

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DETECÇÃO DE FALHAS EM PROCESSOS INDUSTRIAIS
OPERANDO EM MÚLTIPLAS REGIÕES VIA ANÁLISE
EXTERNA COM MÚLTIPLOS MODELOS LINEARES**

DISSERTAÇÃO DE MESTRADO

RENATA TEIXEIRA DAS NEVES FERNANDES

VITÓRIA

2018

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DETECÇÃO DE FALHAS EM PROCESSOS INDUSTRIAIS
OPERANDO EM MÚLTIPLAS REGIÕES VIA ANÁLISE
EXTERNA COM MÚLTIPLOS MODELOS LINEARES**

RENATA TEIXEIRA DAS NEVES FERNANDES

Dissertação de Mestrado apresentada como
requisito parcial para obtenção do título de
Mestre em Engenharia Elétrica.

Orientador:
Prof. Dr. Celso José Munaro.

VITÓRIA
2018

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial Tecnológica,
Universidade Federal do Espírito Santo, ES, Brasil)

F363d Fernandes, Renata Teixeira das Neves, 1991-
Detecção de falhas em processos industriais operando em
múltiplas regiões via análise externa com múltiplos modelos
lineares / Renata Teixeira das Neves Fernandes. – 2018.
79 f. : il.

Orientador: Celso José Munaro.
Dissertação (Mestrado em Engenharia Elétrica) –
Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Controle de processo industrial. 2. Localização de falhas
(Engenharia). 3. Modelos lineares (Estatística). 4. Reator CSTR.
5. Análise externa. I. Munaro, Celso José. II. Universidade
Federal do Espírito Santo. Centro Tecnológico. III. Título.

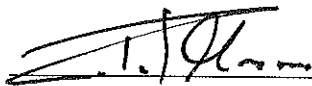
CDU: 621.3

RENATA TEIXEIRA DAS NEVES FERNANDES

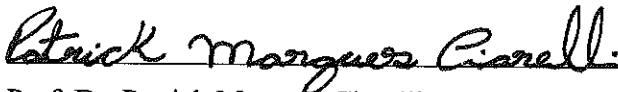
**DETECÇÃO DE FALHAS EM PROCESSOS INDUSTRIAIS
OPERANDO EM MÚLTIPLAS REGIÕES VIA ANÁLISE
EXTERNA COM MÚLTIPLOS MODELOS LINEARES**

Dissertação submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Engenharia Elétrica.

COMISSÃO EXAMINADORA



Prof. Dr. Celso José Munaro
Universidade Federal do Espírito Santo
Orientador



Prof. Dr. Patrick Marques Ciarelli
Universidade Federal do Espírito Santo



Prof. Dr. Marco Antonio de Souza Leite Cuadros
Instituto Federal do Espírito Santo

RESUMO

Métodos convencionais de controle estatístico multivariado vêm sendo muito aplicados para monitoração de processos industriais. Esses métodos assumem que o processo opera em uma única região de operação, ou seja, em um estado estacionário. Para processos que operam em múltiplas regiões de operação são necessários métodos de monitoração que consigam diferenciar mudanças normais nas condições de operação da ocorrência de falhas, caso contrário, altas taxas de falsos alarmes são geradas. Vários estudos recentes propõem soluções para este problema, dentre elas a análise externa, que dispensa a definição de regiões de operação delimitadas e de treinamentos para cada uma delas. Porém, quando há a presença de não linearidades em processos, a aplicação da análise externa linear prejudica a qualidade da monitoração. Nesses casos, uma alternativa é utilizar a análise externa não linear. No entanto, esta solução requer definições complexas de funções e parâmetros que influenciam consideravelmente na qualidade do modelo final. Uma alternativa à análise externa não linear é a utilização de múltiplos modelos lineares na análise externa. Este trabalho tem como objetivo principal investigar a possibilidade da utilização de múltiplos modelos lineares aplicados à análise externa para tratar a presença de não linearidades no processo. Para tal, propõe-se uma metodologia para detectar a necessidade do uso desses múltiplos modelos e para realizar a construção e a aplicação dos múltiplos modelos na monitoração do processo. A metodologia proposta é aplicada na supervisão do processo industrial simulado de um tanque reator continuamente agitado (CSTR) e seu desempenho é comparado ao obtido através da metodologia de análise externa linear tradicional que utiliza um único modelo linear. Os resultados mostraram uma melhoria no desempenho da detecção de falhas através da aplicação da metodologia de múltiplos modelos, que se mostrou mais sensível a falhas que a metodologia de modelo único. Dessa forma, com base nos resultados obtidos neste trabalho, a aplicação de múltiplos modelos lineares mostra-se uma alternativa viável na detecção de falhas em processos industriais.

PALAVRAS-CHAVE: Controle de processos industriais, Detecção de falhas, Análise externa, CSTR, Análise de componentes principais.

ABSTRACT

The conventional methods of Multivariate Statistical Monitoring have been widely applied to industrial processes monitoring. A unique operation condition is usually assumed in these methods, i.e., they operate in a stationary state. To handle processes that operate in multiple operation conditions, monitoring methods capable of differentiating normal changes at operation conditions from processes faults are required. Otherwise, high levels of false alarms are generated. Several recent studies propose solutions to this problem, among them the external analysