta aplicação propõe deslocar este intervalo para $\tilde{m} = [2 \ 4]$, obtendo uma região de incerteza bem abrangente que será utilizada no treinamento do modelo.

Após selecionar os parâmetros iniciais e definir a matriz de partição pseudoaleatória, o primeiro passo para a implementação do algoritmo é calcular os centros dos conjuntos nebulosos, os quais são obtidos pela seguinte equação:

$$\tilde{\mathbf{v}}_i^l = \frac{\sum_{k=1}^N \tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l-1)} \mathbf{Z}_k^{cp}}{\sum_{k=1}^N \tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l-1)}}, \quad 1 \leq i \leq c, \quad (5.28)$$

onde $\tilde{\mathbf{v}}_i \in \mathbb{R}^q$ é o centro intervalar do i -ésimo conjunto nebuloso. Após esta etapa, deve-se calcular a distância entre cada ponto da base de dados e os centros dos conjuntos. Essa métrica será utilizada para avaliar a similaridade entre os dados coletados. No entanto, essa distância deve ser adaptada pela matriz de covariância dos conjuntos e a probabilidade a priori de selecionar o i -ésimo conjunto, os quais são definidos por:

$$\tilde{\mathbf{F}}^i = \frac{\sum_{k=1}^N \tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l-1)} (\mathbf{Z}_k^{cp} - \tilde{\mathbf{v}}_i)(\mathbf{Z}_k^{cp} - \tilde{\mathbf{v}}_i)^T}{\sum_{k=1}^N \tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l-1)}}, \quad (5.29)$$

$$\tilde{P}^i = \frac{1}{N} \sum_{k=1}^N \tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l-1)}, \quad 1 \leq i \leq c, \quad (5.30)$$

onde $\tilde{\mathbf{F}}^i \in \mathbb{R}^{q \times q}$ é a matriz de covariância do i -ésimo conjunto, e \tilde{P}^i é a probabilidade a priori deste conjunto ser selecionado. Estas métricas são aplicadas na norma de distância, para adaptar a métrica de similaridade conforme a distribuição local dos dados e o volume deste conjunto. Nesta etapa, o algoritmo encontra o primeiro problema para sua implementação, que é a necessidade de uma boa inicialização. Isso ocorre, pois, a função exponencial em sua norma de distância faz o algoritmo convergir próximo a um ótimo local. Dessa forma, para evitar o uso de outro agrupamento nebuloso para iniciá-lo, esta metodologia propõe uma norma de distância ponderada, definida pela seguinte expressão:

$$\tilde{\mathcal{D}}_k^i = \begin{cases} \tilde{\mathfrak{W}}_k^i, & C_t \geq 0,45 \\ \frac{C_t - Z_1}{Z_2 - Z_1} (\tilde{\mathfrak{W}}_k^i - \tilde{M}_k^i) + \tilde{M}_k^i, & 0,15 \leq C_t < 0,45 \\ \tilde{M}_k^i & C_t \leq 0,15 \end{cases} \quad (5.31)$$

onde $\tilde{\mathcal{D}}_k^i$ é a norma de distância resultante da amostra \mathbf{Z}_k^{cp} em relação ao conjunto i , C_t é o valor do teste de convergência, $\tilde{\mathfrak{W}}_k^i$ é uma norma de distância não exponencial, \tilde{M}_k^i é a norma de distância exponencial (padrão para o agrupamento de máxima verossimilhança). Tais normas de distâncias são expressas por:

$$\tilde{M}_k^i = \frac{\det(\tilde{\mathbf{F}}^i)^{\frac{1}{2}}}{\tilde{P}^i} \exp\left[\frac{1}{2}(\mathbf{Z}_k^{cp} - \tilde{\mathbf{v}}_i)^T \tilde{\mathbf{F}}^{i-1} (\mathbf{Z}_k^{cp} - \tilde{\mathbf{v}}_i)\right], \quad (5.32)$$

$$\tilde{\mathfrak{W}}_k^i = (\mathbf{Z}_k^{cp} - \tilde{\mathbf{v}}_i)^T [\det(\tilde{\mathbf{F}}^i)^{1/2} \tilde{\mathbf{F}}^{i-1}] (\mathbf{Z}_k^{cp} - \tilde{\mathbf{v}}_i). \quad (5.33)$$

Ao analisar a equação (5.31), pode-se constatar que quando o teste de convergência resulta em um valor alto (próximo a um), a norma de distância resultante leva em conta apenas sua parcela não exponencial, pois se entende, que o algoritmo está longe de convergir. Logo, é mais apropriado utilizar apenas $\tilde{\mathfrak{W}}_k^i$. À medida que o valor de C_t diminui, é realizada uma ponderação entre as duas normas de distância. Porém, quando o algoritmo se aproxima da convergência, é mais apropriado usar apenas \tilde{M}_k^i . Este mecanismo evita o uso de outro agrupamento para inicializar o ANT-2IMLE, ao usar uma norma de distância auxiliar.

Com a norma de distância obtida, deve-se atualizar a matriz de partição $\tilde{\mathbf{U}}$, através de:

$$\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l)} = [\mu_{\mathbf{Z}_k^{cp}}^{i(l)-}, \mu_{\mathbf{Z}_k^{cp}}^{i(l)+}] \in \tilde{\mathbf{U}}, \quad (5.34)$$

$$\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l)} = \left[\frac{1}{\sum_{j=1}^c \left(\frac{\mathcal{D}_k^{i-}}{\mathcal{D}_k^{j-}}\right)^{\left(\frac{2}{m^+ - 1}\right)}}, \frac{1}{\sum_{j=1}^c \left(\frac{\mathcal{D}_k^{i+}}{\mathcal{D}_k^{j+}}\right)^{\left(\frac{2}{m^+ - 1}\right)}} \right], \quad (5.35)$$

$$\mu_{\mathbf{Z}_k^{cp}}^{i(l)+} = \mathbf{max}[\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l)}], \quad \mu_{\mathbf{Z}_k^{cp}}^{i(l)-} = \mathbf{min}[\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(l)}], \quad (5.36)$$

onde os sobrescritos $(.)^-$ e $(.)^+$ indicam respectivamente os limites inferior e superior da variável. Na equação (5.35), encontra-se o segundo problema para implementação. Este ocorre, pois, as amostras de dados que estão distantes do conjunto em questão são improváveis de serem selecionadas. Assim, suas normas de distância são muito elevadas. Quando o valor desta ultrapassa o limite de memória disponível para uma variável, o

Matlab salva o valor desta como infinito. Dessa forma, quando a norma de distância para duas ou mais variáveis ultrapassa esse limite, o seguinte cenário ocorre:

$$\text{Se } \lim_{\mathbf{z}_k \rightarrow \mathbf{z}_z} \mathcal{D}_k^i \rightarrow \infty, \quad e \quad \lim_{\mathbf{z}_k \rightarrow \mathbf{z}_z} \mathcal{D}_k^j \rightarrow \infty, \quad \text{Então:} \quad (5.37)$$

$$\begin{aligned} \mu_{\mathbf{z}_z}^i &= \left(\lim_{\mathbf{z}_k \rightarrow \mathbf{z}_z} \frac{\mathcal{D}_k^i}{\mathcal{D}_k^1} + \dots + \lim_{\mathbf{z}_k \rightarrow \mathbf{z}_z} \frac{\mathcal{D}_k^i}{\mathcal{D}_k^i} + \dots + \lim_{\mathbf{z}_k \rightarrow \mathbf{z}_z} \frac{\mathcal{D}_k^i}{\mathcal{D}_k^j} + \dots + \lim_{\mathbf{z}_k \rightarrow \mathbf{z}_z} \frac{\mathcal{D}_k^i}{\mathcal{D}_k^c} \right)^{-1}, \\ &= \frac{1}{\infty + \dots + 1 + \dots + NaN + \dots + \infty}, \\ &= NaN, \end{aligned} \quad (5.38)$$

onde a sigla *NaN*, do inglês *Not a Number*, representa uma indeterminação matemática. Para eliminar este termo do processo, basta analisar o seguinte expressão, considerando-se que, por se tratar de normas de distância, seus valores são sempre positivos:

$$\lim_{\mathbf{z}_k \rightarrow \mathbf{z}_z} \frac{\mathcal{D}_k^i}{\mathcal{D}_k^j} = \frac{(0 \quad \infty]}{(0 \quad \infty]} = (0 \quad \infty]. \quad (5.39)$$

Portanto, o valor desta divisão será sempre um número positivo. Assim, ao aplicar esse resultado na equação (5.38), tem-se:

$$\begin{aligned} \mu_{\mathbf{z}_z}^i &= \frac{1}{\infty + \dots + 1 + \dots + (0 \quad \infty] + \dots + \infty}, \\ &= \frac{1}{\infty}, \\ &= 0. \end{aligned} \quad (5.40)$$

Se a norma de distância entre o ponto \mathbf{Z}_k^{cp} e o conjunto i tende ao infinito, seu grau de pertinência será nulo. Com essa alteração, a matriz de partição pode ser atualizada sem a aparição de indeterminações matemáticas. Com isso, o teste de convergência é realizado:

$$C_t = \|\bar{\mathbf{U}}^{l-1} - \bar{\mathbf{U}}^l\| < \epsilon, \quad (5.41)$$

onde:

$$\bar{\mathbf{U}} = \frac{1}{2}(\mathbf{U}^+ + \mathbf{U}^-). \quad (5.42)$$

O processo descrito deve ser realizado até que o teste de convergência seja verdadeiro.

5.3.2 Modelagem do Parâmetro de Incerteza

Durante a etapa a *offline* desta estimação, a matriz de partição, os centros dos conjuntos, assim como suas covariâncias e probabilidades a priori foram obtidas. Estas são utilizadas para ponderar o processo de estimação paramétrica do consequente, assim como gerar uma estimativa intervalar para a saída do modelo. Tal estimativa é resultado de uma aplicação da lógica nebulosa tipo-2 intervalar sem a presença de um redutor de tipo. Portanto, os limites inferior e superior da estimativa são definidos por:

$$\hat{\mathbf{y}}_k^- = \sum_{i=1}^c \min[\mu_{\mathbf{z}_k}^{i-} \hat{\mathbf{y}}_k^{i-}, \mu_{\mathbf{z}_k}^{i-} \hat{\mathbf{y}}_k^{i+}, \mu_{\mathbf{z}_k}^{i+} \hat{\mathbf{y}}_k^{i-}, \mu_{\mathbf{z}_k}^{i+} \hat{\mathbf{y}}_k^{i+}], \quad (5.43)$$

$$\hat{\mathbf{y}}_k^+ = \sum_{i=1}^c \max[\mu_{\mathbf{z}_k}^{i-} \hat{\mathbf{y}}_k^{i-}, \mu_{\mathbf{z}_k}^{i-} \hat{\mathbf{y}}_k^{j+}, \mu_{\mathbf{z}_k}^{i+} \hat{\mathbf{y}}_k^{i-}, \mu_{\mathbf{z}_k}^{i+} \hat{\mathbf{y}}_k^{i+}]. \quad (5.44)$$

Por essa formulação, é notável que a estimativa intervalar é centrada em uma previsão para o valor processado da série temporal. Assim, o intervalo estimado deve abranger as componentes da série temporal em volta de sua tendência. Entretanto, ao selecionar um valor fixo para \tilde{m} , perde-se o controle acerca do intervalo de estimação, o que em muitos cenários torna o intervalo bem mais abrangente que o necessário.

Inicialmente, o agrupamento foi realizado com o seguinte parâmetro $\tilde{m} = [2, 4]$. Este foi selecionado para que o intervalo estimado retenha o máximo de informações da série temporal original. Porém, para alguns pontos da base de dados, a abertura do intervalo estimado é bem superior ao necessário, tornando a estimativa pouco confiável. Assim, esta metodologia propõe adaptar o parâmetro \tilde{m} de forma a reduzir o intervalo estimado sem perder precisão. Para isso, considera-se que \tilde{m} é um intervalo centrado em três com abertura de mais ou menos um. Assim, para adaptá-los, deve-se modelar os limites de \tilde{m} através de uma função que expresse o nível de incerteza da série temporal (f_j), como apresentado a seguir:

$$m_j^- = 3 - f_j(I_{adp}^j), \quad (5.45)$$

$$m_j^+ = 3 + f_j(I_{adp}^j), \quad (5.46)$$

onde m_j^- e m_j^+ são respectivamente os limites inferior e superior do expoente de ponderação nebulosa da j -ésima variável, e I_{adp}^j é o nível de incerteza da série temporal, o qual é modelado da seguinte forma:

$$I_{adp}^j = \frac{\Xi^j}{|y_j^p(N)|^2}. \quad (5.47)$$

Observa-se que, como cada variável da série temporal pode possuir ruídos e amplitudes diferentes, a adaptação deve ser realizada para cada variável. Este nível de incerteza é utilizado como argumento da função f_j , a qual realiza uma transição entre a abertura mínima do intervalo (c_{min}) e seu valor padrão (mais ou menos um). O valor mínimo é variável para evitar que o intervalo feche demais e a aplicação perca sua característica intervalar. Assim, a expressão para c_{min} é dada por:

$$c_{min} = \begin{cases} 0,5, & \text{Se } |y_j^p(N)| < 1 \\ 0,025, & \text{Se } |y_j^p(N)| > 800. \\ \exp(\log(0,05) \frac{|y_j^p(N)|}{799} + \log(0,5) + \frac{\log(0,05)}{799}), & \text{Caso Contrário} \end{cases} \quad (5.48)$$

Esta equação seleciona valor inicial utilizado na função de transição. A partir disso, é realizada uma interpolação entre o valor obtido e o valor máximo para o intervalo. Esta equação foi modelada para gerar um intervalo maior para valores menos significativos da série temporal. Isso ocorre, pois, para valores baixos, o intervalo gerado pela equação (5.43) e (5.44) tende a ser menor, logo é necessário aumentar a incerteza na estimação paramétrica.

A função f_j é então modelada para gerar uma transição exponencial entre c_{min} e 1. Assim, a função é modelada por:

$$f_j(I_{adp}) = \begin{cases} c_{min}, & \text{Se } I_{adp} < 0,001 \\ 1, & \text{Se } I_{adp} > 0,3 \\ \exp(I_{adp} \frac{\log(c_{min})}{0,299} + 0,3 \frac{\log(c_{min})}{0,299}), & \text{Caso Contrário.} \end{cases} \quad (5.49)$$

Observa-se que esta função é fundamental para a metodologia, uma vez que ela retira a responsabilidade da escolha de \tilde{m} do projetista, além de adaptar esta dentro da mesma aplicação. Nesse sentido, nota-se, que quando o nível de ruído da série é muito baixo ($< 0,001$), a função seleciona o valor mínimo do intervalo, e quando o nível de ruído é elevado, a função seleciona o intervalo original, pois há a necessidade de aumentar a incerteza na estimação. Caso contrário, o valor obtido será selecionado dentro do intervalo. A figura 10 ilustra este processo, ao apresentar f_j para diferentes valores de c_{min} .

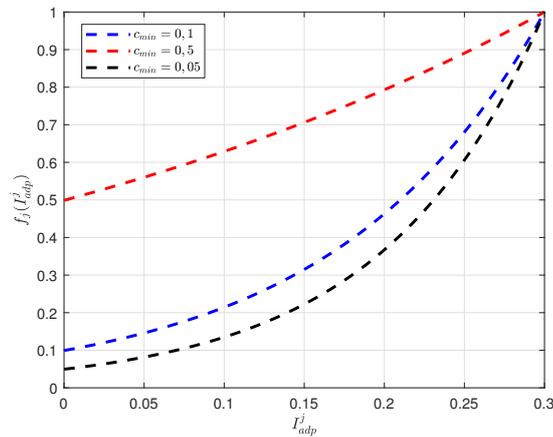


Figura 10 – Modelo de função interpoladora para selecionar a abertura do intervalo de \tilde{m} conforme a incerteza da série temporal. Nela, é demonstrado o comportamento da função de acordo com a seleção do valor mínimo do intervalo.

Ao observar a figura 10, nota-se diferentes comportamentos de acordo com a magnitude do valor medido. Quando $|y_j^p(N)|$ é alto, a função cresce de forma acelerada, pois para níveis de ruídos baixos, uma pequena alteração no intervalo de $\tilde{\mu}_{Z_k}^i$ é necessária para abranger tal variação.

5.3.3 Classificação da Nova Amostra de Dados

Como expresso anteriormente, as novas amostras de dados são classificadas com dois objetivos. O primeiro é para ponderar a estimação paramétrica do antecedente, e o segundo é gerar a estimativa intervalar. Assim, as novas amostras serão classificadas duas vezes, criando assim, dois conjuntos de graus de pertinência.

A primeira classificação é realizada com $\tilde{m} = [2, 4]$, sendo utilizada para gerar submodelos na qual a estimativa abranja o maior intervalo possível. Este grau de pertinência

é calculado utilizando as equações (5.35) e (5.36), e é usado para ponderar a estimação paramétrica do consequente.

A segunda classificação é utilizada para adaptar o intervalo estimado. Assim, os graus de pertinência são obtidos utilizando os novos limites para \tilde{m} :

$$\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(j)} = \left[\frac{1}{\sum_{j=1}^c \left(\frac{\mathcal{D}_k^{i-}}{\mathcal{D}_k^{j-}} \right)^{\left(\frac{2}{m_j^-} - 1 \right)}}, \frac{1}{\sum_{j=1}^c \left(\frac{\mathcal{D}_k^{i+}}{\mathcal{D}_k^{j+}} \right)^{\left(\frac{2}{m_j^+} - 1 \right)}} \right], \quad (5.50)$$

onde $\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(j)}$ é o grau de pertinência que será utilizado para gerar a estimativa do i -ésimo submodelo para a j -ésima variável de \mathbf{Z}_k .

O algoritmo 2 apresenta de forma simplificada o passo-a-passo desenvolvido nesta seção.

5.4 Estimação Paramétrica do Consequente

Para estimar os parâmetros no espaço de estado dos filtros de Kalman alocados na saída de cada regra nebulosa, é utilizado uma combinação dos algoritmos ERA e OKID. Para cumprir os objetivos desta dissertação, são realizadas três alterações na formulação da versão nebulosa deste algoritmo, e duas alterações na simulação do modelo obtido. As alterações na formulação consistem na seleção automática da quantidade de regressores utilizados, a seleção e atualização da ordem de cada submodelo, e por fim, uma atualização refinada para os parâmetros do modelo. As alterações na simulação do modelo consistem em um mecanismo para garantir a estabilidade do modelo durante a mudança de ordem do sistema, e uma mudança na formulação da equação da saída global do modelo, o que é necessário devido à adaptação do parâmetro de incerteza.

5.4.1 OKID Nebuloso Tipo-2 Intervalar

O algoritmo OKID nebuloso tipo-2 intervalar é utilizado para obter os parâmetros de Markov do sistema e do ganho do observador. Estes são obtidos via uma equação ARX. Portanto, é um modelo que depende de variáveis passadas. Logo, um dos parâmetros mais importantes a serem selecionados pelo projetista é a quantidade de regressores (p). Este é utilizado para gerar a matriz de regressores:

$$\mathbf{\Upsilon} = \begin{bmatrix} \mathbf{u}_p & \mathbf{u}_{p+1} & \cdots & \mathbf{u}_{l-1} \\ \mathbf{Z}_{p-1} & \mathbf{Z}_p & \cdots & \mathbf{Z}_{l-2} \\ \mathbf{Z}_{p-2} & \mathbf{Z}_{p-1} & \cdots & \mathbf{Z}_{l-3} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Z}_0 & \mathbf{Z}_1 & \cdots & \mathbf{Z}_{l-p-1} \end{bmatrix}, \quad (5.51)$$

Algoritmo 2: Estimaco Paramtrica do Antecedente: ANT-2IEMV + PCA

```

1 Entrada:  $\mathbf{Z}$ ,  $\epsilon$ ,  $c$ ,  $\tilde{\mathbf{U}}_0$ 
2 Saída:  $\tilde{\mathbf{U}}$ ,  $\tilde{\mu}_{\mathbf{Z}_k^{cp}}^i$ ,  $\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(j)}$ 
3 % Etapa Offline
4 Passo 1: Calcular o ponto mdio da base de dados, Eq. (2.27);
5 Passo 2: Calcular a varincia amostral, Eq. (2.28);
6 Passo 3: Obter os autovalores e autovetores de  $\mathbf{C}_Z$ , Eq. (2.29);
7 Passo 5: Selecionar as componentes principais;
8 Passo 6: Obter  $\mathbf{Z}^{cp}$  atravs da anlise de componentes principais PCA, Eq.
(2.27)-(2.31);
9 while  $C_t \geq \epsilon$  do
10   for  $i = 1 : c$  do
11     Passo 7: Calcular os centros dos conjuntos  $\tilde{\mathbf{v}}_i$ , Eq. (5.28);
12     Passo 8: Calcular a matriz de covarincia  $\tilde{\mathbf{F}}^i$  e probabilidade a priori  $\tilde{P}^i$ ,
Eq. (5.29)-(5.30);
13     Passo 9: Calcular a norma de distncia  $\tilde{\mathcal{D}}_k^i$ , Eq. (5.31)-(5.33);
14     Passo 10: Atualizar a matriz de partio  $\tilde{\mathbf{U}}$ , Eq. (5.34)-(5.36);
15     Passo 11: Teste de convergcia, Eq. (5.41)-(5.42);
16   end
17 end
18 % Etapa Online
19 while  $k > N$  do
20   Passo 12: Reduzir a dimensionalidade da nova amostra de dados  $\mathbf{Z}_k$ , Eq.
(2.31);
21   for  $i = 1 : c$  do
22     Passo 13: Calcular norma de distncia  $\tilde{M}_k^i$ , Eq. (5.32);
23     Passo 14: Calcular o grau de pertincia utilizado na atualizao do modelo
 $\tilde{\mu}_{\mathbf{Z}_k^{cp}}^i$ , Eq. (5.35) e (5.36);
24     for  $j = 1 : m$  do
25       Passo 15: Calcular o nvel de incerteza da srie temporal  $I_{adp}^j$ , Eq.
(5.47)
26       Passo 16: Calcular  $f_j(I_{adp}^j)$ , Eq. (5.48)-(5.49);
27       Passo 17: Obter os novos limites de  $\tilde{m}$ , Eq. (5.45)-(5.46);
28       Passo 18: Calcular os graus de pertincia que sero utilizados para
gerar a estimativa do modelo  $\tilde{\mu}_{\mathbf{Z}_k^{cp}}^{i(j)}$ , Eq. (5.50).
29     end
30   end
31 end

```

onde \mathbf{u}_k  a k -sima medio das variveis de entrada do sistema, e \mathbf{Z}_k  definida por:

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{u}_k \\ \mathbf{y}_k^p \end{bmatrix}. \quad (5.52)$$

Sabe-se que para obter os parmetros de Markov do observador  necessrio calcular a pseudo-inversa de $\mathbf{\Upsilon}$. Assim, a quantidade de parmetros de Markov  definida pelo

número de regressores e a quantidade de variáveis de entrada e saída (JUANG et al., 1993). Dessa forma, espera-se que quanto maior a quantidade de regressores, melhor será a resposta do modelo. Porém, devido a limitações, sejam elas por custo computacional ou a presença de espaço nulo na matriz, o número de regressores não pode ser muito alto (JUANG, 1994). Dessa forma, é importante questionar se existe um valor ideal para p em cada conjunto de dados. Para solucionar tal questionamento, a primeira ação é determinar a quantidade máxima de regressores possíveis para a aplicação p_m . Para isso, deve-se montar a matriz Υ , com uma quantidade inicial de regressores $p_m < 7$, e em seguida, deve-se calcular a seguinte matriz auxiliar:

$$\tilde{\mathfrak{R}}^i = \Upsilon \tilde{\mathbf{W}}^i \Upsilon^T, \quad (5.53)$$

onde $\tilde{\mathfrak{R}}^i$ é a matriz de correlação de dados, e $\tilde{\mathbf{W}}^i$ é a matriz de ponderação nebulosa, a qual é definida por:

$$\tilde{\mathbf{W}}^i = \begin{bmatrix} \tilde{\mu}_{\mathbf{Z}_{p_m}^{cp}}^i & 0 & \cdots & 0 \\ 0 & \tilde{\mu}_{\mathbf{Z}_{p_m+1}^{cp}}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\mu}_{\mathbf{Z}_{N-1}^{cp}}^i \end{bmatrix}. \quad (5.54)$$

Para obter os parâmetros de Markov, deve-se calcular a inversa de $\tilde{\mathfrak{R}}^i$. Porém, se esta possuir espaço nulo, a obtenção destes parâmetros é problemática, sendo assim necessário obter um valor máximo para p_m que gere a matriz $\tilde{\mathfrak{R}}^i$ sem espaço nulo. Assim, para avaliar esta condição, deve-se calcular o posto de $\tilde{\mathfrak{R}}^i$:

$$\tau = \min[\text{Rank}(\tilde{\mathfrak{R}}^{i(-)}), \text{Rank}(\tilde{\mathfrak{R}}^{i(+)})], \quad (5.55)$$

onde $\tilde{\mathfrak{R}}^{i(-)}$ e $\tilde{\mathfrak{R}}^{i(+)}$ são respectivamente os limites inferior e superior de $\tilde{\mathfrak{R}}^i$. Se τ for igual ao número de colunas de matriz de dados, então, a quantidade máxima de regressores p_m para esta aplicação esta definida. Caso contrário, deve-se atualizar o valor máximo possível para p_m . Como a dimensão de $\tilde{\mathfrak{R}}^i$ depende do número de variáveis de entrada, saída e da quantidade de regressores, ela não pode assumir qualquer valor. Portanto, deve-se definir uma nova estimativa para a quantidade máxima de regressores, a qual é definida por:

$$\begin{aligned} \tau &= m + p_m(m + n) \\ p_m^l &= \text{int}\left(\frac{\tau - m}{m + n}\right), \end{aligned} \quad (5.56)$$

onde int representa que apenas a parte inteira desta divisão foi selecionada. Com o novo valor de p_m^l selecionado, deve-se excluir as últimas $(m + n)(p_m^{l-1} - p_m^l)$ linhas e colunas de $\tilde{\mathfrak{R}}^i$, e então repetir este processo até gerar uma matriz de posto completo. A fim de informar as linhas e colunas selecionadas, o sobrescrito p será adicionado à variável $\tilde{\mathfrak{R}}^i$, onde p é a quantidade de regressores utilizados nesta versão. A figura 11 descreve este processo.

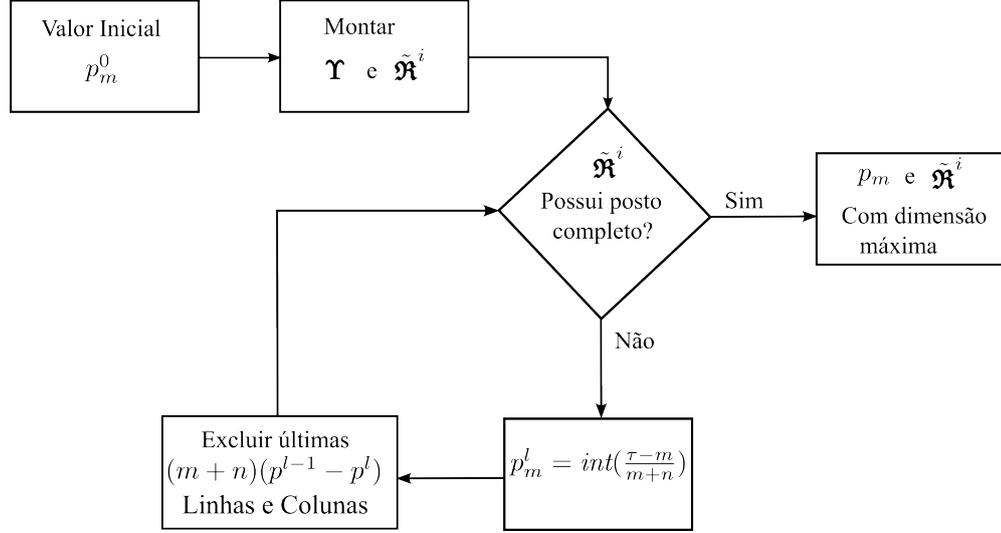


Figura 11 – Fluxograma do processo de seleção da quantidade máxima de regressores p_m .

Com o número máximo de regressores selecionados, têm-se p_m possíveis soluções para os parâmetros de Markov do sistema. Assim, o objetivo é construir o melhor conjunto solução dentre todas as possibilidades. Para isso, deve-se obter todas as possíveis soluções para os parâmetros de Markov do observador, através de:

$$\tilde{\mathbf{Y}}_p^i = \tilde{\mathfrak{G}}_p^i [\tilde{\mathfrak{R}}_p^i]^{-1}, \quad 1 \leq p \leq p_m \quad (5.57)$$

$$\tilde{\mathbf{Y}}_p^i = [\tilde{\mathbf{D}}_p^i \quad \tilde{\mathbf{C}}_p^i \tilde{\mathbf{B}}_p^i \quad \dots \quad \tilde{\mathbf{C}}_p^i \tilde{\mathbf{A}}_p^{p-1} \tilde{\mathbf{B}}_p^i] \quad (5.58)$$

onde $\tilde{\mathbf{Y}}_p^i$ são os parâmetros de Markov do observador obtidos com p regressores, e $\tilde{\mathfrak{G}}_p^i$ é a versão truncada da seguinte matriz:

$$\tilde{\mathfrak{G}}^i = \bar{\mathbf{y}}_p \tilde{\mathbf{W}}^i \mathbf{\Upsilon}^T, \quad (5.59)$$

onde suas últimas $(p_m - p)(m + n)$ colunas são desconsideradas, e $\bar{\mathbf{y}}_p$ é a matriz que contém os dados processados pela análise espectral singular, sendo definido por:

$$\bar{\mathbf{y}}_p = [\mathbf{y}_p^a \quad \mathbf{y}_{p+1}^a \quad \dots \quad \mathbf{y}_N^a]. \quad (5.60)$$

Observa-se que $\bar{\mathbf{y}}_p$ é o valor na qual a estimativa é centrada, e o produto entre $[\bar{\mathbf{y}}_p \tilde{\mathbf{W}}^i]$ determina os limites inferior e superior desejados para a estimação. Assim, o modelo identificado deve gerar uma estimativa na qual o erro quadrático médio em relação a estes limites seja o menor possível. Para determinar qual destes modelos possui a melhor aproximação, deve-se simular todas as soluções obtidas para os parâmetros de Markov através de:

$$\tilde{\hat{\mathbf{y}}}_p^i = \tilde{\mathbf{Y}}_p^i \tilde{\mathbf{\Upsilon}}_p \tilde{\mathbf{W}}_p^i, \quad (5.61)$$

onde $\tilde{\hat{\mathbf{y}}}_p^i$ é a aproximação ponderada da i -ésima regra obtida pela solução dos parâmetros de Markov com p regressores, $\tilde{\mathbf{\Upsilon}}_p$ é a versão truncada de $\mathbf{\Upsilon}$, excluindo suas últimas

$(p_m - p)(m + n)$ linhas, e $\tilde{\mathbf{W}}_p^i$ é a versão truncada da matriz de ponderação nebulosa, obtida ao excluir suas últimas $(p_m - p)(m + n)$ linhas e colunas. A equação (5.61) é implementada com os dados já obtidos, enquanto que o modelo final será utilizado para realizar uma previsão para as variáveis de saída. Por isso a importância de selecionar a melhor solução possível.

Ao analisar a equação (5.61), é notável que para estimar a primeira variável de $\tilde{\mathbf{y}}_p^i$ apenas da primeira linha de $\tilde{\mathbf{Y}}_p^i$ é necessária, o mesmo vale para as demais variáveis. Dessa forma, é possível que a melhor estimativa para cada variável pertença a diferentes soluções para os parâmetros de Markov ($\tilde{\mathbf{Y}}_p^i$). Assim, pode-se combinar as possíveis soluções de forma a construir um conjunto solução que tenha a melhor resposta dentre todos os conjuntos estimados. Para avaliar a qualidade da estimação, é utilizado o critério da informação de Akaike. Este, avalia a qualidade da estimativa através da função de verossimilhança, e a penaliza pela quantidade de parâmetros (BOZDOGAN, 2000; CAVANAUGH; NEATH, 2019). Por estar avaliando submodelos nebulosos, a formulação adotada nesta seção leva em conta os graus de pertinência utilizados para ponderar a estimação. Assim, o critério da informação de Akaike deve ser calculado por:

$$AIC_{r,p}^i = \frac{1}{\sum_{k=p}^N \mu_{\mathbf{Z}_k^{cp}}^{i-} + \mu_{\mathbf{Z}_k^{cp}}^{i+}} (\ln(\mathbb{C}_{j,p}^{2i}) + 2p), \quad (5.62)$$

onde $AIC_{r,p}^i$ é o valor do critério de Akaike obtido pela estimativa do modelo $\tilde{\mathbf{Y}}_p^i$ para sua j -ésima variável de saída, e $\mathbb{C}_{j,p}^{2i}$ é a covariância do erro de estimação obtida por esse mesmo modelo, a qual é definida por:

$$\mathbb{C}_{j,p}^{2i} = \frac{1}{\sum_{k=p}^N \mu_{\mathbf{Z}_k^{cp}}^{i-} + \mu_{\mathbf{Z}_k^{cp}}^{i+}} \sum_{k=p}^N (y_{j,k}^{i-} - \hat{y}_{j,k}^{i-})^2 + (y_{j,k}^{i+} - \hat{y}_{j,k}^{i+})^2, \quad (5.63)$$

onde $y_{j,k}^{i-}$ e $y_{j,k}^{i+}$ são respectivamente os limites inferior e superior da k -ésima observação da j -ésima variável do produto $[\bar{\mathbf{y}}_p \tilde{\mathbf{W}}^i]$, já o sobrescrito $(\hat{\cdot})$ indica que é uma estimativa para esta mesma variável. Assim, os vetores linha pertencentes aos submodelos $\tilde{\mathbf{Y}}_p^i$ que possuem o menor valor para o critério de Akaike serão selecionados para formar o conjunto solução $\tilde{\mathbf{Y}}^i$. Como tais vetores podem possuir quantidades de parâmetros diferentes, $\tilde{\mathbf{Y}}^i$ é dimensionado de forma a abranger o vetor de maior dimensão. Os demais, são incorporados à matriz ao adicionar zeros aos elementos restantes, gerando assim um conjunto solução com um formato semelhante a:

$$\tilde{\mathbf{Y}}^i = \begin{bmatrix} \tilde{Y}_{1,1}^i & \tilde{Y}_{1,2}^i & 0 & 0 \\ \tilde{Y}_{2,1}^i & \tilde{Y}_{2,2}^i & \tilde{Y}_{2,3}^i & \tilde{Y}_{2,4}^i \\ \tilde{Y}_{3,1}^i & 0 & 0 & 0 \end{bmatrix}. \quad (5.64)$$

Esta equação é apenas ilustrativa, e serve para demonstrar como o conjunto solução é armazenado.

De posse da solução para os parâmetros de Markov, o próximo passo é separar $\tilde{\mathbf{Y}}^i$ em parâmetros do sistema e do ganho do observador. Para isso, deve-se ordenar os elementos de $\tilde{\mathbf{Y}}^i$ da seguinte forma:

$$\tilde{\mathbf{Y}}_0^i = \tilde{\mathbf{D}}_k^i, \quad (5.65)$$

$$\tilde{\mathbf{Y}}_j^i = [\tilde{\mathbf{Y}}_j^{(1)i}, -\tilde{\mathbf{Y}}_j^{(2)i}], \quad (5.66)$$

onde $\tilde{\mathbf{Y}}_0^i$ corresponde aos elementos presentes nas primeiras m colunas de $\tilde{\mathbf{Y}}^i$. Com isso, pode-se obter os parâmetros do sistema e do ganho do observador, através de:

$$\tilde{\mathbf{Y}}_j^{S^i} = \begin{cases} \tilde{\mathbf{Y}}_j^{i(1)} - \sum_{l=1}^j \tilde{\mathbf{Y}}_j^{i(2)} \tilde{\mathbf{Y}}_{j-l}^{S^i}, & k = 1, \dots, q \\ -\sum_{l=1}^q \tilde{\mathbf{Y}}_j^{i(2)} \tilde{\mathbf{Y}}_{j-l}^{S^i}, & k = q + 1, \dots, \infty, \end{cases} \quad (5.67)$$

$$\tilde{\mathbf{Y}}_j^{O^i} = \begin{cases} \tilde{\mathbf{Y}}_j^{i(2)}, & k = 1, \\ \tilde{\mathbf{Y}}_j^{(2)i} - \sum_{l=1}^{j-1} \tilde{\mathbf{Y}}_j^{(2)i} \tilde{\mathbf{Y}}_{j-l}^{O^i}, & j = 2, \dots, q, \\ -\sum_{l=1}^{j-1} \tilde{\mathbf{Y}}_j^{(2)i} \tilde{\mathbf{Y}}_{j-l}^{O^i}, & j = q + 1, \dots, \infty, \end{cases} \quad (5.68)$$

onde $\tilde{\mathbf{Y}}_j^{S^i} \in \mathbb{R}^{w \times w}$ são os parâmetros de Markov do sistema, os quais serão utilizados para identificar as matrizes de estado, e $\tilde{\mathbf{Y}}_j^{O^i} \in \mathbb{R}^{w \times n}$ são os parâmetros do ganho do observador, os quais serão utilizados para determinar o ganho de Kalman.

5.4.2 ERA com Seleção Automática de Ordem

O *Eigensystem realization algorithm* é aplicado para obter uma representação no espaço de estado para os parâmetros de Markov identificados na seção anterior. Para isso, o primeiro passo é calcular a matriz de Hankel $\tilde{\mathbf{H}}(j-1)$ para j igual a 1 e 2, a qual pode ser definida pela seguinte expressão:

$$\tilde{\mathbf{H}}(j-1)^i = \begin{bmatrix} \tilde{\mathbf{Y}}_j^{S^i} & \tilde{\mathbf{Y}}_{j+1}^{S^i} & \cdots & \tilde{\mathbf{Y}}_{j+\beta-1}^{S^i} \\ \tilde{\mathbf{Y}}_{j+1}^{S^i} & \tilde{\mathbf{Y}}_{j+2}^{S^i} & \cdots & \tilde{\mathbf{Y}}_{j+\beta}^{S^i} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{Y}}_{j+\alpha-1}^{S^i} & \tilde{\mathbf{Y}}_{j+\alpha}^{O^i} & \cdots & \tilde{\mathbf{Y}}_{j+\alpha+\beta-2}^{S^i} \end{bmatrix} \quad (5.69)$$

onde α e β são, respectivamente, os números de linhas e colunas de $\tilde{\mathbf{H}}(j-1)^i$. Sabe-se que a ordem de realização mínima do modelo a ser identificado é igual ao posto de $\tilde{\mathbf{H}}(0)^i$ (JUANG, 1994). Portanto, para evitar custo computacional desnecessário, deve-se selecionar um valor ideal para α e β . Para isso, deve-se analisar os seguintes pontos:

- O modelo identificado deve ser observável, portanto, partindo de condição inicial nula e utilizando as entradas e saídas do sistema dentro de um intervalo finito de tempo $0 < k \leq t$, os estados do sistema devem ser completamente determinados (JUANG, 1994). Assim, a quantidade de estados deste modelo devem ser suficientes para sintetizar estas informações passadas.

- O modelo a ser obtido é oriundo de uma representação ARX, como apresentado na equação (5.58). Para obter a estimativa atual da saída do sistema, esta equação utiliza a entrada no instante atual, bem como suas p entradas e saídas passadas. Assim, espera-se que seu modelo no espaço de estados determine todos os seus estados dentro do intervalo finito $0 < k < p + 1$. Logo, a quantidade de regressores selecionada influencia diretamente na ordem do modelo no espaço de estado.
- O conjunto solução possui um número de regressores diferente para cada variável estimada. Assim, algumas variáveis podem ser determinadas em um intervalo de tempo inferior a $p_m + 1$.

Analisando tais pontos, percebe-se que a ordem do submodelo está relacionada com a quantidade de regressores e o número de variáveis de entrada e de saída. Dessa forma, para montar uma matriz de Hankel com a dimensão apropriada, deve-se adotar a seguinte formulação:

$$\tilde{\mathbf{H}}(j-1)^i = \begin{bmatrix} \tilde{\mathbf{Y}}_j^{S^i} & \tilde{\mathbf{Y}}_{j+1}^{S^i} & \cdots & \tilde{\mathbf{Y}}_{j+p_m}^{S^i} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{Y}}_{j+p_m}^{S^i} & \tilde{\mathbf{Y}}_{j+p_m+1}^{S^i} & \cdots & \tilde{\mathbf{Y}}_{j+2p_m}^{S^i} \end{bmatrix}. \quad (5.70)$$

Assim, percebe-se que para tal formulação, só é necessário obter $2p_m + 2$ parâmetros de Markov do sistema (os quais são obtidos pela equação (5.67)). Dessa forma, a ordem de cada submodelo pode ser definida por:

$$\iota_i = \text{Rank}(\tilde{\mathbf{H}}(0)^i), \quad (5.71)$$

onde ι_i é a ordem do i -ésimo submodelo. Observa-se que a ordem do sistema é dependente da quantidade de variáveis e do número de regressores. Assim, devido a variação de p dentro do conjunto solução, algumas variáveis necessitam de uma menor quantidade de informações passadas para serem determinadas, o que diminui a ordem de realização mínima do sistema.

Com a matriz de Hankel calculada e a ordem do sistema definida, o próximo passo é decompor $\tilde{\mathbf{H}}(0)^i$ utilizando a SVD:

$$\tilde{\mathbf{H}}(0)^i = \tilde{\mathbf{R}}^i \tilde{\mathcal{E}}^i \tilde{\mathcal{S}}^{iT}, \quad (5.72)$$

onde $\tilde{\mathbf{R}}^i \in \mathbb{R}^{wp_m \times wp_m}$ é a matriz que contém os vetores singulares a esquerda de $\tilde{\mathbf{H}}(0)^i$, $\tilde{\mathcal{E}}^i \in \mathbb{R}^{wp_m \times np_m}$ é a matriz que contém os valores singulares, e $\tilde{\mathcal{S}}^i \in \mathbb{R}^{np_m \times wp_m}$ é a matriz que contém os vetores singulares a direita de $\tilde{\mathbf{H}}(0)^i$. Para determinar as matrizes de estado do sistema, é necessário apenas as primeiras ι_i linhas e colunas de $\tilde{\mathbf{R}}^i$ e $\tilde{\mathbf{H}}(0)^i$, portanto, o subscrito ι será empregado para informar que as demais linhas forma desconsideradas

nestas matrizes. Para identificar as matrizes no espaço de estado de cada submodelo, deve-se primeiro calcular as matrizes de observabilidade ($\tilde{\mathcal{P}}_\iota$) e controlabilidade ($\tilde{\mathcal{Q}}_\iota$):

$$\tilde{\mathcal{P}}_\iota^i = \tilde{\mathbf{R}}_\iota^i \tilde{\mathcal{E}}_\iota^{i\frac{1}{2}}, \quad (5.73)$$

$$\tilde{\mathcal{Q}}_\iota = \tilde{\mathcal{E}}_\iota^{i\frac{1}{2}} \tilde{\mathcal{S}}_\iota^{iT}. \quad (5.74)$$

A partir destas, pode-se calcular as matrizes no espaço de estado, bem como o ganho de Kalman para cada submodelo:

$$\tilde{\mathbf{A}}_k^i = (\tilde{\mathcal{E}}_\iota^i)^{-1/2} \tilde{\mathbf{R}}_\iota^T \tilde{\mathbf{H}}(1)^i \tilde{\mathcal{S}}_\iota^{iT} (\tilde{\mathcal{E}}_\iota^i)^{-1/2}, \quad (5.75)$$

$$\tilde{\mathbf{B}}_k^i = \text{Primeiras } m \text{ colunas de } \tilde{\mathcal{Q}}_\iota^i, \quad (5.76)$$

$$\tilde{\mathbf{C}}_k^i = \text{Primeiras } n \text{ linhas de } \tilde{\mathcal{P}}_\iota^i, \quad (5.77)$$

$$\tilde{\mathbf{D}}_k^i = \tilde{\mathbf{Y}}_0^i \quad (5.78)$$

$$\tilde{\mathbf{K}}_k^i = -[\tilde{\mathcal{P}}_\iota^{iT} \tilde{\mathcal{P}}_\iota^i]^{-1} \tilde{\mathbf{Y}}_\iota^{o^i}. \quad (5.79)$$

5.4.3 Atualização Recursiva e Refinada

O modelo inicial foi determinado na seção anterior. Este pode ser utilizado para gerar uma estimativa de 1 passo à frente (instante $k + 1$), ou simulação livre (usado em caso de falta de medições). Nesse sentido, para manter a qualidade do modelo obtido, deve-se atualizar o conjunto solução a cada nova medição disponível, mantendo atenção para possíveis alterações no número ideal de regressores. Assim, deve-se buscar por uma atualização recursiva para tais parâmetros.

O primeiro passo desse processo é montar o vetor de regressores \mathbf{v}_k :

$$\mathbf{v}_k = \begin{bmatrix} \mathbf{u}_{k+1} \\ \mathbf{Z}_k \\ \mathbf{Z}_{k-1} \\ \vdots \\ \mathbf{Z}_{k-p} \end{bmatrix}, \quad 1 \leq p \leq p_m \quad (5.80)$$

o qual é equivalente a uma nova coluna da matriz Υ decorrente da chegada de uma nova medição. Em seguida, o vetor de ganho ($\tilde{\mathfrak{F}}_p^i(k)$) deve ser calculado por:

$$\tilde{\mathfrak{F}}_p^i(k) = \frac{\tilde{\mu}_{\mathbf{Z}_k}^i \mathbf{v}_k^T \tilde{\mathfrak{F}}_p^i(k-1)}{1 + \tilde{\mu}_{\mathbf{Z}_k}^i \mathbf{v}_k^T \tilde{\mathfrak{F}}_p^i(k-1) \mathbf{v}_k}, \quad (5.81)$$

onde $\tilde{\mathfrak{F}}_p^i$ é a inversa da matriz de correlação de dados $\tilde{\mathfrak{R}}_p^i$. Portanto, esta deve ser atualizada através de:

$$\tilde{\mathfrak{F}}_p^i(k) = \tilde{\mathfrak{F}}_p^i(k-1) [\mathbf{I} - \tilde{\mu}_{\mathbf{Z}_k}^i \mathbf{v}_k \tilde{\mathfrak{F}}_p^i(k)], \quad (5.82)$$

onde \mathbf{I} é uma matriz identidade de ordem apropriada. Neste ponto, os parâmetros de Markov do observador podem ser atualizados recursivamente:

$$\tilde{\mathbf{Y}}_p^i(k+1) = \tilde{\mathbf{Y}}_p^i(k) + \tilde{\boldsymbol{\delta}}_p^i(k+1)\tilde{\boldsymbol{\mathfrak{F}}}_p^i(k), \quad (5.83)$$

onde $\tilde{\boldsymbol{\delta}}_p^i$ é o vetor de resíduo de estimação obtido pelo submodelo i com p regressores. Este é calculado por:

$$\tilde{\boldsymbol{\delta}}_p^i = \tilde{\boldsymbol{\mu}}_{\mathbf{Z}_k^{cp}}^i[\mathbf{y}(k+1) - \tilde{\hat{\mathbf{y}}}^i(k+1)], \quad (5.84)$$

onde $\mathbf{y}(k+1)$ é a saída esperada (obtida pela AES), e $\tilde{\hat{\mathbf{y}}}(k+1)$ é a saída do submodelo, obtida por:

$$\tilde{\hat{\mathbf{y}}}_p^i(k+1) = \tilde{\mathbf{Y}}_p^i(k)\mathbf{v}_k. \quad (5.85)$$

Os parâmetros de Markov obtidos serão utilizados para atualizar o modelo no espaço de estado. Portanto, tais parâmetros não podem variar mais rápido do que os estados estimados. Para evitar tal cenário, a variação dos parâmetros será submetida a um filtro passa-baixa, obtendo assim a seguinte atualização refinada:

$$\tilde{\mathbf{Y}}_p^{r_i}(k+1) = \tilde{\mathbf{Y}}_p^i(k+1) + \frac{1}{T_d}(\tilde{\mathbf{Y}}_p^i(k+1) - \tilde{\mathbf{Y}}_p^i(k)), \quad (5.86)$$

onde $\tilde{\mathbf{Y}}_p^{r_i}(k+1)$ é a atualização refinada para o i -ésimo submodelo obtida com p regressores, T_d é uma constante de tempo definida pelo projetista. Após a primeira atualização, a solução refinada será utilizada na equação (5.85) para obter as estimativas dos submodelos.

Com os modelos atualizados, deve-se montar o conjunto solução com a melhor resposta segundo o critério da informação de Akaike. Para avaliar tal quesito, deve-se atualizar recursivamente a covariância do erro de estimação:

$$\mathbb{C}_{j,p}^{2^i} = \frac{1}{\varrho_u^i + \mu_{\mathbf{Z}_k^{cp}}^{i-} + \mu_{\mathbf{Z}_k^{cp}}^{i+}}[\varrho_\delta^i + (\delta_{j,p}^{i-}(k+1))^2 + (\delta_{j,p}^{i+}(k+1))^2], \quad (5.87)$$

onde ϱ_u^i e ϱ_δ^i são respectivamente os valores anteriores para a soma dos graus de pertinência e resíduo de estimação, sendo expresso por:

$$\varrho_u^i = \sum_{j=p}^k \mu_{\mathbf{Z}_j^{cp}}^{i-} + \mu_{\mathbf{Z}_j^{cp}}^{i+}, \quad (5.88)$$

$$\varrho_\delta^i = \sum_{j=p}^k (y_{j,k}^{i-} - \hat{y}_{j,k}^{i-})^2 + (y_{j,k}^{i+} - \hat{y}_{j,k}^{i+})^2. \quad (5.89)$$

Com a covariância do ruído de estimação atualizada, pode-se usar a equação (5.62) para atualizar o valor do critério da informação de Akaike para todas as possíveis soluções, e então, escolher aquelas com menor valor. Após formar o novo conjunto solução, deve-se extrair os parâmetros do sistema e do ganho do observador através das equações (5.65)-(5.68).

5.4.3.1 Mecanismo de Variação Temporal de Ordem

Após a atualização do conjunto solução e a extração dos parâmetros do sistema e do ganho do observador, a representação no espaço de estado deve ser obtida. Nesse sentido, os seguintes cenários são possíveis:

- A configuração do conjunto solução permanece inalterada, ou seja, os submodelos com menor valor para o critério de Akaike permanecem os mesmos após a atualização.
- O valor do critério de Akaike sugere que a melhor solução para uma ou mais variáveis pertence a um submodelo diferente do inicial. Portanto, a configuração do conjunto solução deve ser alterada.

Na primeira situação, ocorre apenas uma adaptação dos parâmetros de acordo com o erro de estimação dos modelos. Assim, a ordem do modelo permanece a mesma, e o modelo atualizado pode ser obtido repetindo o processo apresentado nas equações (5.70)-(5.79).

Na segunda situação, a configuração do conjunto solução pode variar, ou seja, para algumas variáveis, a melhor solução pode ser obtida com um número de regressores diferente da solução anterior. Nesse caso, os estados do modelo anterior não serão mais os mesmos, pois devem sintetizar uma quantidade diferente de informações passadas. Assim, após calcular a matriz de Hankel $\tilde{\mathbf{H}}(0)_{k+1}^i$, deve-se atualizar a ordem do sistema através de:

$$l_i^{k+1} = \text{Rank}(\tilde{\mathbf{H}}(0)_{k+1}^i). \quad (5.90)$$

Após a atualização da ordem, deve-se obter a nova realização no espaço de estados através das equações (5.72)-(5.79). Como os estados deste novo sistema são diferentes dos estados do modelo anterior, é necessário simular o sistema estimado com suas $p + 1$ entradas e saídas anteriores para determinar todos os seus estados. Em seguida, deve-se calcular a estimativa do modelo ao aplicar o conjunto de medições atual.

Ao aplicar a metodologia no rastreamento e previsão de séries temporais, a variação temporal da ordem indica que o modelo estimado se adapta às características e necessidades locais da base de dados, alterando a ordem para obter o melhor resultado entre os modelos possíveis. No entanto, espera-se um comportamento diferente para sistemas físicos, onde suas variáveis de entrada e saída são relacionadas. Isso ocorre, pois tais sistemas possuem uma ordem fixa.

O algoritmo 3 apresenta de forma resumida os passos computacionais desenvolvidos na estimação paramétrica do consequente.

Algoritmo 3: Estimaco paramtrica do conseqente

```

1 Entrada:  $y^a$ ,  $\mathbf{Z}$ ,  $\tilde{\mathbf{U}}$  e  $p_0$ .
2 Saída:  $\tilde{\mathbf{A}}_k^i$ ,  $\tilde{\mathbf{B}}_k^i$ ,  $\tilde{\mathbf{C}}_k^i$ ,  $\tilde{\mathbf{D}}_k^i$  e  $\tilde{\mathbf{K}}_k^i$ .
3 %Etapa Offline
4 for  $i = 1 : c$  do
5     Passo 1: Montar a matriz de correlaco de dados  $\tilde{\mathfrak{R}}^i$  com dimenso mxima,
        como mostrado na figura 11;
6     for  $1 \leq p \leq p_m$  do
7         Passo 2: Calcular os parmetros de Markov do observador  $\tilde{\mathbf{Y}}_p^i$ , Eq.
            (5.57)-(5.59);
8         Passo 3: Simular o modelo ARX, Eq. (5.61);
9         Passo 4: Montar o conjunto soluo com base no critrio de Akaike  $\tilde{\mathbf{Y}}^i$ , Eq.
            (5.62)-(5.63);
10    end
11    Passo 5: Montar a matriz de Hankel  $\tilde{\mathbf{H}}(0)^i$ , Eq. (5.70);
12    Passo 6: Determinar a ordem do sistema  $\tau_i$ , Eq. (5.71);
13    Passo 7: Estimar as matrizes  $\tilde{\mathbf{A}}_k^i$ ,  $\tilde{\mathbf{B}}_k^i$ ,  $\tilde{\mathbf{C}}_k^i$ ,  $\tilde{\mathbf{D}}_k^i$  e  $\tilde{\mathbf{K}}_k^i$ , Eq. (5.75)-(5.79);
14 end
15 %Etapa online  $k + 1$ 
16 for  $i = 1 : c$  do
17     Passo 8: Montar o vetor de regressores  $\mathbf{v}_k$ , Eq. (5.80);
18     for  $1 \leq p \leq p_m$  do
19         Passo 9: Montar o vetor de resduos  $\tilde{\boldsymbol{\delta}}_p^i$ , Eq. (5.84)-(5.85);
20         Passo 10: Obter a soluo refinada  $\tilde{\mathbf{Y}}_p^{r^i}(k + 1)$ , Eq. (5.81)-(5.83) e (5.86);
21         Passo 11: Atualizar o critrio de Akaike e atualizar o conjunto soluo, Eq.
            (5.87)-(5.89);
22     end
23     Passo 12: Montar a matriz de Hankel  $\tilde{\mathbf{H}}(0)^i$ , Eq. (5.70);
24     if Houve alteraco no configurao de  $\tilde{\mathbf{Y}}^i$  then
25         Passo 13: Atualizar a ordem do modelo  $\tau_i^{k+1}$ , Eq. (5.90);
26         Passo 14: Estimar  $\tilde{\mathbf{A}}_k^i$ ,  $\tilde{\mathbf{B}}_k^i$ ,  $\tilde{\mathbf{C}}_k^i$ ,  $\tilde{\mathbf{D}}_k^i$  e  $\tilde{\mathbf{K}}_k^i$ ;
27         Passo 15: Determinar os novos estados simulando o sistema com suas  $p + 1$ 
            entradas e sadas anteriores;
28         else
29             Passo 16: Atualizar os parmetros no espao de estado, Eq.
                (5.75)-(5.79).
30         end
31     end
32 end

```

5.5 Simulação do Modelo

Após a estimativa de todos os parâmetros, o modelo obtido é utilizado para realizar a estimativa acerca da variável de saída. Nesta dissertação, os seguintes cenários são explorados:

- **Simulação um passo a frente:** quando novas medições estão disponíveis, deve-se utilizá-las para atualizar os parâmetros do modelo e então realizar a estimação da variável de saída para o próximo instante de tempo (AGUIRRE, 2007).
- **Simulação mista:** quando ocorre um problema com os sensores e novas medições não estão disponíveis, deve-se utilizar o modelo disponível para estimar os valores da variável de saída até que novas medições estejam disponíveis.

5.5.1 Simulação um Passo a Frente

Neste tipo de simulação, todas as informações estão disponíveis. Assim, com base na medição para variável de saída \mathbf{y}_{k-1} e do ganho de Kalman, pode-se propagar os estados e obter o vetor $\tilde{\mathbf{x}}_k^i$. Portanto, a estimativa para o próximo instante de tempo \mathbf{y}_k pode ser obtida utilizando \mathbf{u}_k , em que \mathbf{y}_k é a saída do i -ésimo submodelo. Para obter a saída global do sistema, utiliza-se os graus de pertinência para ponderar os submodelos. Nesse sentido, esta metodologia utiliza o grau de pertinência intervalar com parâmetro de incerteza adaptado pela Eq. (5.50). Portanto, a formulação para a saída global do modelo é dada por:

$$\tilde{\mathbf{y}}_k = \sum_{i=1}^c \tilde{\mathbf{u}}^i \tilde{\mathbf{y}}_k^i, \quad (5.91)$$

$$\tilde{\mathbf{u}}^i = \begin{bmatrix} \tilde{\mu}_{\mathbf{Z}_k}^{i(1)} & 0 & \cdots & 0 \\ 0 & \tilde{\mu}_{\mathbf{Z}_k}^{i(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\mu}_{\mathbf{Z}_k}^{i(m)} \end{bmatrix} \quad (5.92)$$

onde $\tilde{\mu}_{\mathbf{Z}_k}^{i(m)}$ é o grau de pertinência adaptado pelo nível de incerteza da m -ésima variável de saída.

5.5.2 Simulação Livre

A simulação livre é realizada quando não há novas medições da variável de saída. Nesse caso, as estimativas dos estados do sistema não podem ser corrigidas com base em dados reais. Para gerar as estimativas, o primeiro passo é determinar os graus de pertinência. Como não há novas medições, deve-se montar o vetor \mathbf{Z}_k utilizando o ponto

médio do intervalo da última estimativa como variável de saída:

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{u}_k \\ \frac{1}{2}(\hat{\mathbf{y}}_k^+ + \hat{\mathbf{y}}_k^-) \end{bmatrix}. \quad (5.93)$$

Após esta etapa, deve-se extrair as componentes principais e então calcular o grau de pertinência adaptado a cada variável. Como não há novas medições, a covariância do ruído de medição não pode ser atualizada. Portanto, durante a simulação livre, usa-se a última atualização para o expoente de ponderação nebulosa \tilde{m}_j para calcular os graus de pertinência, conforme a equação (5.50).

Para estimar os estados, os parâmetros do consequente foram obtidos de forma ponderada. Portanto, os submodelos cujos os graus de pertinência são baixos tendem a ter uma resposta menos precisa que a dos demais modelos. A longo prazo, estes erros se acumulam, fazendo com que a estimativa se distancie cada vez mais do valor ideal, o que causa problemas na estimativa global do modelo. Assim, deve-se buscar uma forma de corrigir os estados dos submodelos. Para isso, observe que, para os algoritmos de agrupamento baseado em critério de máxima verossimilhança, os graus de pertinência podem ser interpretados como a probabilidade de selecionar o submodelo dado a variável linguística do antecedente \mathbf{Z}_k (BABUŠKA, 2012). Portanto, a média ponderada entre a saída dos submodelos é uma estimativa mais provável para a variável de saída:

$$\hat{\mathbf{y}}_{ref} = \frac{\sum_{i=1}^c \mu_{\mathbf{Z}_k}^{i+} \hat{\mathbf{y}}_k^{i+} + \mu_{\mathbf{Z}_k}^{i-} \hat{\mathbf{y}}_k^{i-}}{\sum_{i=1}^c \mu_{\mathbf{Z}_k}^{i+} + \mu_{\mathbf{Z}_k}^{i-}}. \quad (5.94)$$

Esta variável será utilizada para simular o cálculo do resíduo de estimação, e corrigir os estados através do ganho de Kalman, como apresentado a seguir:

$$\tilde{\mathbf{x}}_k^i = \tilde{\mathbf{A}}_k^i \tilde{\mathbf{x}}_{k-1}^i + \tilde{\mathbf{B}}_k^i \mathbf{u}_{k-1} + \tilde{\mathbf{K}}_k^i (\hat{\mathbf{y}}_{ref} - \tilde{\mathbf{y}}_{k-1}^i). \quad (5.95)$$

Com os estados atualizados, deve-se utilizar a equação (5.91) para obter a saída do modelo.

A partir do momento em que novas medições estão disponíveis, deve-se realizar a simulação livre por mais p amostras e coletar as medições durante este intervalo de tempo. Após este período, deve-se utilizar as medições coletadas para montar o vetor de regressores e realizar a atualização recursiva do modelo.

Parte IV

RESULTADOS

6 ANÁLISE COMPUTACIONAL

Para a análise computacional da metodologia proposta, um conjunto de dados simulados acerca da formação PRISMA de satélites foi desenvolvido. Este conjunto de dados possui medições simuladas acerca da posição e velocidade relativa entre os satélites no referencial *Local-Vertical-Local-Horizontal* (LVLH).

6.1 Previsão intervalar multivariável de posição e velocidade relativas entre satélites em formação PRISMA

A formação PRISMA consiste em dois satélites voando de forma coordenada e relativamente próxima (D'AMICO et al., 2012). Tais satélites são, então, denominados:

1. Tango, o qual serve como o *target* ou referência para a formação;
2. Mango, o qual serve como o *chaser* da formação.

6.1.1 Formulação matemática de satélites em formação PRISMA

Para simular o movimento dos satélites, faz-se necessário modelar a dinâmica de seus movimentos através da seguinte equação diferencial:

$$\ddot{\mathbf{r}} = \mathbf{\bar{a}} + \sum_{t=1}^4 \mathbf{\bar{a}}_t, \quad (6.1)$$

onde $\mathbf{r} \in \mathbb{R}^3$ é o vetor posição, os termos $\mathbf{\bar{a}}_t \in \mathbb{R}^3$ são perturbações não Keplerianas observadas no movimento, e $\mathbf{\bar{a}} \in \mathbb{R}^3$ representa a equação da gravitação universal de Newton, a qual é modelada por:

$$\mathbf{\bar{a}} = -\zeta_t \left(\frac{\mathbf{r}}{|\mathbf{r}|^3} \right), \quad (6.2)$$

onde ζ_t é a constante gravitacional da Terra. Na formulação desenvolvida, são adicionadas quatro perturbações, a primeira é referente ao formato oblíquo da Terra $\mathbf{\bar{a}}_1$ (BLITZER; WHEELON, 1957), a qual é modelada por:

$$a_{1_x} = \frac{\zeta_t r_x}{|\mathbf{r}|^3} \left(\frac{3}{2} J_2 \right) \left(\frac{R_E}{|\mathbf{r}|} \right)^2 \left(5 \frac{r_z^2}{|\mathbf{r}|^2} - 1 \right), \quad (6.3)$$

$$a_{1_y} = \frac{\zeta_t r_y}{|\mathbf{r}|^3} \left(\frac{3}{2} J_2 \right) \left(\frac{R_E}{|\mathbf{r}|} \right)^2 \left(5 \frac{r_z^2}{|\mathbf{r}|^2} - 1 \right), \quad (6.4)$$

$$a_{1_z} = \frac{\zeta_t r_z}{|\mathbf{r}|^3} \left(\frac{3}{2} J_2 \right) \left(\frac{R_E}{|\mathbf{r}|} \right)^2 \left(5 \frac{r_z^2}{|\mathbf{r}|^2} - 3 \right), \quad (6.5)$$

$$\vec{\mathbf{a}}_j = \begin{bmatrix} a_{1_x} \\ a_{1_y} \\ a_{1_z} \end{bmatrix}, \quad (6.6)$$

onde $J_2 = 1082.625 \times 10^{-6}$ é um parâmetro planetário, R_E é o raio da Terra, r_x , r_y , e r_z são respectivamente as componentes nos eixos X , Y , e Z do vetor posição $\vec{\mathbf{r}}$. A segunda perturbação é relativa a influência de um terceiro corpo no movimento $\vec{\mathbf{a}}_2$ (ROSCOE; VADALI; ALFRIEND, 2013). Neste caso, é considerada a influência que o Sol e a Lua exercem no movimento dos satélites. Tal perturbação é modelada matematicamente através de:

$$\vec{\mathbf{a}}_2 = \zeta_i \left(\frac{\vec{\mathbf{r}}_{mi} - \vec{\mathbf{r}}}{|\vec{\mathbf{r}}_{mi} - \vec{\mathbf{r}}|^3} - \frac{\vec{\mathbf{r}}_{mi}}{|\vec{\mathbf{r}}_{mi}|^3} \right), \quad (6.7)$$

onde $\vec{\mathbf{r}}_{mi}$ é a posição do terceiro corpo (Sol ou Lua) no referencial inercial da Terra. A terceira perturbação é causada pela força de arrasto atmosférico (MORGAN et al., 2012; HATTEN; RUSSELL, 2017), a qual é representada matematicamente por:

$$\vec{\mathbf{a}}_3 = -0.5 \rho_{atm} |\vec{\mathbf{v}}_r|^2 \frac{\mathcal{C}_d \mathcal{A}_d}{m_s} \frac{\vec{\mathbf{v}}_r}{|\vec{\mathbf{v}}_r|}, \quad (6.8)$$

onde \mathcal{C}_d é o coeficiente de arrasto, \mathcal{A}_d é a área exposta da superfície do satélite, m_s é a massa do satélite, $\vec{\mathbf{v}}_r$ é o vetor velocidade do satélite em relação a atmosfera, e ρ_{atm} é a densidade atmosférica. A última perturbação é causada pela pressão de radiação solar (VOKROUHLICKY; FARINELLA; MIGNARD, 1993), sendo modelada por:

$$\vec{\mathbf{a}}_4 = -V_{il} P_N \frac{C_r A_s}{m_s} (AU^2) \frac{\vec{\mathbf{r}}_{sun} - \vec{\mathbf{r}}}{|\vec{\mathbf{r}}_{sun} - \vec{\mathbf{r}}|^3}, \quad (6.9)$$

onde P_N é a pressão de radiação solar, C_r é o coeficiente de radiação solar, A_s é a área da superfície exposta à radiação solar, AU é a unidade astronômica, $\vec{\mathbf{r}}_{sun}$ é o vetor posição do Sol em relação a Terra, e V_{il} é o fator de iluminação, o qual recebe o valor unitário quando o satélite esta iluminado pelo Sol, e o valor nulo para os demais casos.

Todos os termos modelados nas equações (6.2)-(6.9) são acelerações causadas pela influência de alguma força. Assim, para obter as posições e velocidade de cada satélite, deve-se integrar duas vezes a equação (6.1), utilizando como integrador numérico o algoritmo Runge-Kutta de ordem 4. Nesse ponto, têm-se a velocidade e posição de cada satélite no referencial *Earth-Centred Inertial* (ECI). Para simular um processo de medição, as variáveis são corrompidas por ruído branco, conforme apresentado por:

$$\vec{\mathbf{r}}_{co} = \vec{\mathbf{r}} + \mathcal{N}(0, h_r), \quad h_r = 1.2m, \quad (6.10)$$

$$\vec{\mathbf{v}}_{co} = \vec{\mathbf{v}} + \mathcal{N}(0, h_v), \quad h_v = 0.03m/s, \quad (6.11)$$

onde $\mathcal{N}(0, h)$ representa uma variável aleatória com distribuição normal e desvio padrão h , enquanto que $\vec{\mathbf{v}}_{co}$ e $\vec{\mathbf{r}}_{co}$ são respectivamente os vetores velocidade e posição corrompidos por ruído. Para obter os vetores de velocidade ($\vec{\mathbf{v}}_m$) e posição ($\vec{\mathbf{r}}_m$) relativa, as medições

devem ser convertidas para o referencial *Local-Vertical Local-Horizontal* (LVLH). Este processo é representado por:

$$\vec{\mathbf{r}}_m = \mathcal{S}^T(\vec{\mathbf{r}}_c - \vec{\mathbf{r}}_t), \quad (6.12)$$

$$\vec{\mathbf{v}}_m = \mathcal{S}^T(\vec{\mathbf{v}}_c - \vec{\mathbf{v}}_t - \mathcal{W}(\vec{\mathbf{v}}_c - \vec{\mathbf{r}}_t)), \quad (6.13)$$

onde

$$\mathcal{S} = \begin{bmatrix} \frac{\vec{\mathbf{r}}_t}{|\vec{\mathbf{r}}_t|}, & \frac{\vec{\mathbf{r}}_t \times \vec{\mathbf{v}}_t}{|\vec{\mathbf{r}}_t \times \vec{\mathbf{v}}_t|} \times \frac{\vec{\mathbf{r}}_t}{|\vec{\mathbf{r}}_t|}, & \frac{\vec{\mathbf{r}}_t \times \vec{\mathbf{v}}_t}{|\vec{\mathbf{r}}_t \times \vec{\mathbf{v}}_t|} \end{bmatrix}, \quad (6.14)$$

$$\mathcal{W} = \frac{\mathbf{r}^\times(\vec{\mathbf{v}}_c - \vec{\mathbf{r}}_t)}{\|\mathbf{r}^\times\|^2}, \quad (6.15)$$

$$\mathbf{r}^\times = \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix}, \quad (6.16)$$

onde \mathbf{r}_t é o vetor posição do *Target*, \mathbf{r}_c é o vetor posição do *Chaser*, \mathbf{v}_t é o vetor velocidade do *Target* e \mathbf{v}_c é o vetor velocidade do *Chaser*. Este processo de conversão de referenciais pode alterar as características do ruído.

Para simular o processo descrito acima, faz-se necessário utilizar o ambiente *MatLab/Simulink*. Nesse sentido, a figura 12 apresenta o diagrama de blocos desta operação.

Para validar a metodologia proposta, 100 realizações ruidosas foram testadas. Nesse sentido, as figuras 13 e 14 apresentam, respectivamente um exemplo acerca das medições da posição e velocidade relativa entre os satélites em formação; enquanto que a figura 15 apresenta a distribuição do ruído de medição presente em cada uma das variáveis de saída.

As distribuições de ruído apresentadas na figura 15 apresentam características diferentes entre o eixo Z e os demais eixos. Ao aplicar o teste de Kolmogorov-Smirnov (BERGER; ZHOU, 2014), apenas o eixo Z foi classificado como uma distribuição normal. Os demais eixos possuem média zero, porém suas covariâncias não são constantes, o que pode ser observado claramente na figura 14.

Para esta aplicação, foram realizados dois testes, um para estimar a posição relativa dos satélites e outro para estimar a velocidade relativa. Para ambos os testes, as primeiras 5900 medições foram utilizadas para treinar o modelo *offline*, e as 5900 medições restantes foram utilizadas para testar o modelo *online*. Por não possuir variável de entrada externa, será utilizada a medição com um atraso puro de um período amostral como entrada para o sistema:

$$\mathbf{u}_k = \mathbf{y}_{k-1}. \quad (6.17)$$

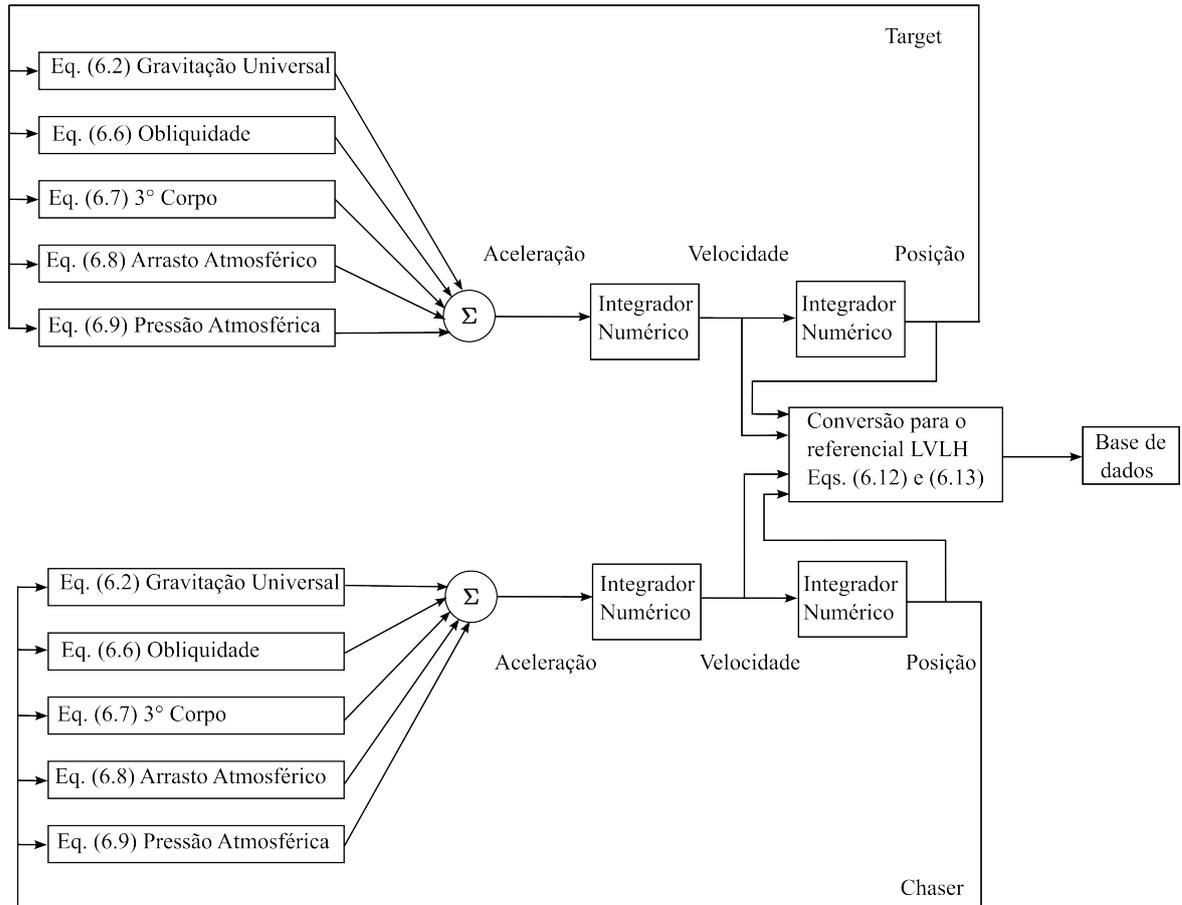


Figura 12 – Diagrama de blocos para a simulação do movimento dos satélites da formação PRISMA.

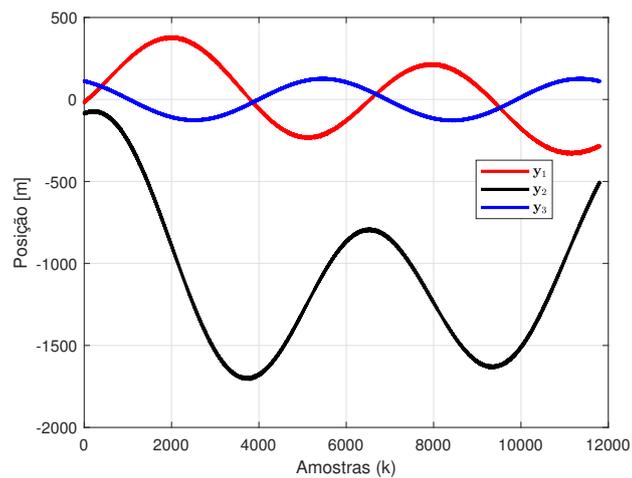


Figura 13 – Posição relativa entre os satélites em formação PRISMA. A variável y_1 representa o eixo X do vetor posição, y_2 representa o eixo Y deste vetor, e y_3 representa o eixo Z .

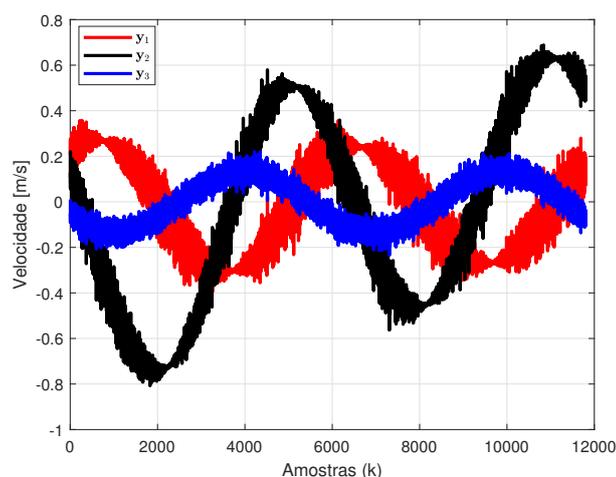


Figura 14 – Posição relativa entre os satélites em formação. A variável y_1 representa o eixo X do vetor velocidade, y_2 representa o eixo Y deste vetor, e y_3 representa o eixo Z.

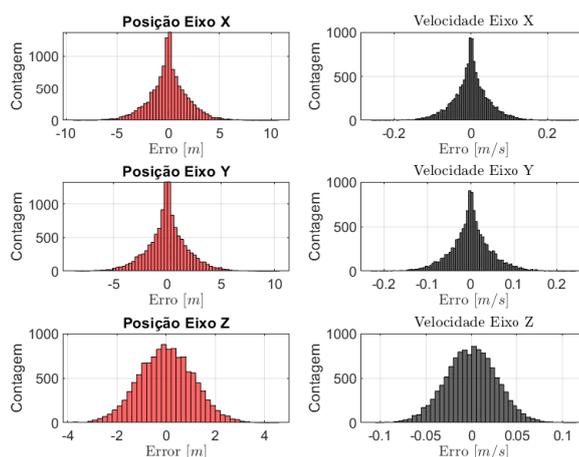


Figura 15 – Distribuição do ruído de medição presente em cada uma das variáveis de saída.

6.1.2 Posição Relativa

Para a realização dos testes, as 100 realizações foram inicializadas pelos mesmos parâmetros iniciais. Estes, são apresentados na tabela 1.

Tabela 1 – Parâmetros iniciais para a estimação paramétrica da metodologia.

Parâmetros	Descrição	Valor
c	Quantidade de Conjuntos	5
ϵ	Critério de Convergência	10^{-4}
N	Dimensão do conjunto de dados inicial	5900
δ	Número de Linhas da Matriz Trajetória	700

Após a etapa de pré-processamento, foram extraídas as componentes espectrais

mais significativas em cada variável. Esta informação é apresentada na tabela 2.

Tabela 2 – Quantidade de componentes espectrais selecionadas.

Variável	Componentes extraídas
Posição no Eixo X y_1	3
Posição no Eixo Y y_2	3
Posição no Eixo Z y_3	2

A variável linguística do antecedente é formada por seis variáveis: três variáveis de entrada e três variáveis de saída. A figura 16 apresenta a partição da base de dados *offline* decomposta em duas componentes principais, que representam mais de 99% da variação total do conjunto. Completando tal informação, a figura 17 apresenta as funções de pertinência intervalar obtida pelo ANT-2IEMV. Os resultados apresentados nestas figuras foram obtidos pela realização com o resultado mais próximo à média obtida entre as 100 realizações.

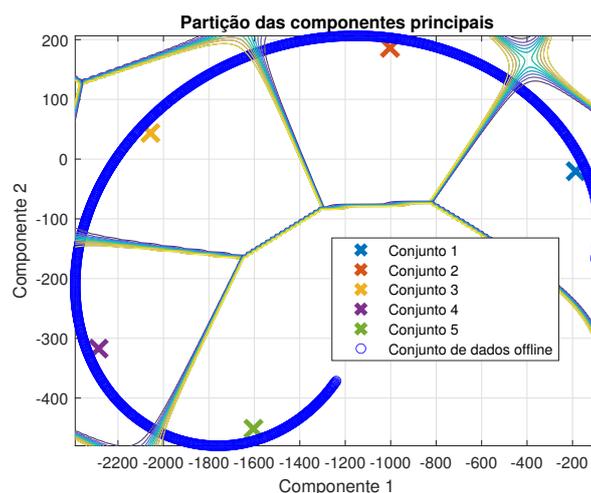


Figura 16 – Partição nebulosa para o conjunto de dados da posição relativa com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.

6.1.2.1 Simulação um Passo a Frente

Após a partição inicial dos dados e a estimação do modelo inicial, a metodologia proposta estima as variáveis de saída referente à próxima amostra e atualiza os parâmetros do filtro de Kalman nebuloso tipo-2 intervalar. Nesse sentido, as figuras 18-20 apresentam a comparação entre o valor real das variáveis de saída e o intervalo de estimação das mesmas.

Nas figuras 18 e 20, é possível observar com alguma clareza a diferença entre o intervalo estimado e a variável em questão. Porém, para a figura 19, esta diferença é quase

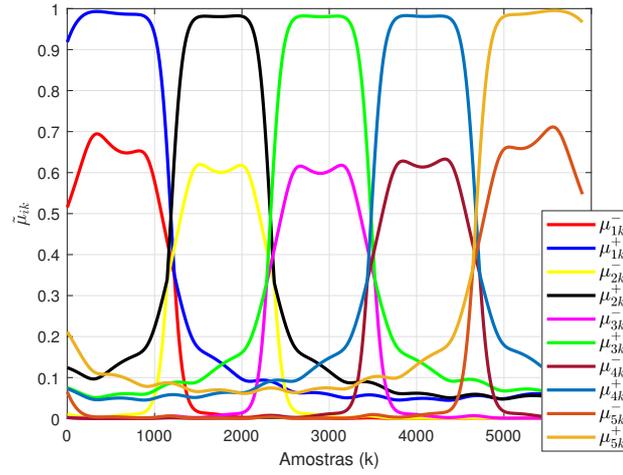


Figura 17 – Funções de pertinência intervalar para os cinco conjuntos nebulosos estimados.

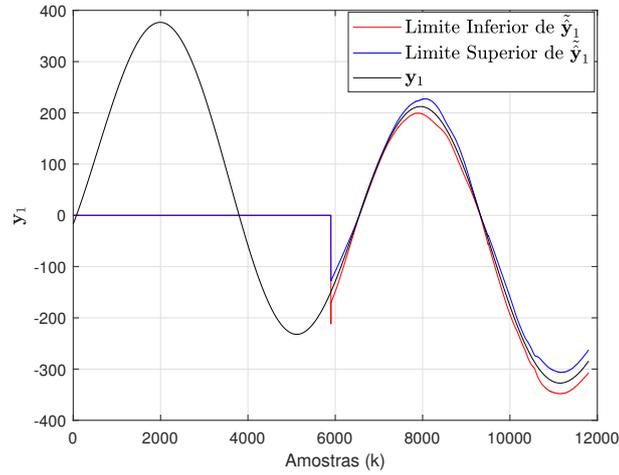


Figura 18 – Comparação entre o intervalo estimado e o valor real da distância no eixo X entre os satélites da formação PRISMA no referencial $LVLH$.

imperceptível. Assim, para avaliar a qualidade da resposta do modelo, os critérios do erro 3D-RMS (e_{3D}) e da acurácia (ACC) foram utilizados. Tais critérios são definidos por:

$$e_{3D} = \sqrt{\frac{1}{N} \sum_{i=1}^N [e_x^2 + e_y^2 + e_z^2]}, \quad (6.18)$$

$$ACC = \frac{\text{Previsões Corretas}}{\text{Total de Previsões}}, \quad (6.19)$$

onde e_x é o erro da estimativa na componente X . Nesse sentido, a tabela 3 apresenta as duas métricas citadas acima, e as compara com outras metodologias presentes na literatura.

Os trabalhos referenciados na tabela 3 consistem em diferentes aplicações do filtro de Kalman. Os modelos Q-MLE-AEKF e Q-FAEKF apresentados em Fraser e

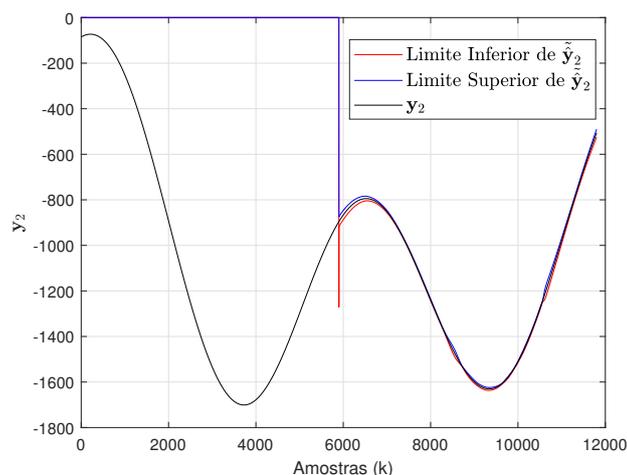


Figura 19 – Comparação entre o intervalo estimado e o valor real da distância no eixo Y entre os satélites da formação PRISMA no referencial $LVLH$.

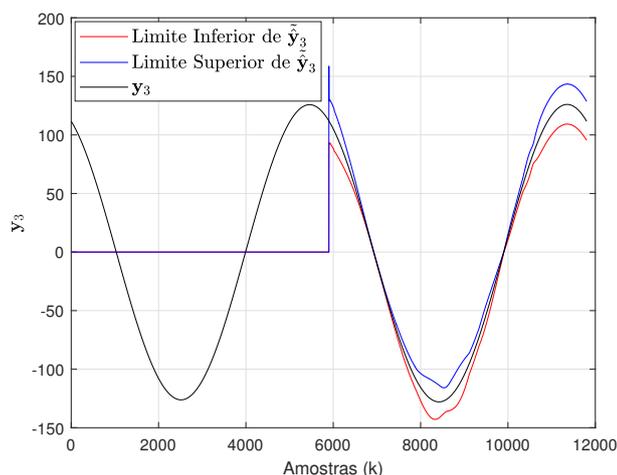


Figura 20 – Comparação entre o intervalo estimado e o valor real da distância no eixo Z entre os satélites da formação PRISMA no referencial $LVLH$.

Tabela 3 – Comparação com outras metodologias presentes na literatura.

Método	e_{3D} (desvio padrão)	ACC (desvio padrão)
Metodologia proposta	0,0422m (0,0554m)	99,4675% (0,5323%)
Filtro de Kalman Estendido	2,991m	NA
Q-MLE-AEKF (FRASER; ULRICH, 2021)	0.2389m	NA
Q-FAEKF(FRASER; ULRICH, 2021)	0,2646m	NA
Versão proposta em Montel e Serra (2023)	0,1104m	94,06%

Ulrich (2021) consistem em versões adaptativas do filtro de Kalman estendido. Neles, foram utilizados a estimação em máxima verossimilhança (MLE) e um sistema de regras nebulosas (representado na sigla pelo F) para adaptar a matriz de covariância do ruído de processo Q . Já a versão proposta em Montel e Serra (2023) consiste em uma versão nebulosa

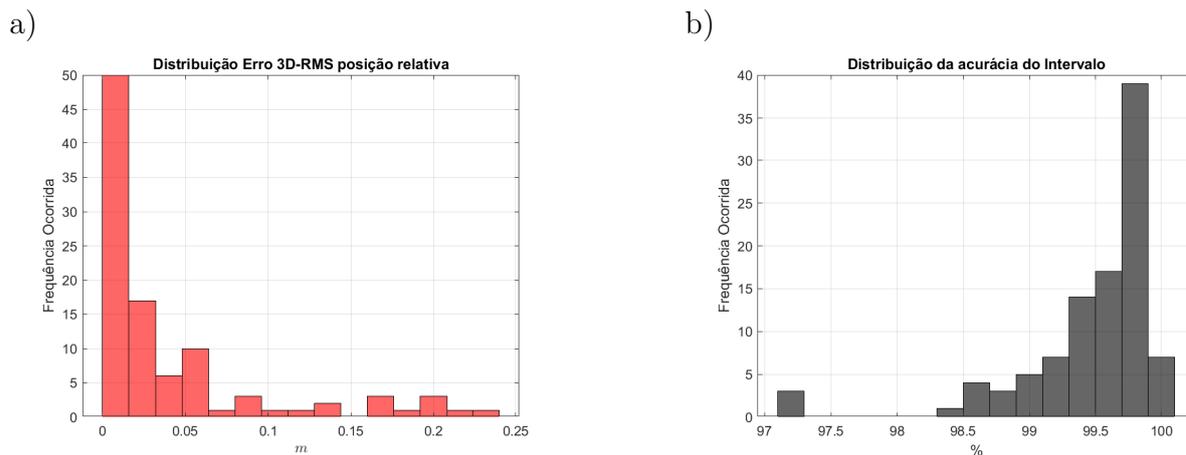


Figura 21 – Distribuição do erro e acurácia obtida nas 100 realizações ruidosas.

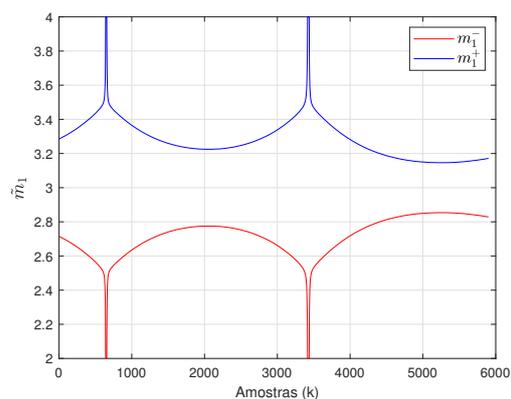
do algoritmo OKID, onde não é feita a adaptação do parâmetro de incerteza \tilde{m} e nem a seleção da quantidade ideal de regressores. Dentre todas as metodologias apresentadas, o método proposto nesta dissertação foi o que apresentou o melhor desempenho. Como este resultado é um valor médio, deve-se observar como este valor está distribuído. Nesse sentido, a figura 21 apresenta um histograma da distribuição dos erros 3D-RMS e da acurácia obtida.

Nos dois gráficos apresentados na figura 21, percebe-se que há uma alta concentração de resultados em um valor pouco deslocado da média informada, o que pode ser considerado como devido à influência de uma minoria de resultados levemente inferiores à média. Porém, é importante salientar que em ambos os casos, a região de maior concentração é localizada em um valor superior em qualidade à média observada.

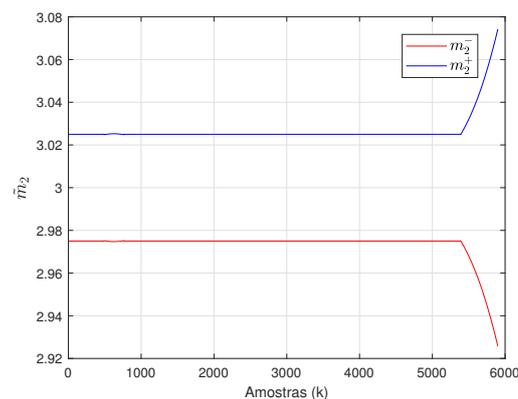
A figura 22 apresenta a variação do expoente de ponderação nebulosa durante a operação *online*. Para o eixo X , o intervalo máximo do expoente é sugerido em dois pontos da base de dados, o que ocorre nos dois pontos em que esta série temporal se aproxima do valor nulo; ademais, se não levarmos em conta estes dois pontos, é possível observar (pelas linhas do gráfico) que o nível de incerteza da série temporal diminui com tempo. Para o eixo Y , o valor da série temporal é alto em magnitude durante toda a operação, necessitando assim do intervalo mínimo de incerteza na maior parte do tempo. O cenário do eixo Z , é similar ao observado no eixo X , porém, devido à covariância do ruído aproximadamente constante, não é possível observar uma variação expressiva no valor de \tilde{m} , exceto nos pontos onde a série temporal se aproxima de zero.

Na figura 23 é apresentada a variação temporal na ordem de cada submodelo. Para analisá-la, primeiro deve-se informar que a quantidade máxima de regressores para todas as regras foi selecionada como $p_m = 2$. Desse modo, observa-se que a ordem dos submodelos varia entre 6 e 9. Além disso, a ordem de todos os submodelos não varia rapidamente, o que pode indicar que a configuração escolhida para os parâmetros de Markov é bem

a) Eixo X



b) Eixo Y



c) Eixo Z

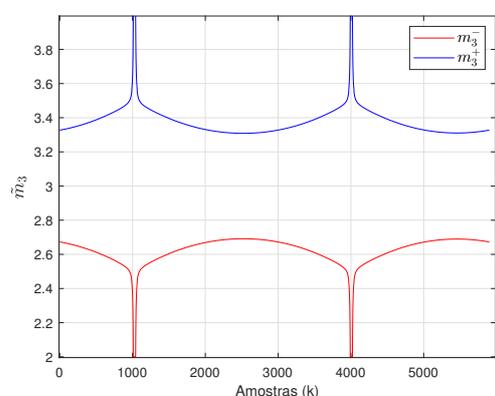


Figura 22 – Limites superior e inferior para o expoente de ponderação nebulosa \tilde{m} adaptado segundo o nível de incerteza das medições de cada variável.

superior às demais.

Por fim, como se trata de uma aplicação *online*, deve-se avaliar a tempo utilizado pela metodologia para atualizar seus parâmetros. Neste conjunto de teste, foi utilizado um notebook com 8g Gb de memória RAM, com o processador 11th Gen Intel(R) Core(TM) *i7* – 1165G7@2.80GHz. Em média, foi necessário 0,02 *s/it.* sendo bem inferior ao período de amostragem das medições da formação PRISMA, que é de 1 *s.*

6.1.2.2 Simulação Mista

Para o teste envolvendo simulação mista, o modelo com o desempenho médio foi utilizado. Para o caso da posição relativa, a simulação livre se torna imprecisa em um horizonte de eventos muito distante (o motivo será explicado nesta seção). Por isso, durante a etapa *online*, foi simulado de forma arbitrária dez pontos de falta de medição com duração de 30 segundos cada. Tais falhas foram alocadas de forma a estressar o sistema e observar seu comportamento em diferentes regiões de operação. A primeira falha foi alocada na região onde a previsão um passo a frente obteve o pior desempenho. As falhas 3 a 6 foram alocadas em uma região intermediária da operação *online*. As demais

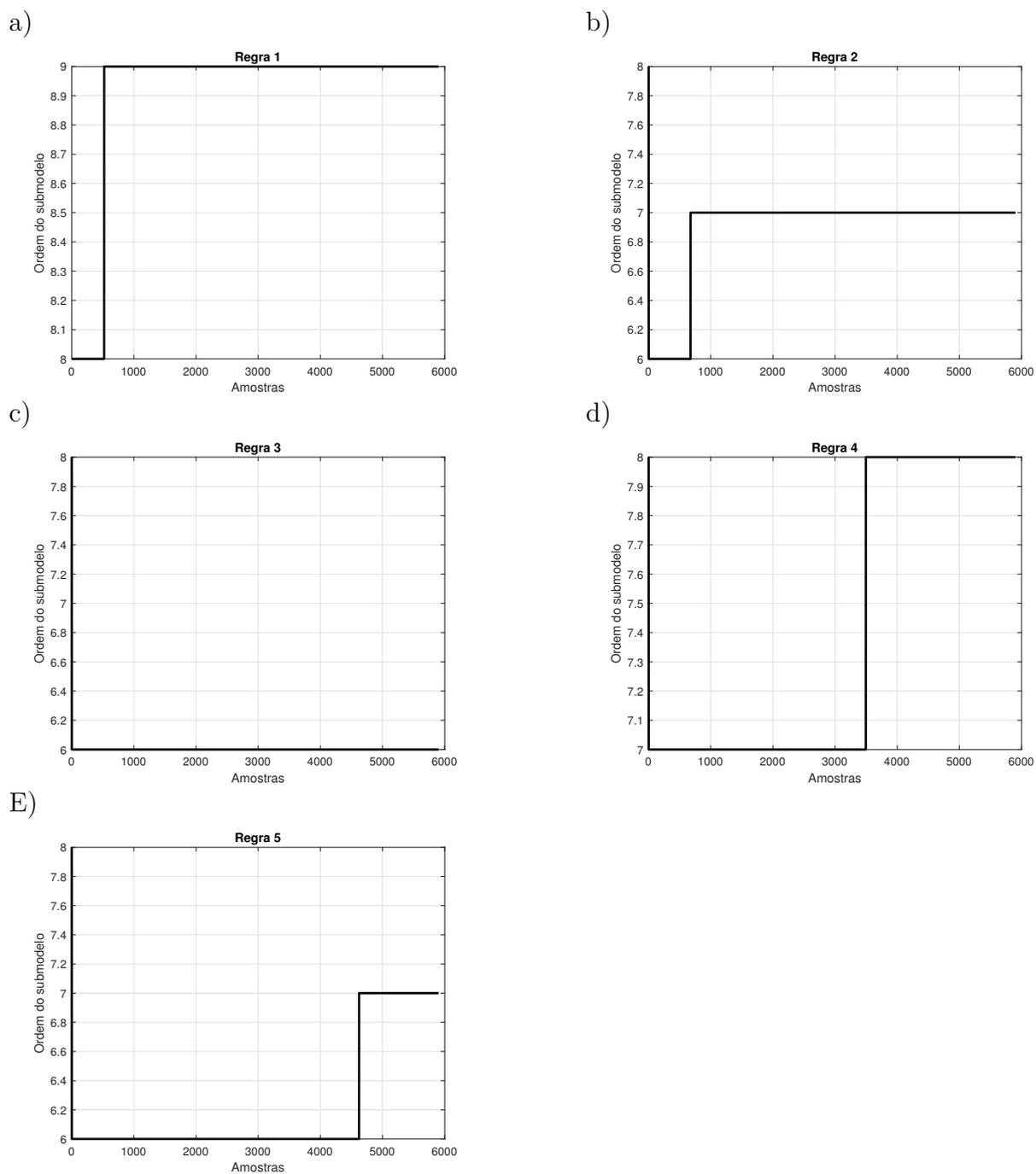


Figura 23 – Variação temporal na ordem de cada um dos cinco submodelos.

falhas foram alocadas próximo ao final da operação *online*. Por se tratar de vários pontos de curta duração, é difícil observar graficamente estes resultados; assim, as figuras 24-26 apresentam para cada variável a comparação entre o intervalo estimado e o valor real da variável. Além disso, a tabela 4 apresenta os intervalos de tempo onde a simulação livre ocorreu, bem como uma comparação entre a acurácia obtida pela simulação mista e um passo a frente neste mesmo intervalo.

Tabela 4 – Desempenho da simulação livre.

	Início	Fim	Acurácia	Referência
Falha 1	6521s	6551s	73,3%	89,7%
Falha 2	6620s	6650s	100%	100%
Falha 3	7892s	7922s	90%	100%
Falha 4	8573s	8603s	91,1%	100%
Falha 5	8798s	8828s	94,4%	100%
Falha 6	9397s	9427s	98,9%	100%
Falha 7	10797s	10827s	100%	100%
Falha 8	11226s	11256s	100%	100%
Falha 9	11318s	11348s	100%	100%
Falha 10	11616s	11646s	100%	100%

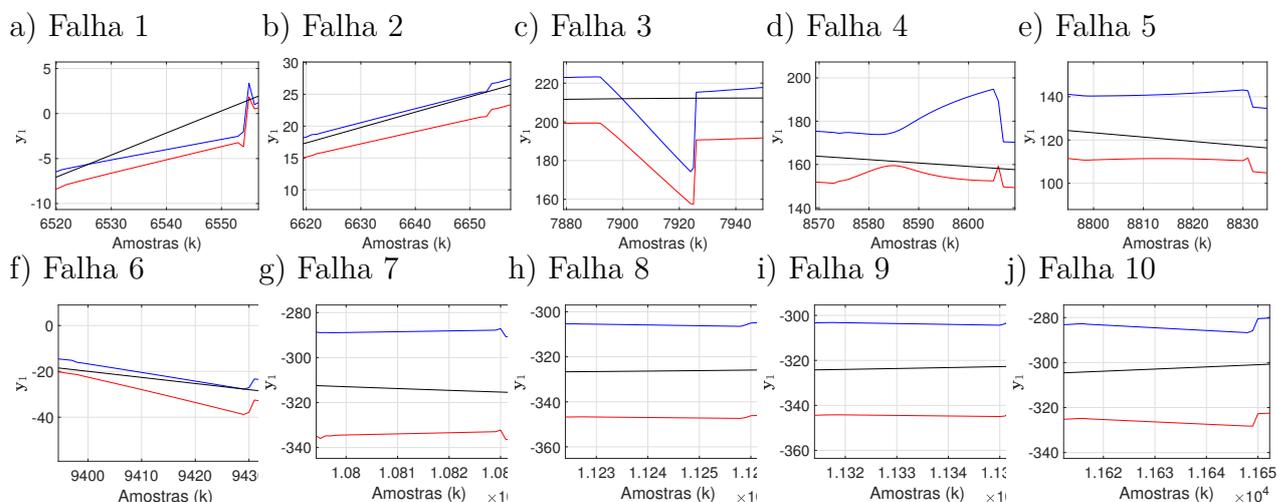


Figura 24 – Previsão intervalar para o eixo X realizada durante as situações de falta de medição na simulação mista.

A tabela 4 mostra que a primeira falha teve o pior resultado de acurácia. Isso era esperado, pois a falha ocorreu em um dos piores recortes da base de dados. Para os pontos mais afastados do conjunto de dados inicial, a simulação livre obteve os melhores resultados. No entanto, para as falhas 3 a 6, a acurácia da simulação livre é inferior ao resultado esperado. Este é o principal problema para simulações livres com um horizonte de eventos mais distante. Para analisar os motivos desta diferença, a figura 27 mostra a decomposição em componentes principais da base de dados completa, junto com a partição de dados obtida pelo ANT-2IEMV. Essa figura pode ajudar a explicar a diferença observada.

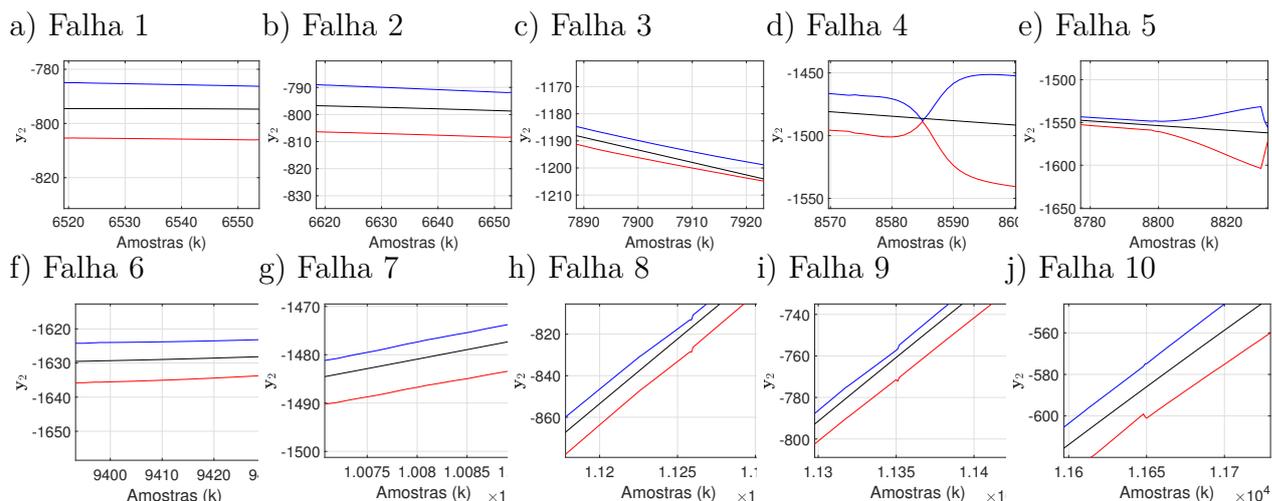


Figura 25 – Previsão intervalar para o eixo Y realizada durante as situações de falta de medição na simulação mista.

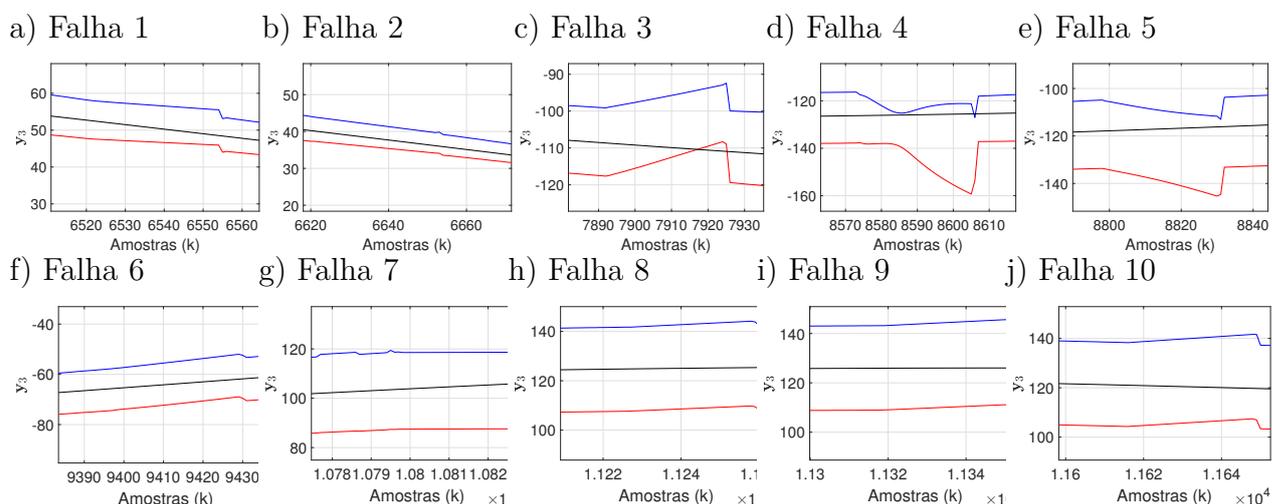


Figura 26 – Previsão intervalar para o eixo Z realizada durante as situações de falta de medição na simulação mista.

A figura 27 mostra que os novos dados obtidos após a etapa de processamento *offline* estão em uma região diferente da região dos dados observados previamente. Essa região não é bem representada pelos modelos iniciais, então é necessário atualizá-los. No entanto, se não houver medições, não é possível atualizar os modelos. Isso explica a queda na qualidade da simulação livre nas falhas 3 a 6. Assim, à medida que os dados se afastam temporalmente, eles se aproximam de uma região que já é bem representada pelos submodelos. Portanto, a resposta da simulação livre tende a melhorar, o que é demonstrado nas falhas 7 a 10.

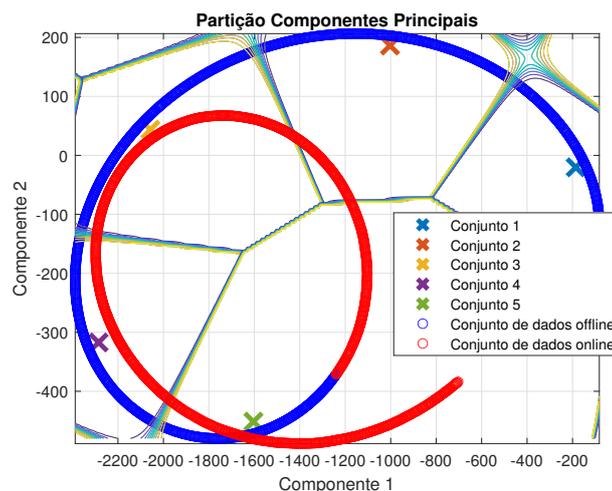


Figura 27 – Decomposição da base de dados em função de suas duas componentes principais, classificadas de acordo com os cinco conjuntos nebulosos adotados

6.1.3 Velocidade Relativa

Para testar a metodologia com o conjunto de dados da velocidade relativa, foram realizadas 100 simulações ruidosas. Os resultados serão apresentados em termos de média e desvio padrão. As imagens apresentadas serão referentes à simulação que obteve os resultados mais próximos da média. Os parâmetros apresentados na tabela 5 foram utilizados para inicializar a metodologia.

Tabela 5 – Parâmetros iniciais para a estimação paramétrica da metodologia.

Parâmetros	Descrição	Valor
c	Quantidade de Conjuntos	5
ϵ	Critério de Convergência	10^{-4}
N	Dimensão do conjunto de dados inicial	5900
δ	Número de Linhas da Matriz Trajetória	700

A tabela 6 apresenta a quantidade de componentes espectrais extraídas em cada variável medida.

Tabela 6 – Quantidade de componentes espectrais selecionadas.

Variável	Componentes extraídas
Velocidade no Eixo X y_1	2
Velocidade no Eixo Y y_2	2
Velocidade no Eixo Z y_3	2

A variável linguística do antecedente é formada por seis variáveis, sendo três de entrada e três de saída. A figura 28 mostra como essas seis variáveis são representadas em duas componentes principais, que representam mais de 99% da variação total do conjunto

de dados. A figura 29 mostra como essas duas componentes principais são usadas para gerar as funções de pertinência intervalar.

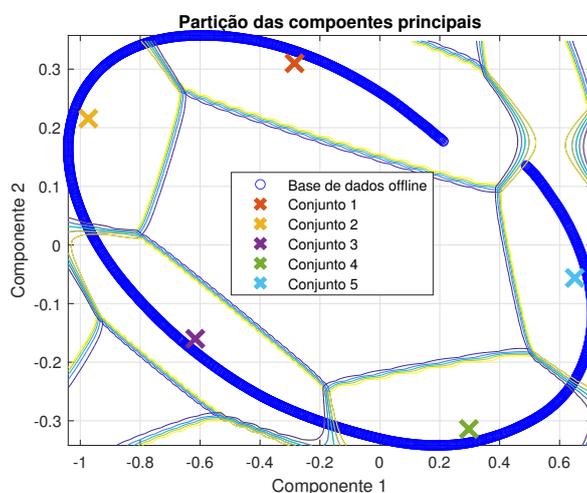


Figura 28 – Partição nebulosa para o conjunto de dados da velocidade relativa com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.

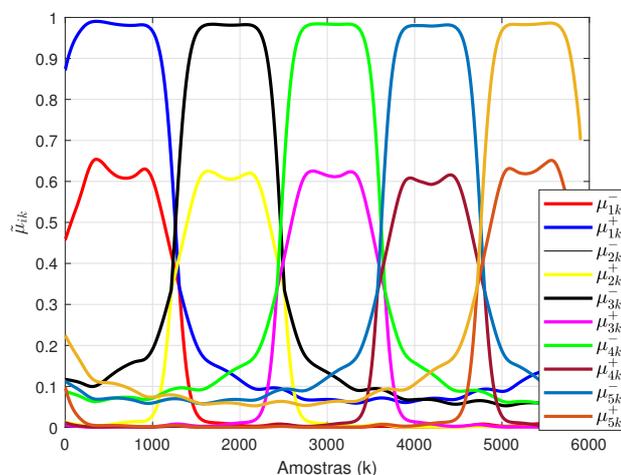


Figura 29 – Funções de pertinência intervalar para os cinco conjuntos nebulosos estimados.

6.1.3.1 Simulação um Passo a Frente

A metodologia proposta divide os dados em cinco conjuntos nebulosos e estima um modelo inicial para cada conjunto. Em seguida, ela usa esses modelos para estimar os valores da variável para o próximo período amostral. Para isso, ela atualiza os parâmetros dos modelos com base nos dados do período anterior. As figuras 30-32 mostram a comparação entre o valor real das variáveis de saída e o intervalo estimado. Nelas, é possível observar com clareza que na maioria dos casos, o intervalo estimado inclui o valor real da variável.

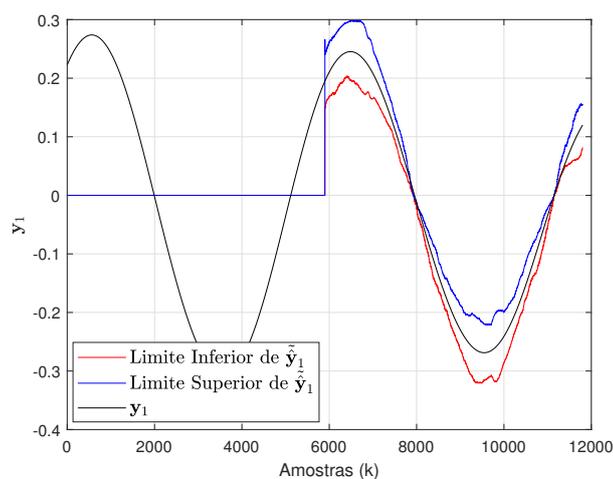


Figura 30 – Comparação entre o intervalo estimado e o valor real da velocidade no eixo X entre os satélites da formação PRISMA no referencial $LVLH$.

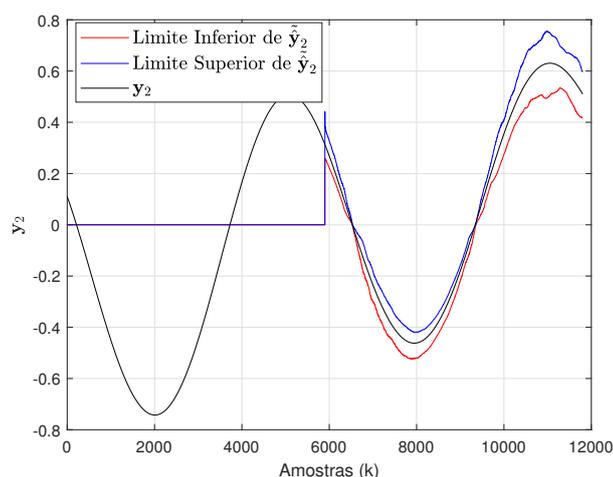


Figura 31 – Comparação entre o intervalo estimado e o valor real da velocidade no eixo Y entre os satélites da formação PRISMA no referencial $LVLH$.

A tabela 7 compara a eficiência do resultado obtido com outras metodologias presentes na literatura, mostrando o erro 3D-RMS e a acurácia obtida por cada metodologia. Entre os resultados apresentados, a metodologia desenvolvida nesta dissertação possui os melhores resultados.

Tabela 7 – Comparação com outras metodologias presentes na literatura.

Método	e_{3D} (desvio padrão)	ACC (desvio padrão)
Metodologia proposta	$9.10^{-4}cm/s$ ($3.10^4cm/s$)	97,342% (0,644%)
Filtro de Kalman Estendido	3,76cm/s	NA
Q-MLE-AEKF (FRASER; ULRICH, 2021)	0,1cm/s	NA
Q-FAEKF(FRASER; ULRICH, 2021)	0,1cm/s	NA
Versão proposta em Montel e Serra (2023)	0,02cm/s	89,36%

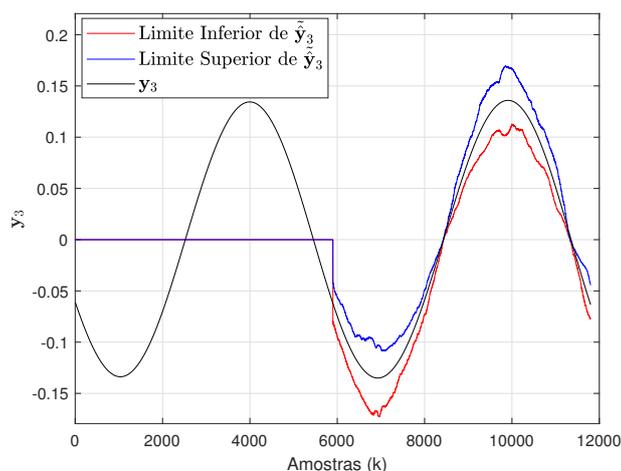


Figura 32 – Comparação entre o intervalo estimado e o valor real da velocidade no eixo Z entre os satélites da formação PRISMA no referencial $LVLH$.

O valor apresentado na tabela 7 é a média entre os valores obtidos nas 100 simulações. Assim, a figura 33 apresenta a distribuição observada entre o e_{3D} e a Acurácia do intervalo. Nela, pode-se observar que os resultados se distribuem ao redor da média, com esta pertencendo ao intervalo de maior recorrência.

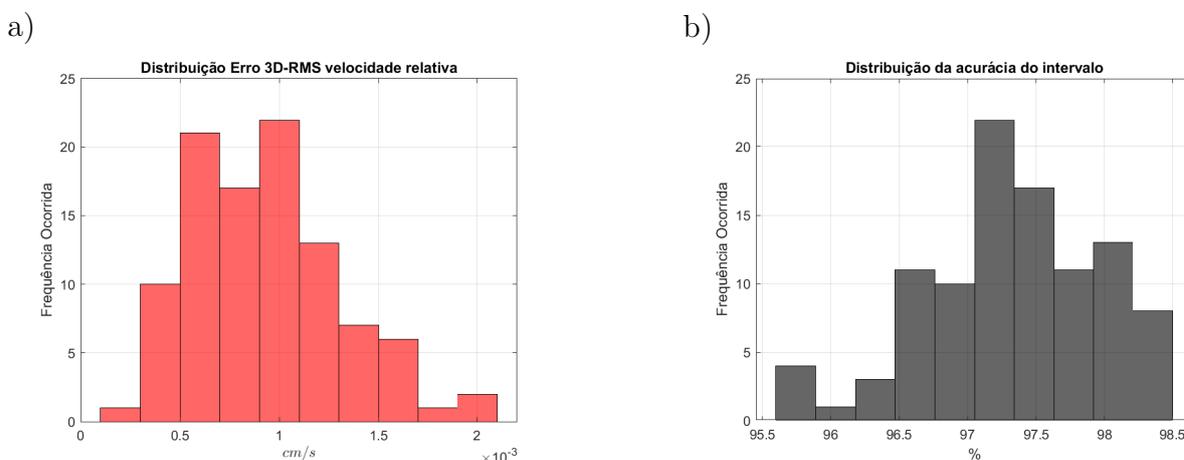
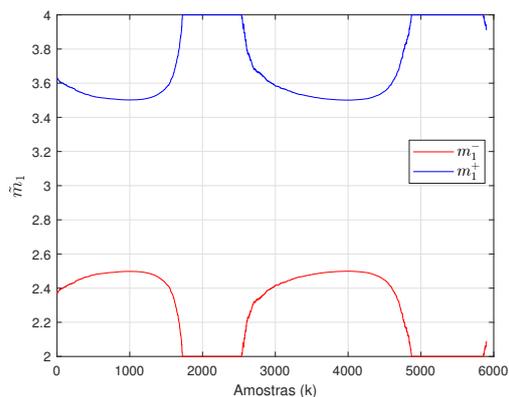


Figura 33 – Distribuição do erro e acurácia obtida nas 100 realizações ruidosas.

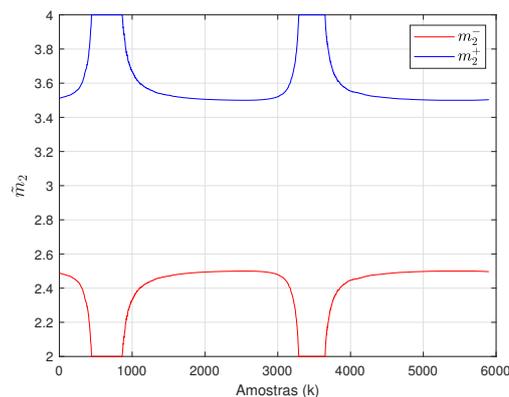
A figura 34 mostra que o intervalo de \tilde{m} é maior para a velocidade relativa do que para a posição. Isso ocorre, pois as séries temporais desta base de dados possuem uma relação sinal ruído (SNR) de $12,02dB$ para a variável \mathbf{y}_1 , $19,11dB$ para \mathbf{y}_2 , e $10,6148dB$ para \mathbf{y}_3 . Enquanto que para a posição relativa, estes valores são de $41,19dB$ para \mathbf{y}_1 , $55,85dB$ para \mathbf{y}_2 , e de $38,15dB$ para \mathbf{y}_3 . Portanto, a potência do ruído de medição é bem mais significativo para a base de dados da velocidade relativa.

Na figura 35 é apresentada a variação temporal na ordem de cada submodelo. A quantidade máxima de regressores para cada regra foi selecionada como $p_m = 2$. No geral,

a) Eixo X



b) Eixo Y



c) Eixo Z

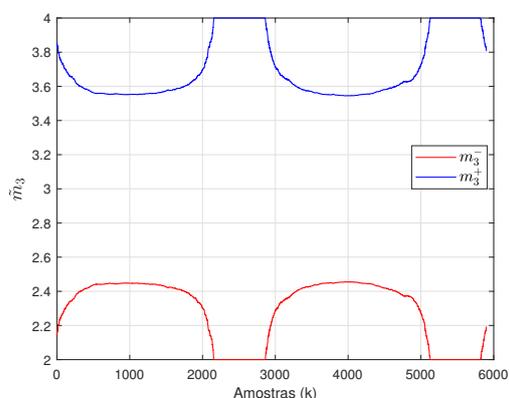


Figura 34 – Limites superior e inferior para o expoente de ponderação nebulosa \tilde{m} adaptado segundo o nível de incerteza das medições de cada variável.

a ordem dos submodelos varia entre 6 e 9. Diferente do teste anterior, é observado uma variação mais rápida da ordem em alguns pontos da base de dados.

Por fim, como se trata de uma aplicação *online*, deve-se analisar o tempo utilizado pela metodologia para atualização em cada iteração. Para este conjunto de teste, foi necessário em média $0,02 \text{ s/it}$. Dessa forma, ao somar esse resultado com o obtido no teste da posição relativa, observa-se que a metodologia proposta necessita de aproximadamente $0,04 \text{ s/it}$, sendo inferior ao período de amostragem de 1 s utilizado na simulação.

6.1.3.2 Simulação Mista

Para este conjunto de testes, o modelo estimado possui uma boa resposta a longo prazo em simulações livre. Assim, foram simulados cinco pontos de falhas de medição com duração de 200 segundos cada. Os pontos de falhas foram selecionados arbitrariamente, de forma a estressar o sistema em cada uma de suas três variáveis de saída. As regiões de falha estão listadas na tabela 8.

Neste cenário, como a janela de tempo da simulação livre é grande, é possível representá-la graficamente. As figuras 36-38 mostram os resultados da simulação mista

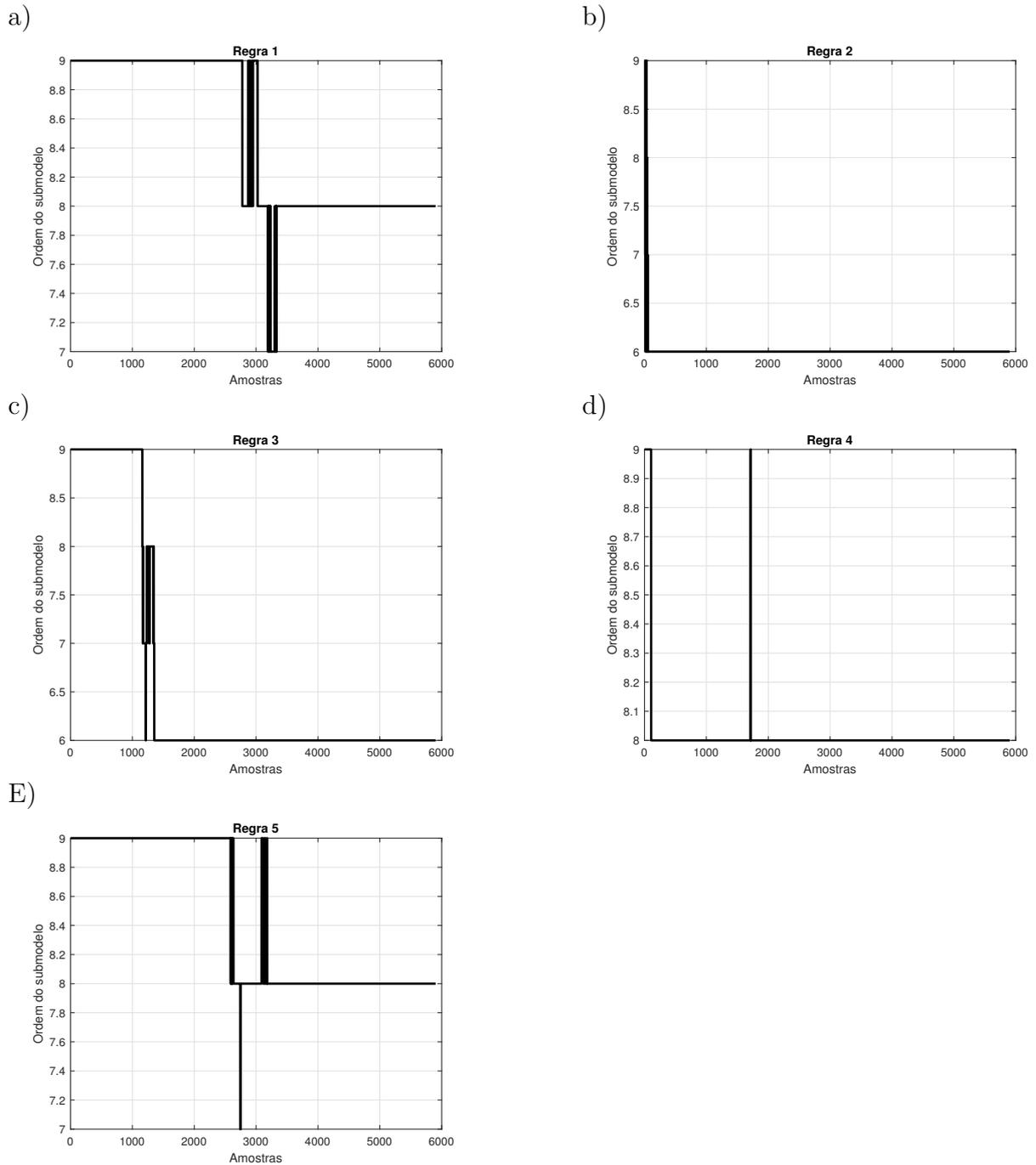


Figura 35 – Variação temporal na ordem de cada um dos cinco submodelos.

Tabela 8 – Regiões onde a simulação livre foi realizada.

	Início	Fim
Falha 1	8300s	8500s
Falha 2	9598s	9798s
Falha 3	10427s	10627s
Falha 4	10784s	10984s
Falha 5	11283s	11483s

do modelo. Em todas as figuras, é difícil notar a diferença entre as regiões com e sem simulação livre. Essas regiões foram sombreadas para destacar essa informação. Além disso, a metodologia proposta não apresenta problemas de instabilidade na retomada da atualização *online*.

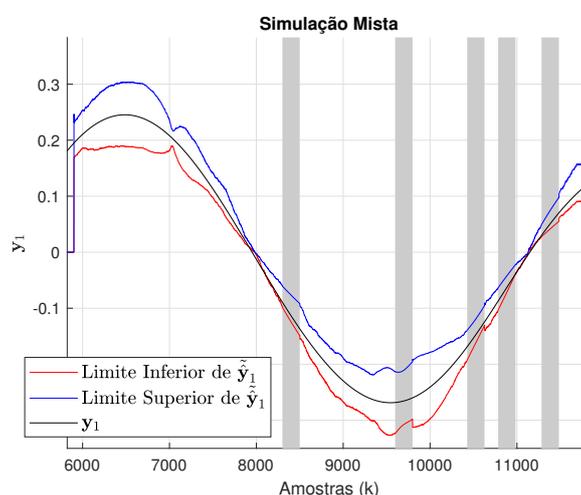


Figura 36 – Simulação com ausência de medições para a velocidade relativa no eixo X . As áreas sombreadas representam as regiões onde a simulação livre ocorreu.

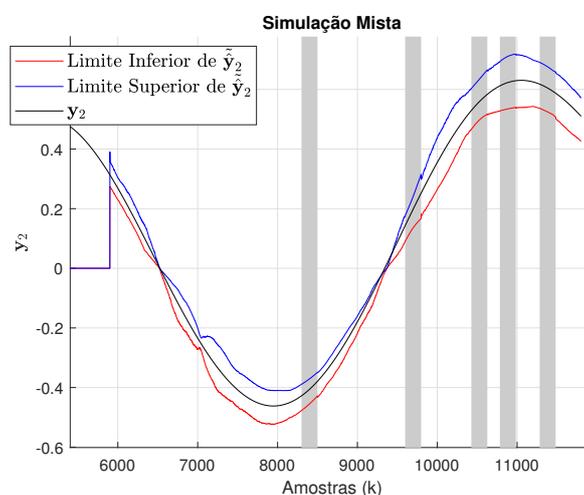


Figura 37 – Simulação com ausência de medições para a velocidade relativa no eixo Y . As áreas sombreadas representam as regiões onde a simulação livre ocorreu.

O desempenho da simulação mista pode ser explicado pela figura 39, a qual apresenta a decomposição da base de dados em suas duas componentes principais. Nela, pode-se observar que o conjunto de dados *offline* é bem abrangente, de modo que durante a etapa *online*, as novas amostras estão sempre próximas a uma região já abrangida pelos submodelos, logo, a resposta em simulação livre tende a ser satisfatória.

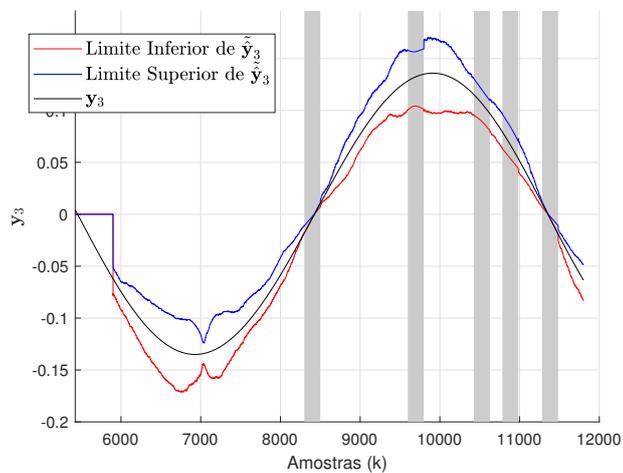


Figura 38 – Simulação com ausência de medições para a velocidade relativa no eixo Z . As áreas sombreadas representam as regiões onde a simulação livre ocorreu.

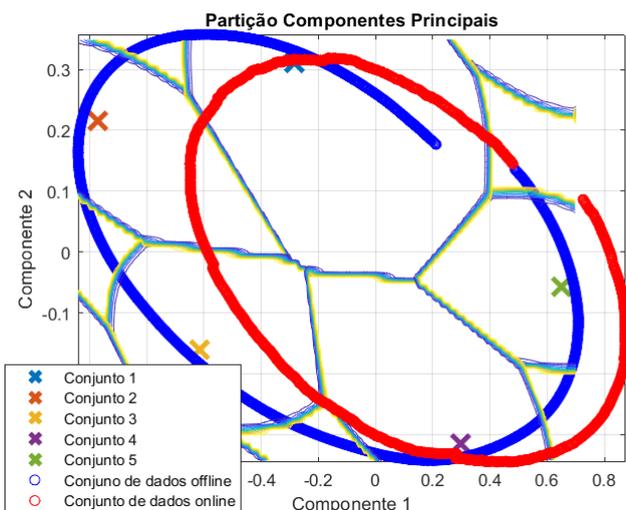


Figura 39 – Decomposição da base de dados em função de suas duas componentes principais, classificadas de acordo com os cinco conjuntos nebulosos adotados

7 ANÁLISE EXPERIMENTAL

Neste capítulo, a metodologia proposta será avaliada utilizando conjunto de dados experimentais. Para isso, dois problemas serão propostos, o primeiro consiste em um helicóptero com dois graus de liberdade, e o segundo consiste no veículo suborbital VS-30.

7.1 Rastreamento intervalar da elevação e azimute de um helicóptero 2DoH

O segundo problema com dados experimentais abordado é o desenvolvimento de um modelo para o helicóptero com dois graus de liberdade, apresentado na figura 40. O helicóptero possui duas variáveis de saída: o ângulo de azimute ($\mathbf{y}_1 \in [-360^\circ, +360^\circ]$) e o ângulo de elevação ($\mathbf{y}_2 \in [-40^\circ, +40^\circ]$). Essas variáveis são controladas por dois sinais de tensão elétrica: $\mathbf{u}_1 \in [-24V, +24V]$ e $\mathbf{u}_2 \in [-24V, +24V]$.

A base de dados obtida para este teste possui 1491 medições de cada variável de entrada e saída. Nesse sentido, as figuras apresentam o conjunto de dados obtidos. Destes, foram utilizados 750 para o treinamento *offline* do modelo e as 741 restantes foram utilizadas na etapa *online* da metodologia. Nesse sentido, as figuras 41 e 42 apresentam o conjunto total de medições para as variáveis de entrada e saída.

A tabela 9 apresenta os parâmetros iniciais informados à metodologia proposta. Neles, percebe-se que o número de linhas da matriz trajetória foi selecionado com um valor bem inferior aos demais testes. Isso ocorreu pois a influência de ruído na base de dados é bem inferior quando comparado aos demais testes.

Tabela 9 – Parâmetros iniciais para a estimação paramétrica da metodologia.

Parâmetros	Descrição	Valor
c	Quantidade de Conjuntos	3
ϵ	Critério de Convergência	10^{-4}
N	Dimensão do conjunto de dados inicial	750
δ	Número de Linhas da Matriz Trajetória	40

Na etapa de pré-processamento, a análise espectral singular foi aplicada apenas as variáveis de saída, de forma a extrair suas componentes espectrais mais significativas. A tabela 10 apresenta a quantidade de componentes extraídas. Durante a etapa de estimação paramétrica do antecedente, a análise de componentes principais reduziu a dimensionalidade da base de dados para 2 componentes, retendo mais de 99% da variação do conjunto original. As figuras 43 e 44 mostram respectivamente a distribuição do conjunto de dados

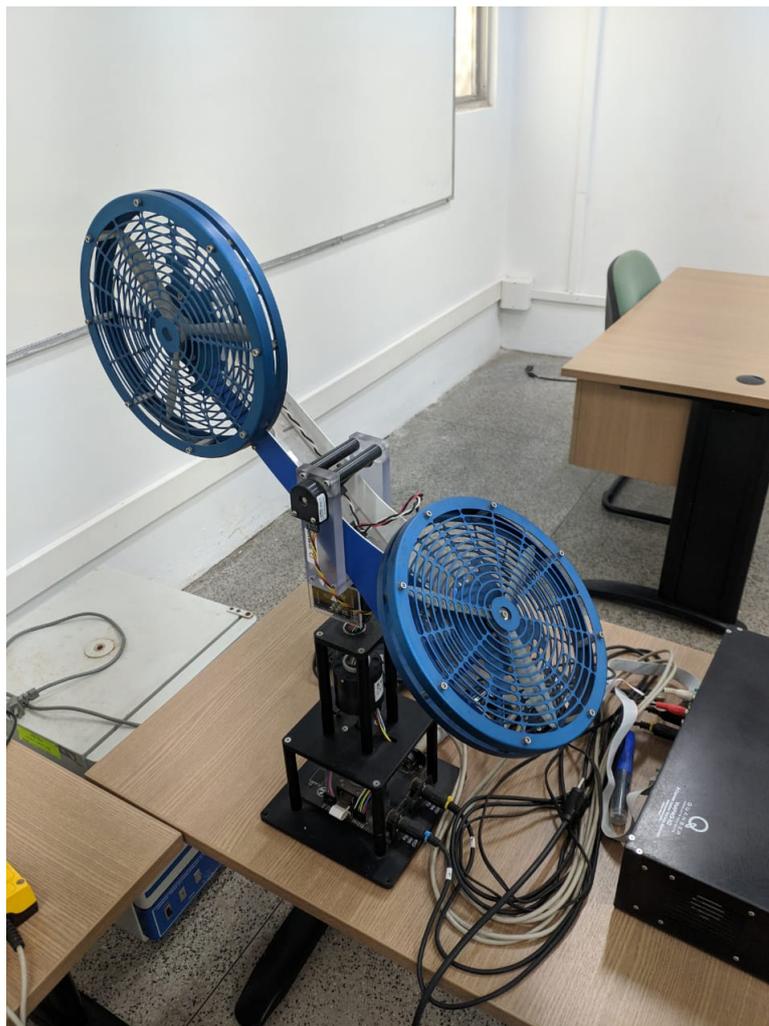


Figura 40 – Helicóptero com dois graus de liberdade.

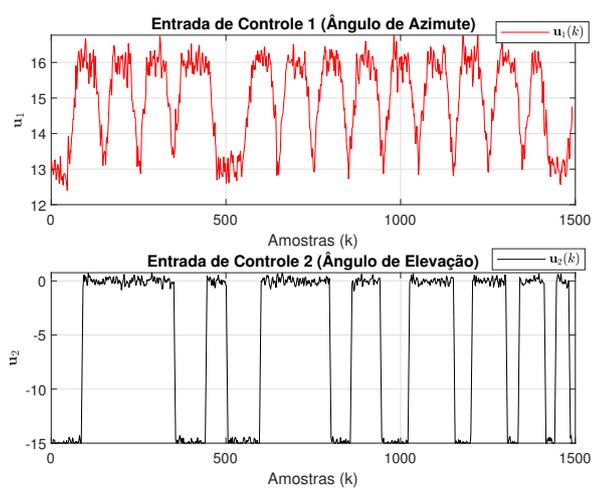


Figura 41 – Entradas de controle do Helicóptero.

offline no novo espaço de duas dimensões, bem como os graus de pertinência intervalar de cada ponto de dados às duas variáveis.

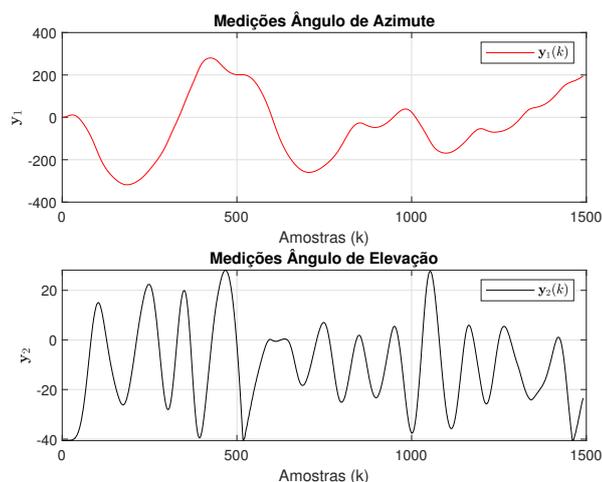


Figura 42 – Variáveis de Saída do Helicóptero.

Tabela 10 – Quantidade de componentes espectrais selecionadas.

Variável	Componentes extraídas
Ângulo de Azimute y_1	5
Ângulo de Elevação y_2	8

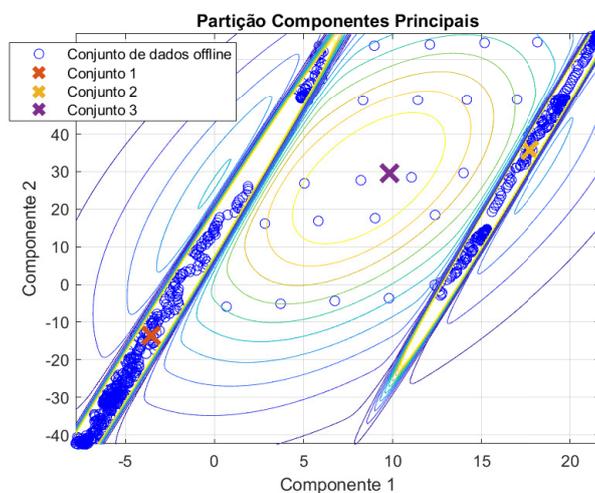


Figura 43 – Partição nebulosa para o conjunto de dados com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.

Observe que na figura 43, o conjunto de dados com dimensão reduzida é apresentado junto com a informação de qual conjunto este dado esta melhor relacionado. Além disso, percebe-se que o conjunto 3 está em uma zona de transição entre dois pontos de operação, o que pode ser constatado ao observar sua função de pertinência na figura 44. Este conjunto é necessário, pois estes dados possuem uma orientação bem diferente da apresentada pelos conjunto 1 e 2, além de possuírem uma variação mais rápida.

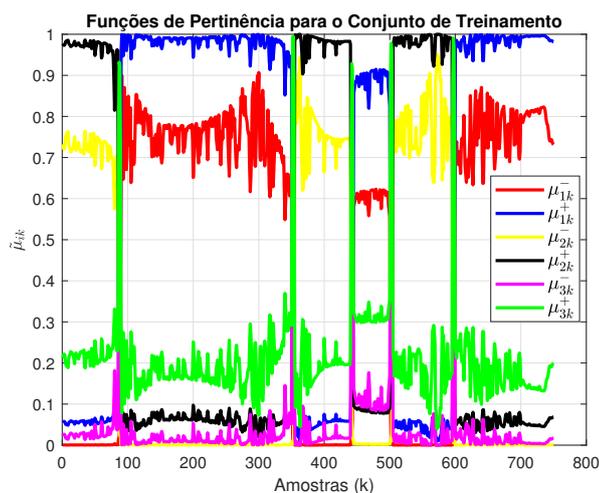
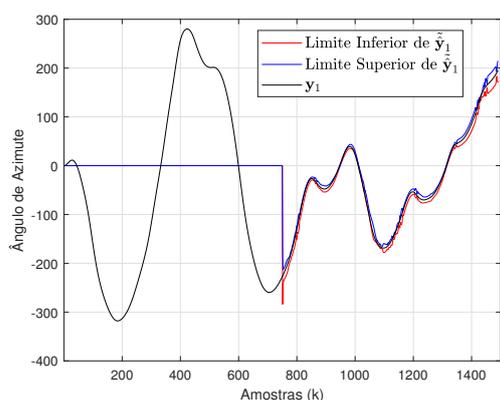


Figura 44 – Funções de pertinência para a etapa *offline* da metodologia.

a) Ângulo de azimute



b) Ângulo de elevação

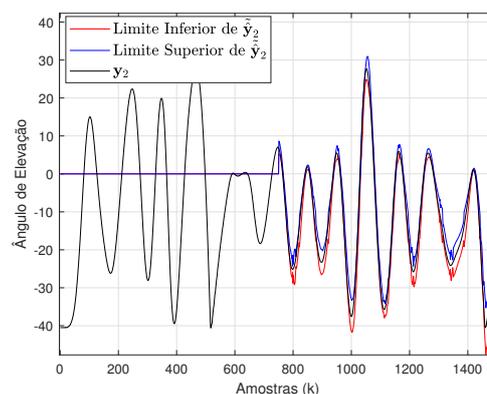


Figura 45 – Comparação entre o intervalo estimado e o valor real para o ângulo de elevação e azimute.

Por possuir uma tensão de controle como variável de entrada, esta base de dados será submetida apenas a simulação de um passo a frente. Nesse sentido, a figura 45 apresenta a comparação entre o valor da variável de saída e o intervalo estimado pelo modelo.

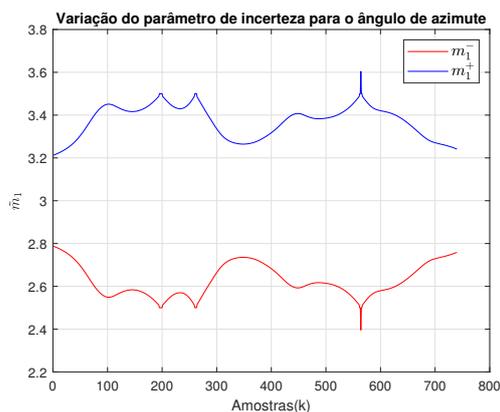
Em ambas as figuras, pode-se observar que o intervalo estimado é bem próximo ao valor da variável, sendo difícil de observar em algumas regiões. Nesse sentido, a tabela 11 apresenta os valores de acurácia e RMSE obtidos pela metodologia proposta, e os compara com outras versões nebulosas do algoritmo ERA/OKID.

Nota-se que tais adaptações do algoritmo ERA/OKID foram desenvolvidas no decorrer desta dissertação. A sigla SSA indica que o método utilizou de componentes espectrais não observáveis para estimar seus parâmetros. A sigla T2 indica que foi adotado o algoritmo de agrupamento ANT-2IEMV para a estimação paramétrica do antecedente.

Tabela 11 – Acurácia e erro da estimativa obtida.

Método	ACC	RMSE y_1	RMSE y_2
Metodologia proposta	91,36%	0,296°	0,286°
OKID/SSA/T2	74,1%	1,3435°	15,7163°
OKID/SSA/T2/m	83,77%	1,07°	1,269°
OKID/SSA/T2/m/PCA	91,36%	0,296°	0,286°

a) Ângulo de Azimute



b) Ângulo de Elevação

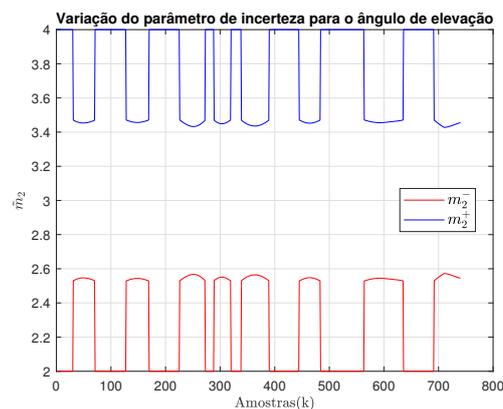


Figura 46 – Limites superior e inferior para o expoente de ponderação nebulosa \tilde{m} adaptado segundo o nível de incerteza das medições de cada variável.

A letra m indica que foi utilizada a modelagem do parâmetro de incerteza para adaptar o intervalo \tilde{m} durante a operação online. A sigla PCA indica que a base de dados teve sua dimensão reduzida pela análise de componentes principais.

Observa-se que a variação na ordem de cada submodelo não surtiu efeito neste conjunto de testes. Isso ocorreu pois neste caso, as ordens do modelo se mantiveram constantes. Além disso, o mecanismo de adaptação do intervalo \tilde{m} gerou uma melhora significativa os resultados para o ângulo de elevação (y_2). Tal cenário ocorreu pois os valores de \tilde{m} para cada variável de saída são bem distintos, portanto, um valor fixo e comum a ambas as variáveis pode gerar resultados discrepantes.

A figura 46 apresenta a variação do expoente de ponderação nebulosa durante a etapa *online*. Na figura 46(a), é possível observar que o intervalo de \tilde{m} varia pouco com o tempo, enquanto que na figura 46(b) o intervalo assume varias vezes o seu valor máximo. Em ambas as variáveis, a relação sinal ruído é superior a 50dB, porém para o ângulo de elevação, o valor medido passa varias vezes pelo ponto zero, o que justifica o formato para a evolução temporal de \tilde{m}_j apresentado na figura 46(b).

Para este cenário de teste, o mecanismo de seleção e atualização determinou a ordem 6 para o modelo durante toda a operação (*online* e *offline*). Ou seja, a melhor representação para todas as regras foi obtida com dois regressores. Tal comportamento era esperado, pois se trata de uma aplicação em um sistema físico onde as variáveis de saída

são controladas por dois sinais elétricos. Logo, este sistema possui uma ordem fixa.

7.2 Rastreamento e previsão intervalares da trajetória de um veículo suborbital

Apresentado na figura 47, o VS-30 é um veículo de porte médio criado pelo Comando da Força Aérea Brasileira (COMAER), sendo utilizado para testes e projetos de lançamentos em condições adversas. Este foguete possui 0,3 m de diâmetro e e 5,4 m de comprimento, com uma massa de 490 kg.

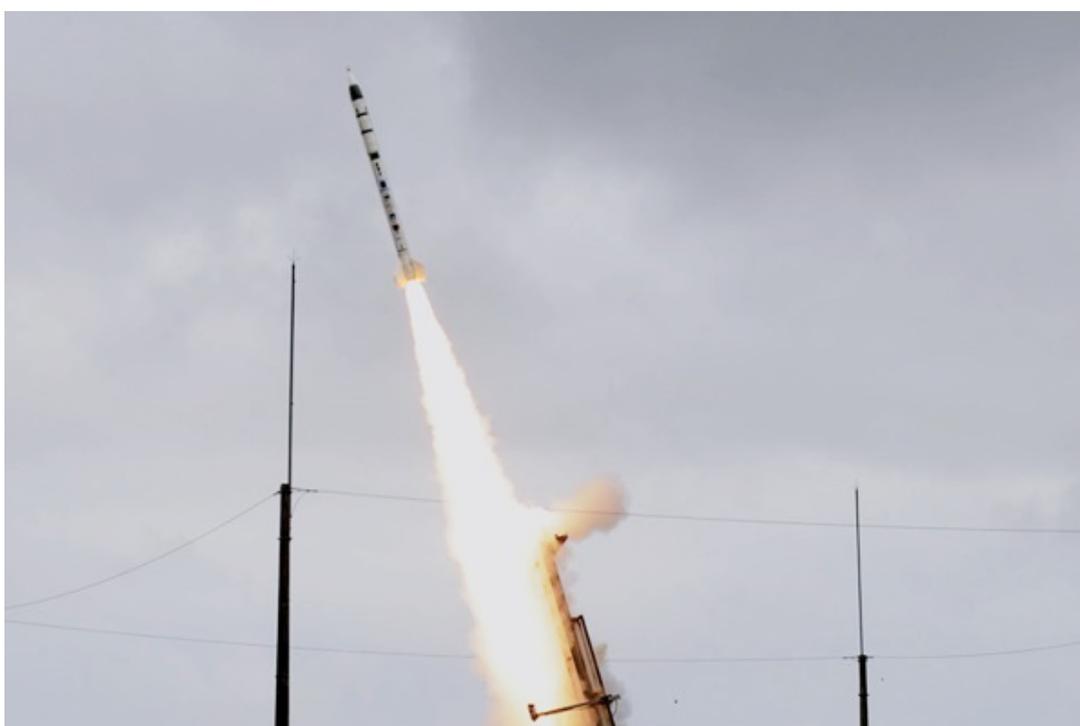


Figura 47 – Veículo suborbital VS-30 (COMAER). Fonte: Força aérea brasileira (FAB MIL).

O conjunto de dados foi disponibilizado por membros do laboratório de inteligência computacional aplicado a tecnologia que atuam no Centro de Lançamento de Alcântara (CLA). Este conjunto de medições contém dados sobre o ângulo de elevação y_1 , ângulo de azimute y_2 e distância percorrida y_3 . Para a aplicação na metodologia proposta, as variáveis de entrada são definidas como o conjunto de medições com um atraso de tempo de 1 segundo, ou seja:

$$\mathbf{u}_k = \mathbf{y}_{k-1}. \quad (7.1)$$

Foram utilizadas 2900 medições deste conjunto de dados experimentais, os quais 1000 foram utilizados para o treinamento e os restantes 1900 para a etapa *online* da metodologia.

Na tabela 12, apresenta-se os parâmetros utilizados para inicializar o método proposto nesta dissertação.

Tabela 12 – Parâmetros iniciais para a estimação paramétrica da metodologia.

Parâmetros	Descrição	Valor
c	Quantidade de Conjuntos	2
ϵ	Critério de Convergência	10^{-4}
N	Dimensão do conjunto de dados inicial	1000
δ	Número de Linhas da Matriz Trajetória	400

Após a etapa de pré-processamento, foram extraídas as componentes espectrais mais significativas em cada variável. Esta informação é apresentada na tabela 13.

Tabela 13 – Quantidade de componentes espectrais selecionadas.

Variável	Componentes extraídas
Ângulo de Elevação y_1	3
Ângulo de Azimute y_2	2
Distância y_3	3

Durante a etapa de estimação paramétrica do antecedente, a análise de componentes principais reduziu a dimensionalidade da base de dados de 6 para 2 componentes, retendo mais de 99% da variação do conjunto original. As figuras 48 e 49 mostram a distribuição dos dados no novo espaço de duas dimensões, bem como os graus de pertinência intervalar de cada ponto de dados às duas variáveis.

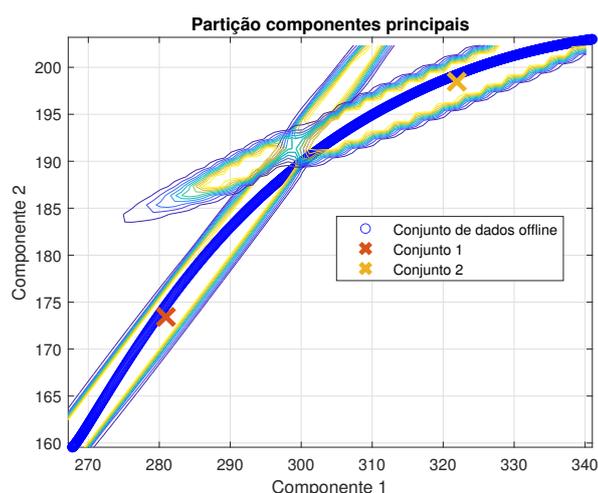


Figura 48 – Partição nebulosa para o conjunto de dados com dimensão reduzida pelo PCA. As curvas de nível representam a os graus de pertinência, valores altos possuem cor amarela.

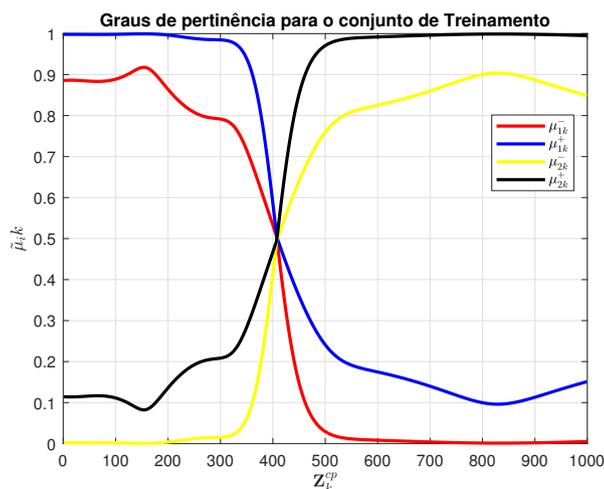


Figura 49 – Funções de pertinência para a etapa *offline* da metodologia.

Serão realizadas duas simulações para avaliar a etapa *online* da metodologia. A primeira foi feita usando apenas a estimação de um passo a frente, enquanto que o segundo teste simulará a falta de medições durante a etapa *online*.

7.2.1 Simulação um passo a frente

Para este tipo de simulação, será avaliado o mecanismo de seleção e atualização de ordem, a acurácia da previsão um passo a frente do modelo, além do tempo médio por iteração necessário para a metodologia atualizar seus parâmetros e calcular sua estimativa. Nesse sentido, as figuras 50-52 apresentam a comparação entre o valor real das variáveis e o intervalo estimado.

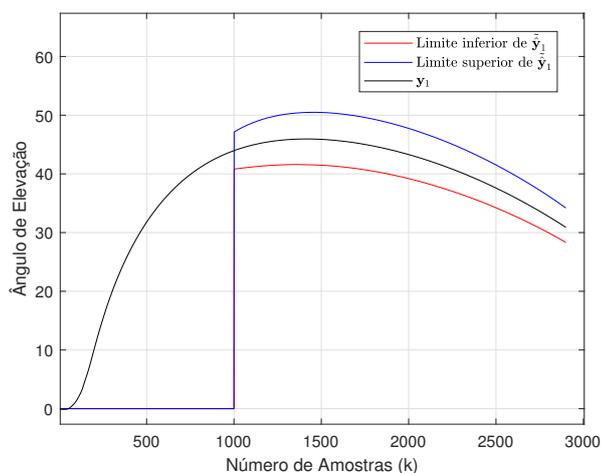


Figura 50 – Comparação entre o intervalo estimado e o valor real do ângulo de elevação do VS-30 em graus.

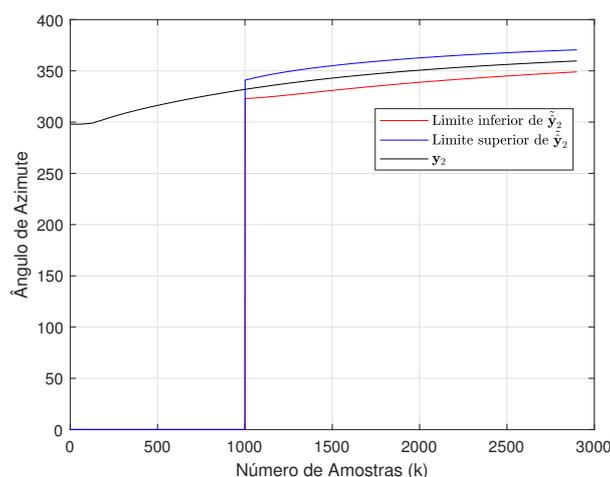


Figura 51 – Comparação entre o intervalo estimado e o valor real do ângulo de azimute do VS-30 em graus.

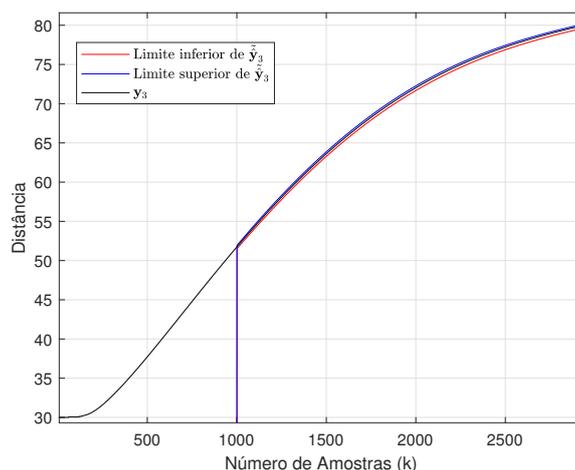


Figura 52 – Comparação entre o intervalo estimado e o valor real da distância percorrida pelo VS-30 em km .

Na figura 50, pode-se observar que o valor real do ângulo de elevação está dentro do intervalo estimado durante toda a etapa *online*. O mesmo ocorre na figura 51, onde é apresentado o resultado obtido para o ângulo de azimute. Na figura 52, a abertura do intervalo estimado é bem menos significativo do que nos outros exemplos, porém ao aplicar um zoom sobre a imagem, é possível observar que a estimativa do modelo está correta durante toda a simulação.

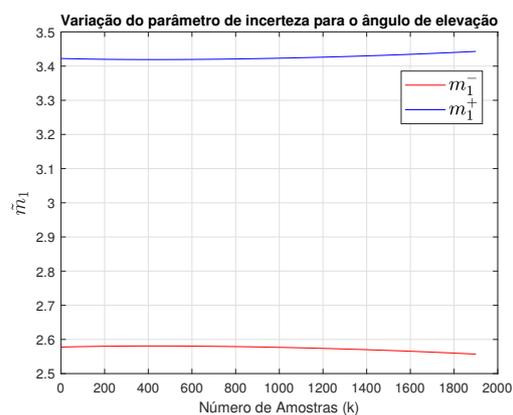
A tabela 14 apresenta a comparação entre a acurácia obtida e outras versões nebulosas do algoritmo ERA/OKID. É possível notar que cada adaptação proposta nesta dissertação apresenta uma leve melhora no resultado final.

Ao observar a abertura do intervalo estimado, percebe-se que este é menos significativo para a distância percorrida. Para analisar a razão por trás deste ocorrido, deve-se observar

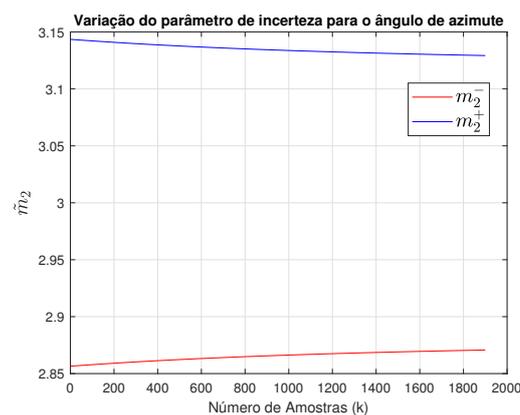
Tabela 14 – Acurácia e erro da estimativa obtida.

Método	ACC
Metodologia proposta	100%
OKID/SSA/T2	95,3%
OKID/SSA/T2/m	97,42%
OKID/SSA/T2/m/PCA	99,45%

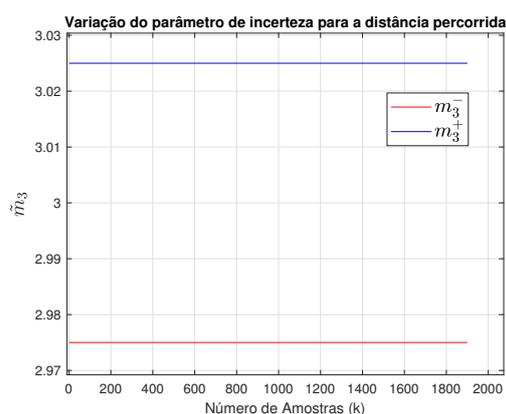
a) Ângulo de Elevação



b) Ângulo de Azimute



c) Distância

Figura 53 – Limites superior e inferior para o expoente de ponderação nebulosa \hat{m} adaptado segundo o nível de incerteza das medições de cada variável.

a seleção e variação do expoente de ponderação nebulosa para cada variável. Esta variação é apresentada pela figura 53.

Ao analisar a variação do expoente de ponderação nebulosa, observa-se que para a distância percorrida, a metodologia sugere um intervalo próximo ao mínimo possível, isso ocorre pois a amplitude do valor medido é alta (na casa dos km), e o ruído estimado é pouco significativo quando comparado a esta. Para todas as variáveis de interesse, percebe-se que o valor estimado permanece aproximadamente constante, isso ocorre pois a covariância do ruído estimado sofreu pouca alteração com a chegada de novas medições.

O mecanismo para seleção da quantidade máxima de regressores selecionou $p_m = 2$ para ambas as regras. Assim, espera-se que a ordem dos submodelos varie entre 6 e 9.

A figura 54 apresenta a variação temporal na ordem de cada submodelo. Nela, pode-se perceber que a ordem do submodelo alocado na segunda regra sofre duas alterações durante a etapa *online*. Inicialmente, a ordem deste submodelo foi selecionada como 9, o que sugere que a melhor aproximação para todas as variáveis foi obtida com 2 regressores. Durante a maior parte da operação *online*, a ordem do submodelo ficou em 6, portanto, a melhor solução é oriunda dos parâmetros de Markov com apenas 1 regressor. Ao final da etapa *online*, o mecanismo atualizou a ordem do modelo para 7, construindo assim um conjunto solução com um número variável de regressores.

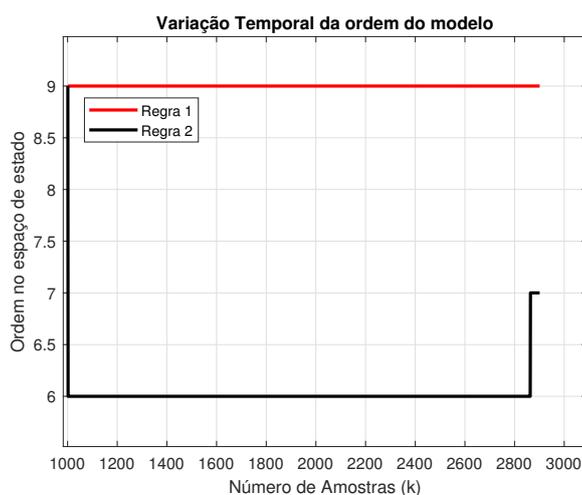


Figura 54 – Variação temporal da ordem para os dois submodelos estimados.

Por se tratar de uma metodologia recursiva que visa uma operação *online*, é importante avaliar o tempo médio por iteração gasto na atualização dos submodelos e na obtenção da estimativa. Para esta aplicação, a metodologia proposta gastou em média $0,0055 \text{ s/it}$, o que permite aplicações com período amostral um décimo de segundo. Este tempo pode ser diminuído ao selecionar uma dimensão inferior para a matriz trajetória. Porém, isso tende a diminuir a acurácia e qualidade da estimativa.

7.2.2 Simulação Mista

Para esta simulação, foram escolhidos arbitrariamente três regiões onde os medidores não informam medições. Nestas, os submodelos realizam as medições baseados em seus parâmetros e configurações selecionadas no momento da falha. Cada uma das falhas tem duração de 150 amostras. A tabela 15 mostra o período de tempo no qual os submodelos não têm acesso a medições.

As figuras 55-57 apresentam o resultado obtido para a simulação com falta de medições. Nelas, as áreas sombreadas representam as regiões onde a simulação livre ocorreu. Percebe-se que em todos os cenários, a simulação livre conseguiu estimar a variável de saída

Tabela 15 – Parâmetros iniciais para a estimação paramétrica da metodologia.

Indicador	Início	Fim
Falha 1	Amostra 1113	Amostra 1263
Falha 2	Amostra 1924	Amostra 2074
Falha 3	Amostra 2597	Amostra 2747

com precisão, sendo bem difícil observar uma diferença no comportamento da estimativa sem medições disponível.

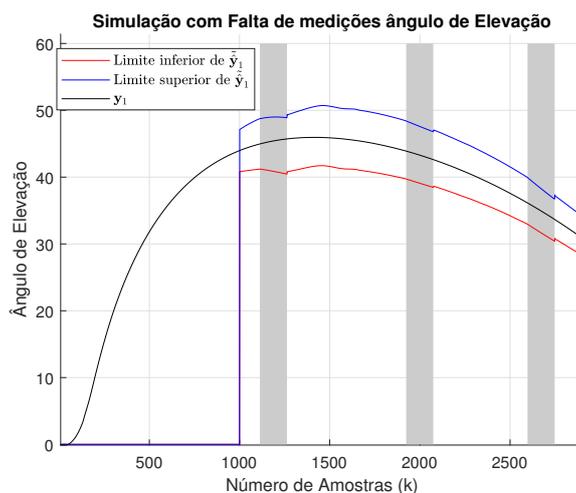


Figura 55 – Simulação com ausência de medições para o ângulo de elevação. As áreas sombreadas representam as regiões onde a simulação livre ocorreu.

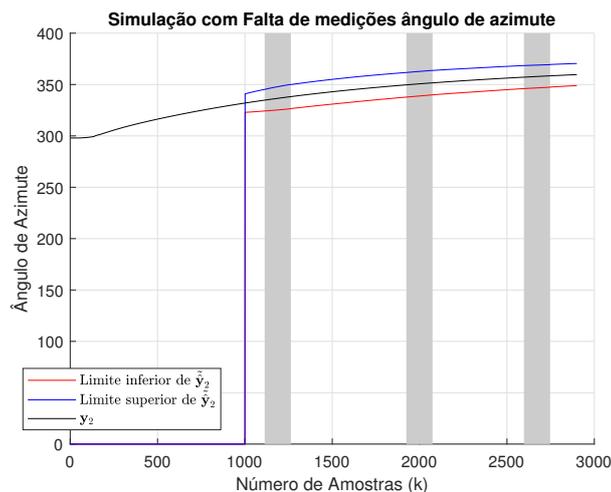


Figura 56 – Simulação com ausência de medições para o ângulo de azimute. As áreas sombreadas representam as regiões onde a simulação livre ocorreu.

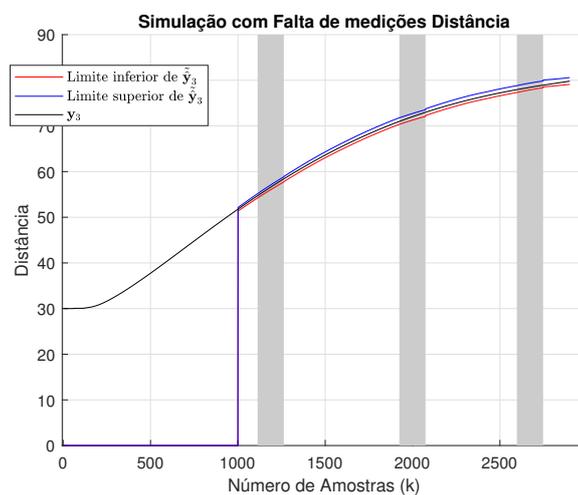


Figura 57 – Simulação com ausência de medições para Distância percorrida. As áreas sombreadas representam as regiões onde a simulação livre ocorreu.

Parte V

CONTEXTUALIZAÇÃO FINAL

8 CONSIDERAÇÕES FINAIS

Nesta dissertação, uma metodologia para projeto de filtros de Kalman nebulosos tipo-2 intervalar baseado em critério de máxima verossimilhança e com adaptação de parâmetros e redução de dimensionalidade foi proposta. A metodologia consiste em uma estimação *offline* e atualizações recursivas. Devido a utilização da lógica nebulosa tipo-2 intervalar, a estimativa do modelo é um intervalo de confiança, o qual é adaptado durante o processamento *offline*.

Para o desenvolvimento da metodologia, foi formulada uma versão em batelada e uma atualização recursiva para a análise espectral singular, a qual se adapta à forma como a metodologia processa os dados, reduzindo operações redundantes. Para a estimação paramétrica do antecedente, foi aplicada a análise de componentes principais para reduzir a dimensionalidade dos dados, e então foi formulado uma versão nebulosa tipo-2 intervalar do algoritmo de agrupamento nebuloso de estimação em máxima verossimilhança (ANT-2IEMV). Este algoritmo evita problemas com indeterminações matemáticas, é inicializado por uma matriz de partição aleatória, e utiliza uma estimativa da covariância do ruído de medição para adaptar \tilde{m} durante a etapa *online*. Na estimação paramétrica do consequente, é formulada uma versão do algoritmo ERA/OKID, na qual a quantidade de regressores ideal para cada variável é selecionada através do critério da informação de Akaike, criando assim submodelos com ordens distintas no espaço de estado. Estes submodelos são atualizados de forma recursiva e refinada, além de possuírem um mecanismo de estabilidade em mudanças de ordem. Por fim, é criada uma adaptação ao filtro, que permite a execução de um esquema de simulação mista.

Os resultados computacionais demonstram a eficiência da metodologia proposta na estimação um passo a frente em ambas as bases de dados. Nestes resultados, pode-se observar a variação temporal da ordem dos submodelos e a evolução do parâmetro de incerteza. Os resultados foram superiores aos outros trabalhos presentes na literatura. Para a simulação mista, a metodologia proposta obteve os melhores resultados para a estimação da velocidade relativa, onde a qualidade da estimativa foi satisfatória em uma situação onde novas medições se tornaram indisponíveis. Em ambas as bases de dados, a análise do tempo computacional não apresentou nenhum problema para implementação.

Os resultados experimentais foram abordados em dois testes. O primeiro deles ocorreu com o foguete de treinamento intermediário (FOGTREIN) do centro de lançamento de Alcântara, obtendo resultados satisfatórios tanto na simulação um passo a frente quanto na simulação mista. Na simulação mista foram simuladas três regiões de ausência de informações. O segundo teste experimental ocorreu com o helicóptero com dois graus de

liberdade. Devido às características da base de dados, foram realizados apenas testes com a simulação um passo a frente. Esta obteve resultados satisfatórios tanto em erro quanto em tempo de processamento.

8.1 Propostas Futuras

Após a análise dos resultados obtidos pela metodologia proposta nesta dissertação, e as características apresentadas pela metodologia, surge o interesse em aplicar os métodos desenvolvidos em outros contextos. Criando assim a seguinte lista de propostas de implementações futuras:

- Desenvolver a metodologia proposta nesta dissertação no contexto dos sistemas nebulosos tipo-2 evolutivos, permitindo que o algoritmo selecione e atualize a quantidade de conjuntos necessários.

Referências

- AGUIRRE, L. A. *Introdução à Identificação de Sistemas—Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. [S.l.]: Editora UFMG, 2007. Citado 6 vezes nas páginas [22](#), [23](#), [25](#), [41](#), [59](#) e [82](#).
- ALLEN, M. R.; SMITH, L. A. Monte carlo ssa: Detecting irregular oscillations in the presence of colored noise. *Journal of climate*, v. 9, n. 12, p. 3373–3404, 1996. Citado na página [52](#).
- ANDERSON, B. D.; MOORE, J. B. *Optimal filtering*. [S.l.]: Courier Corporation, 2012. Citado na página [22](#).
- ASKARI, S. Fuzzy c-means clustering algorithm for data with unequal cluster sizes and contaminated with noise and outliers: Review and development. *Expert Systems with Applications*, Elsevier, v. 165, p. 113856, 2021. Citado na página [31](#).
- ÅSTRÖM, K. J.; EYKHOFF, P. System identification—a survey. *Automatica*, Elsevier, v. 7, n. 2, p. 123–162, 1971. Citado na página [22](#).
- AVVENTI, E.; LINDQUIST, A. G.; WAHLBERG, B. Arma identification of graphical models. *IEEE Transactions on Automatic Control*, IEEE, v. 58, n. 5, p. 1167–1178, 2012. Citado na página [23](#).
- BABUKA, R.; VEEN, P. Van der; KAYMAK, U. Improved covariance estimation for gustafson-kessel clustering. In: IEEE. *2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No. 02CH37291)*. [S.l.], 2002. v. 2, p. 1081–1085. Citado na página [31](#).
- BABUŠKA, R. *Fuzzy modeling for control*. [S.l.]: Springer Science & Business Media, 2012. v. 12. Citado 8 vezes nas páginas [22](#), [25](#), [30](#), [31](#), [33](#), [35](#), [66](#) e [83](#).
- BAR-SHALOM, Y.; LI, X. R.; KIRUBARAJAN, T. *Estimation with applications to tracking and navigation: theory algorithms and software*. [S.l.]: John Wiley & Sons, 2001. Citado na página [25](#).
- BAS, E.; EGRIOGLU, E. A fuzzy regression functions approach based on gustafson-kessel clustering algorithm. *Information Sciences*, Elsevier, v. 592, p. 206–214, 2022. Citado na página [31](#).
- BERGER, V. W.; ZHOU, Y. Kolmogorov–smirnov test: Overview. *Wiley statsref: Statistics reference online*, Wiley Online Library, 2014. Citado na página [87](#).
- BEZDEK, J. C. A physical interpretation of fuzzy isodata. In: *Readings in Fuzzy Sets for Intelligent Systems*. [S.l.]: Elsevier, 1993. p. 615–616. Citado na página [35](#).
- BEZDEK, J. C.; DUNN, J. C. Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions. *IEEE Transactions on Computers*, IEEE, v. 100, n. 8, p. 835–838, 1975. Citado 2 vezes nas páginas [31](#) e [32](#).

- BEZDEK, J. C.; EHRlich, R.; FULL, W. Fcm: The fuzzy c-means clustering algorithm. *Computers & geosciences*, Elsevier, v. 10, n. 2-3, p. 191–203, 1984. Citado na página 35.
- BLITZER, L.; WHEELON, A. D. Oblateness perturbation of elliptical satellite orbits. *Journal of Applied Physics*, American Institute of Physics, v. 28, n. 2, p. 279–279, 1957. Citado na página 85.
- BOWDEN, R. J.; TURKINGTON, D. A. *Instrumental variables*. [S.l.]: Cambridge university press, 1990. Citado na página 23.
- BOZDOGAN, H. Akaike's information criterion and recent developments in information complexity. *Journal of mathematical psychology*, Elsevier, v. 44, n. 1, p. 62–91, 2000. Citado na página 75.
- BRO, R.; SMILDE, A. K. Principal component analysis. *Analytical methods*, Royal Society of Chemistry, v. 6, n. 9, p. 2812–2831, 2014. Citado na página 39.
- CAVANAUGH, J. E.; NEATH, A. A. The akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics*, Wiley Online Library, v. 11, n. 3, p. e1460, 2019. Citado na página 75.
- CHEN, G.; XIE, Q.; SHIEH, L. S. Fuzzy kalman filtering. *Information Sciences*, Elsevier, v. 109, n. 1-4, p. 197–209, 1998. Citado na página 23.
- CHOI, B.-I.; RHEE, F. C.-H. Interval type-2 fuzzy membership function generation methods for pattern recognition. *Information Sciences*, Elsevier, v. 179, n. 13, p. 2102–2122, 2009. Citado na página 39.
- CUI, Z.; DAI, J.; SUN, J.; LI, D.; WANG, L.; WANG, K. Hybrid methods using neural network and kalman filter for the state of charge estimation of lithium-ion battery. *Mathematical Problems in Engineering*, Hindawi Limited, v. 2022, p. 1–11, 2022. Citado na página 22.
- DAYAN, P.; SAHANI, M.; DEBACK, G. Unsupervised learning. *The MIT encyclopedia of the cognitive sciences*, MIT Press, p. 857–859, 1999. Citado na página 30.
- DING, F.; CHEN, T. Identification of hammerstein nonlinear armax systems. *Automatica*, Elsevier, v. 41, n. 9, p. 1479–1489, 2005. Citado na página 23.
- DOVŽAN, D.; ŠKRJANC, I. Recursive clustering based on a gustafson–kessel algorithm. *Evolving systems*, Springer, v. 2, n. 1, p. 15–24, 2011. Citado na página 31.
- DUDA, R. O.; HART, P. E. et al. *Pattern classification and scene analysis*. [S.l.]: Wiley New York, 1973. v. 3. Citado na página 32.
- D'AMICO, S.; BODIN, P.; DELPECH, M.; NOTEBORN, R. Prisma. In: *Distributed space missions for earth system monitoring*. [S.l.]: Springer, 2012. p. 599–637. Citado na página 85.
- ELIASON, S. R. *Maximum likelihood estimation: Logic and practice*. [S.l.]: Sage, 1993. Citado na página 23.

- FRASER, C. T.; ULRICH, S. Adaptive extended kalman filtering strategies for spacecraft formation relative navigation. *Acta Astronautica*, Elsevier, v. 178, p. 700–721, 2021. Citado 2 vezes nas páginas 92 e 100.
- GHAHRAMANI, Z. Unsupervised learning. In: *Summer school on machine learning*. [S.l.]: Springer, 2003. p. 72–112. Citado na página 30.
- GOLYANDINA, N.; SHLEMOV, A. Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series. *arXiv preprint arXiv:1308.4022*, 2013. Citado na página 52.
- GOLYANDINA, N.; ZHIGLJAVSKY, A. Basic ssa. In: _____. *Singular Spectrum Analysis for Time Series*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. p. 11–70. Citado 2 vezes nas páginas 52 e 54.
- GOMES, D. C. dos S.; SERRA, G. L. de O. Computational approach for real-time interval type-2 fuzzy kalman filtering and forecasting via unobservable spectral components of experimental data. *Journal of Control, Automation and Electrical Systems*, Springer, v. 32, p. 326–355, 2021. Citado 2 vezes nas páginas 24 e 25.
- GOMES, D. C. dos S.; SERRA, G. L. de O. Machine learning model for computational tracking and forecasting the covid-19 dynamic propagation. *IEEE Journal of Biomedical and Health Informatics*, IEEE, v. 25, n. 3, p. 615–622, 2021. Citado 2 vezes nas páginas 24 e 25.
- GREENACRE, M.; GROENEN, P. J.; HASTIE, T.; D'ENZA, A. I.; MARKOS, A.; TUZHILINA, E. Principal component analysis. *Nature Reviews Methods Primers*, Nature Publishing Group UK London, v. 2, n. 1, p. 100, 2022. Citado na página 39.
- GREKOUSIS, G.; THOMAS, H. Comparison of two fuzzy algorithms in geodemographic segmentation analysis: The fuzzy c-means and gustafson–kessel methods. *Applied Geography*, Elsevier, v. 34, p. 125–136, 2012. Citado na página 31.
- GUSTAFSON, D. E.; KESSEL, W. C. Fuzzy clustering with a fuzzy covariance matrix. In: IEEE. *1978 IEEE conference on decision and control including the 17th symposium on adaptive processes*. [S.l.], 1979. p. 761–766. Citado na página 31.
- HASSANI, H.; THOMAKOS, D. A review on singular spectrum analysis for economic and financial time series. *Statistics and its Interface*, International Press of Boston, v. 3, n. 3, p. 377–397, 2010. Citado na página 52.
- HASSANI, H.; ZHIGLJAVSKY, A. Singular spectrum analysis: methodology and application to economics data. *Journal of Systems Science and Complexity*, Springer, v. 22, p. 372–394, 2009. Citado na página 52.
- HATTEN, N.; RUSSELL, R. P. A smooth and robust harris-priester atmospheric density model for low earth orbit applications. *Advances in Space Research*, Elsevier, v. 59, n. 2, p. 571–586, 2017. Citado na página 86.
- HENDRICKS, E.; JANNERUP, O.; SØRENSEN, P. H. *Linear systems control: deterministic and stochastic methods*. [S.l.]: Springer, 2008. Citado na página 59.

- HOSSAIN, M.; HAQUE, M.; ARIF, M. T. Kalman filtering techniques for the online model parameters and state of charge estimation of the li-ion batteries: A comparative analysis. *Journal of Energy Storage*, Elsevier, v. 51, p. 104174, 2022. Citado na página 22.
- HUO, Y.; CAI, Z.; GONG, W.; LIU, Q. A new adaptive kalman filter by combining evolutionary algorithm and fuzzy inference system. In: IEEE. *2014 IEEE Congress on Evolutionary Computation (CEC)*. [S.l.], 2014. p. 2893–2900. Citado na página 24.
- HWANG, C.; RHEE, F. C.-H. Uncertain fuzzy clustering: Interval type-2 fuzzy approach to *c*-means. *IEEE Transactions on fuzzy systems*, IEEE, v. 15, n. 1, p. 107–120, 2007. Citado 3 vezes nas páginas 30, 37 e 38.
- JING, L.; DENG, D.; YU, J. Weighting exponent selection of fuzzy *c*-means via jacobian matrix. In: SPRINGER. *Knowledge Science, Engineering and Management: 7th International Conference, KSEM 2014, Sibiu, Romania, October 16-18, 2014. Proceedings 7*. [S.l.], 2014. p. 115–126. Citado na página 35.
- JOHANSSON, R.; ROBERTSSON, A.; NILSSON, K.; VERHAEGEN, M. State-space system identification of robot manipulator dynamics. *Mechatronics*, Elsevier, v. 10, n. 3, p. 403–418, 2000. Citado na página 41.
- JOLLIFFE, I. T. Introduction. In: _____. *Principal Component Analysis*. New York, NY: Springer New York, 2002. p. 1–9. ISBN 978-0-387-22440-4. Citado 2 vezes nas páginas 39 e 40.
- JUANG, J.-N. *Applied system identification*. [S.l.]: Prentice-Hall, Inc., 1994. Citado 7 vezes nas páginas 23, 41, 49, 50, 51, 73 e 76.
- JUANG, J.-N.; PHAN, M.; HORTA, L. G.; LONGMAN, R. W. Identification of observer/kalman filter markov parameters-theory and experiments. *Journal of Guidance, Control, and Dynamics*, v. 16, n. 2, p. 320–329, 1993. Citado na página 73.
- JÚNIOR, S. E. R.; SERRA, G. L. de O. Intelligent forecasting of time series based on evolving distributed neuro-fuzzy network. *Computational Intelligence*, Wiley Online Library, v. 36, n. 3, p. 1394–1413, 2020. Citado na página 52.
- JÚNIOR, S. E. R.; SERRA, G. L. de O. An approach for evolving neuro-fuzzy forecasting of time series based on parallel recursive singular spectrum analysis. *Fuzzy Sets and Systems*, Elsevier, v. 443, p. 1–29, 2022. Citado 3 vezes nas páginas 61, 62 e 63.
- KALHORI, M. R. N.; ZARANDI, M. F. Interval type-2 credibilistic clustering for pattern recognition. *Pattern Recognition*, Elsevier, v. 48, n. 11, p. 3652–3672, 2015. Citado na página 39.
- KALMAN, R. E. A new approach to linear filtering and prediction problems. 1960. Citado na página 22.
- KARIYA, T.; KURATA, H. *Generalized least squares*. [S.l.]: John Wiley & Sons, 2004. Citado na página 23.
- KARNIK, N. N.; MENDEL, J. M. Introduction to type-2 fuzzy logic systems. In: IEEE. *1998 IEEE international conference on fuzzy systems proceedings. IEEE world congress on computational intelligence (Cat. No. 98CH36228)*. [S.l.], 1998. v. 2, p. 915–920. Citado na página 36.

- KEFER, K.; HANGHOFER, R.; KEFER, P.; STÖGER, M.; HOFER, B.; AFFENZELLER, M.; WINKLER, S. Simulation-based optimization of residential energy flows using white box modeling by genetic programming. *Energy and Buildings*, Elsevier, v. 258, p. 111829, 2022. Citado na página 22.
- KRIEGEL, H.-P.; KRÖGER, P.; SANDER, J.; ZIMEK, A. Density-based clustering. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, Wiley Online Library, v. 1, n. 3, p. 231–240, 2011. Citado na página 30.
- LIANG, Q.; MENDEL, J. M. Interval type-2 fuzzy logic systems: theory and design. *IEEE Transactions on Fuzzy systems*, IEEE, v. 8, n. 5, p. 535–550, 2000. Citado 3 vezes nas páginas 30, 36 e 37.
- LIAO, G.; LING, B. W.-K.; HO, C. Y.-F. Alzheimer’s detection with speech using singular spectrum analysis trend extraction method. In: IEEE. *2022 13th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*. [S.l.], 2022. p. 417–421. Citado na página 52.
- LINDA, O.; MANIC, M. General type-2 fuzzy c-means algorithm for uncertain fuzzy clustering. *IEEE Transactions on Fuzzy Systems*, IEEE, v. 20, n. 5, p. 883–897, 2012. Citado na página 38.
- LIU, H.-C.; YIH, J.-M.; WU, D.-B.; LIU, S.-W. Fuzzy c-mean clustering algorithms based on picard iteration and particle swarm optimization. In: IEEE. *2008 international workshop on education technology and training & 2008 international workshop on geoscience and remote sensing*. [S.l.], 2008. v. 2, p. 838–842. Citado na página 33.
- LJUNG, L. System identification. In: *Signal analysis and prediction*. [S.l.]: Springer, 1998. p. 163–173. Citado na página 22.
- LOYOLA-GONZALEZ, O. Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view. *IEEE Access*, IEEE, v. 7, p. 154096–154113, 2019. Citado na página 22.
- LU, Y.; MA, T.; YIN, C.; XIE, X.; TIAN, W.; ZHONG, S. Implementation of the fuzzy c-means clustering algorithm in meteorological data. *International Journal of Database Theory and Application*, v. 6, n. 6, p. 1–18, 2013. Citado na página 31.
- MADBOULY, E. E.; ABDALLA, A.; BANBY, G. M. E. Fuzzy adaptive kalman filter for multi-sensor system. In: IEEE. *2009 International Conference on Networking and Media Convergence*. [S.l.], 2009. p. 141–145. Citado na página 24.
- MADHULATHA, T. S. An overview on clustering methods. *arXiv preprint arXiv:1205.1117*, 2012. Citado na página 30.
- MASYITOH, A. D.; ZUKHRONAH, E.; SLAMET, I. Forecasting for monthly rainfall in gondang reservoir lamongan using singular spectrum analysis (ssa) method. In: AIP PUBLISHING. *AIP Conference Proceedings*. [S.l.], 2022. v. 2391, n. 1. Citado na página 52.
- MENDEL, J. M.; JOHN, R. I.; LIU, F. Interval type-2 fuzzy logic systems made simple. *IEEE transactions on fuzzy systems*, IEEE, v. 14, n. 6, p. 808–821, 2006. Citado na página 37.

- MOHAN, K. K.; VERMA, A.; SRIVIDYA, A.; PAPIC, L. Integration of black-box and white-box modeling approaches for software reliability estimation. *International Journal of Reliability, Quality and Safety Engineering*, World Scientific, v. 17, n. 03, p. 261–273, 2010. Citado na página [22](#).
- MONTEL, B.-H. M. M.; SERRA, G. L. de O. A data-driven interval type-2 fuzzy kalman filter of minimum realization for forecasting spacecraft formation on low earth orbit. In: *Global Perspectives on Robotics and Autonomous Systems: Development and Applications*. [S.l.]: IGI Global, 2023. p. 93–129. Citado 2 vezes nas páginas [92](#) e [100](#).
- MOOR, B. D.; OVERSCHEE, P. V.; FAVOREEL, W. Algorithms for subspace state-space system identification: an overview. *Applied and Computational Control, Signals, and Circuits: Volume 1*, Springer, p. 247–311, 1999. Citado na página [41](#).
- MORGAN, D.; CHUNG, S.-J.; BLACKMORE, L.; ACIKMESE, B.; BAYARD, D.; HADAEGH, F. Y. Swarm-keeping strategies for spacecraft under j_2 and atmospheric drag perturbations. *Journal of Guidance, Control, and Dynamics*, v. 35, n. 5, p. 1492–1506, 2012. Citado na página [86](#).
- MYUNG, I. J. Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, Elsevier, v. 47, n. 1, p. 90–100, 2003. Citado na página [23](#).
- NAYAK, J.; NAIK, B.; BEHERA, H. Fuzzy c-means (fcm) clustering algorithm: a decade review from 2000 to 2014. In: SPRINGER. *Computational Intelligence in Data Mining-Volume 2: Proceedings of the International Conference on CIDM, 20-21 December 2014*. [S.l.], 2015. p. 133–149. Citado na página [31](#).
- NEWHEY, W. K.; POWELL, J. L. Instrumental variable estimation of nonparametric models. *Econometrica*, Wiley Online Library, v. 71, n. 5, p. 1565–1578, 2003. Citado na página [23](#).
- NGO, L. T.; MAI, D. S.; PEDRYCZ, W. Semi-supervising interval type-2 fuzzy c-means clustering with spatial information for multi-spectral satellite image classification and change detection. *Computers & geosciences*, Elsevier, v. 83, p. 1–16, 2015. Citado na página [39](#).
- OBEID, S.; AHMADI, G.; JHA, R. Narmax identification based closed-loop control of flow separation over naca 0015 airfoil. *Fluids*, MDPI, v. 5, n. 3, p. 100, 2020. Citado na página [23](#).
- OLIVEIRA, J. V. D.; PEDRYCZ, W. *Advances in fuzzy clustering and its applications*. [S.l.]: John Wiley & Sons, 2007. Citado na página [30](#).
- PAL, N. R.; BEZDEK, J. C. On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy systems*, IEEE, v. 3, n. 3, p. 370–379, 1995. Citado na página [35](#).
- PÁRAMO-CARRANZA, L. A.; MEDA-CAMPAÑA, J. A.; RUBIO, J. de J.; TAPIA-HERRERA, R.; CURTIDOR-LÓPEZ, A.; GRANDE-MEZA, A.; CÁZARES-RAMÍREZ, I. Discrete-time kalman filter for takagi-sugeno fuzzy models. *Evolving Systems*, Springer, v. 8, n. 3, p. 211–219, 2017. Citado na página [24](#).

- PEDDANENI, H.; ERDOGMUS, D.; RAO, Y. N.; HEGDE, A.; PRINCIPE, J. C. Recursive principal components analysis using eigenvector matrix perturbation. In: IEEE. *Proceedings of the 2004 14th IEEE Signal Processing Society Workshop Machine Learning for Signal Processing, 2004*. [S.l.], 2004. p. 83–92. Citado na página 56.
- PEDRYCZ, W.; GOMIDE, F. *Fuzzy systems engineering: toward human-centric computing*. [S.l.]: John Wiley & Sons, 2007. Citado na página 31.
- PEETERS, L.; BEINTEMA, G. I.; FORGIONE, M.; SCHOUKENS, M. Narx identification using derivative-based regularized neural networks. In: IEEE. *2022 IEEE 61st Conference on Decision and Control (CDC)*. [S.l.], 2022. p. 1515–1520. Citado na página 23.
- PEI, J.; YANG, X.; GAO, X.; XIE, W. Weighting exponent m in fuzzy c-means (fcm) clustering algorithm. In: SPIE. *Object Detection, Classification, and Tracking Technologies*. [S.l.], 2001. v. 4554, p. 246–251. Citado 3 vezes nas páginas 35, 38 e 66.
- PINTELAS, E.; LIVIERIS, I. E.; PINTELAS, P. A grey-box ensemble model exploiting black-box accuracy and white-box intrinsic interpretability. *Algorithms*, MDPI, v. 13, n. 1, p. 17, 2020. Citado na página 22.
- PIRES, D.; SERRA, G. An approach for fuzzy kalman filter modeling based on evolving clustering of experimental data. *Journal of Intelligent & Fuzzy Systems*, IOS Press, v. 35, n. 2, p. 1819–1834, 2018. Citado 3 vezes nas páginas 23, 24 e 25.
- PIRES, D. S.; SERRA, G. L. d. O. Methodology for evolving fuzzy kalman filter identification. *International Journal of Control, Automation and Systems*, Springer, v. 17, n. 3, p. 793–800, 2019. Citado 2 vezes nas páginas 24 e 25.
- PITCHFORTH, D. J.; ROGERS, T. J.; TYGESEN, U. T.; CROSS, E. J. Grey-box models for wave loading prediction. *Mechanical Systems and Signal Processing*, Elsevier, v. 159, p. 107741, 2021. Citado na página 22.
- QIU, C.; XIAO, J.; YU, L.; HAN, L.; IQBAL, M. N. A modified interval type-2 fuzzy c-means algorithm with application in mr image segmentation. *Pattern Recognition Letters*, Elsevier, v. 34, n. 12, p. 1329–1338, 2013. Citado 2 vezes nas páginas 30 e 37.
- RICHARDS, F. S. A method of maximum-likelihood estimation. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Oxford University Press, v. 23, n. 2, p. 469–475, 1961. Citado na página 23.
- ROBIN, W. Solving differential equations using modified picard iteration. *International Journal of Mathematical Education in Science and Technology*, Taylor & Francis, v. 41, n. 5, p. 649–665, 2010. Citado na página 33.
- Rocha Filho, O. D.; de Oliveira Serra, G. L. Recursive fuzzy instrumental variable based evolving neuro-fuzzy identification for non-stationary dynamic system in a noisy environment. *Fuzzy Sets and Systems*, v. 338, p. 50–89, 2018. ISSN 0165-0114. Theme: Fuzzy Systems. Citado 2 vezes nas páginas 25 e 33.
- ROKACH, L.; MAIMON, O. Clustering methods. *Data mining and knowledge discovery handbook*, Springer, p. 321–352, 2005. Citado na página 30.

- ROSCOE, C. W. T.; VADALI, S. R.; ALFRIEND, K. T. Third-body perturbation effects on satellite formations. *The Journal of the Astronautical Sciences*, Springer, v. 60, n. 3, p. 408–433, 2013. Citado na página 86.
- RUBIO, E.; CASTILLO, O. Interval type-2 fuzzy clustering for membership function generation. In: IEEE. *2013 IEEE Workshop on Hybrid Intelligent Models and Applications (HIMA)*. [S.l.], 2013. p. 13–18. Citado na página 39.
- SARIOĞLU, A.; KURAL, A. Modeling and arx identification of a quadrotor miniuav. In: IEEE. *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*. [S.l.], 2015. p. 1196–1200. Citado na página 23.
- SASIADEK, J.; WANG, Q.; ZEREMBA, M. Fuzzy adaptive kalman filtering for ins/gps data fusion. In: IEEE. *Proceedings of the 2000 IEEE International Symposium on Intelligent Control. Held jointly with the 8th IEEE Mediterranean Conference on Control and Automation (Cat. No. 00CH37147)*. [S.l.], 2000. p. 181–186. Citado na página 24.
- SERRA, G. L. de O. *Propostas de metodologias para identificação e controle inteligentes*. Tese (Doutorado) — [sn], 2005. Citado na página 22.
- SIMON, D. Kalman filtering for fuzzy discrete time dynamic systems. *Applied Soft Computing*, v. 3, n. 3, p. 191–207, 2003. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494603000346>>. Citado na página 24.
- SJÖBERG, J.; ZHANG, Q.; LJUNG, L.; BENVENISTE, A.; DELYON, B.; GLORENNEC, P.-Y.; HJALMARSSON, H.; JUDITSKY, A. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, Elsevier, v. 31, n. 12, p. 1691–1724, 1995. Citado na página 22.
- SOARES, A. R.; BOTTURA, C. Identificação de séries temporais multivariadas no espaço de estado pelo método de akaike baseado em correlações canônicas com parâmetro iterativo. *Anais do X Simpósio Brasileiro de Automação Inteligente, Universidade Federal de Sao Joao del-Rei, Sao Joao del-Rei, MG, Brasil*, p. 314–319, 2011. Citado na página 41.
- SÖDERSTRÖM, T.; STOICA, P. Instrumental variable methods for system identification. *Circuits, Systems and Signal Processing*, Springer, v. 21, n. 1, p. 1–9, 2002. Citado na página 23.
- SORENSEN, H. W. Least-squares estimation: from gauss to kalman. *IEEE spectrum*, IEEE, v. 7, n. 7, p. 63–68, 1970. Citado na página 23.
- STREJC, V. Least squares parameter estimation. *Automatica*, Elsevier, v. 16, n. 5, p. 535–550, 1980. Citado na página 23.
- SUNG, W.; LEE, S.; YOU, K. Ultra-precision positioning using adaptive fuzzy-kalman filter observer. *Precision Engineering*, v. 34, n. 1, p. 195–199, 2010. ISSN 0141-6359. CIRP-CAT 2007. Citado na página 24.
- TRAUWAERT, E.; KAUFMAN, L.; ROUSSEEUW, P. Fuzzy clustering algorithms based on the maximum likelihood principle. *Fuzzy Sets and Systems*, Elsevier, v. 42, n. 2, p. 213–227, 1991. Citado na página 31.

- UMAMAGESWARI, A.; IGNATIOUS, J. J.; VINODHA, R. A comparative study of kalman filter, extended kalman filter and unscented kalman filter for harmonic analysis of the non-stationary signals. *International Journal of Scientific & Engineering Research*, Citeseer, v. 3, n. 7, p. 1–9, 2012. Citado na página 25.
- VALASEK, J.; CHEN, W. Observer/kalman filter identification for online system identification of aircraft. *Journal of Guidance, Control, and Dynamics*, v. 26, n. 2, p. 347–353, 2003. Citado na página 41.
- VAUTARD, R.; GHIL, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D: Nonlinear Phenomena*, Elsevier, v. 35, n. 3, p. 395–424, 1989. Citado na página 52.
- VELEZ, M.; JAMSHIDI, P.; SIEGMUND, N.; APEL, S.; KÄSTNER, C. White-box analysis over machine learning: Modeling performance of configurable systems. In: IEEE. *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*. [S.l.], 2021. p. 1072–1084. Citado na página 22.
- VERHAEGEN, M. Identification of the deterministic part of mimo state space models given in innovations form from input-output data. *Automatica*, Elsevier, v. 30, n. 1, p. 61–74, 1994. Citado na página 41.
- VOKROUHLICKY, D.; FARINELLA, P.; MIGNARD, F. Solar radiation pressure perturbations for earth satellites. 1: A complete theory including penumbra transitions. *Astronomy and Astrophysics*, v. 280, p. 295–312, 1993. Citado na página 86.
- WAN, E. A.; MERWE, R. V. D. The unscented kalman filter for nonlinear estimation. In: IEEE. *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*. [S.l.], 2000. p. 153–158. Citado na página 25.
- WEI, N.; YIN, L.; LI, C.; LI, C.; CHAN, C.; ZENG, F. Forecasting the daily natural gas consumption with an accurate white-box model. *Energy*, Elsevier, v. 232, p. 121036, 2021. Citado na página 22.
- WINDMANN, S.; HAEB-UMBACH, R. Parameter estimation of a state-space model of noise for robust speech recognition. *IEEE transactions on audio, speech, and language processing*, IEEE, v. 17, n. 8, p. 1577–1590, 2009. Citado na página 59.
- WU, D. Approaches for reducing the computational cost of interval type-2 fuzzy logic systems: overview and comparisons. *IEEE Transactions on Fuzzy Systems*, IEEE, v. 21, n. 1, p. 80–99, 2012. Citado na página 37.
- WU, T.; LIU, X.-W. An interval type-2 fuzzy clustering solution for large-scale multiple-criteria group decision-making problems. *Knowledge-Based Systems*, Elsevier, v. 114, p. 118–127, 2016. Citado na página 30.
- WUNSCH, D.; XU, R. *Clustering*. [S.l.]: John Wiley & Sons, 2008. Citado na página 30.
- XU, J.; FENG, G.; ZHAO, T.; SUN, X.; ZHU, M. Remote sensing image classification based on semi-supervised adaptive interval type-2 fuzzy c-means algorithm. *Computers & geosciences*, Elsevier, v. 131, p. 132–143, 2019. Citado na página 37.

- XU, Y.; HU, M.; ZHOU, A.; LI, Y.; LI, S.; FU, C.; GONG, C. State of charge estimation for lithium-ion batteries based on adaptive dual kalman filter. *Applied Mathematical Modelling*, Elsevier, v. 77, p. 1255–1272, 2020. Citado na página 22.
- YIN, Y.; SHENG, Y.; QIN, J. Interval type-2 fuzzy c-means forecasting model for fuzzy time series. *Applied Soft Computing*, Elsevier, v. 129, p. 109574, 2022. Citado 2 vezes nas páginas 30 e 37.
- YU, C.; LJUNG, L.; VERHAEGEN, M. Identification of structured state-space models. *Automatica*, Elsevier, v. 90, p. 54–61, 2018. Citado na página 41.
- YU, J.; CHENG, Q.; HUANG, H. Analysis of the weighting exponent in the fcm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, IEEE, v. 34, n. 1, p. 634–639, 2004. Citado na página 35.
- YU, X.; SKEIE, K. S.; KNUDSEN, M. D.; REN, Z.; IMSLAND, L.; GEORGES, L. Influence of data pre-processing and sensor dynamics on grey-box models for space-heating: Analysis using field measurements. *Building and Environment*, Elsevier, v. 212, p. 108832, 2022. Citado na página 22.
- ZHOU, L.; LUO, Z.; SHEN, T.; ZHANG, J.; ZHEN, M.; YAO, Y.; FANG, T.; QUAN, L. Kfnet: Learning temporal camera relocalization using kalman filtering. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 4919–4928. Citado na página 22.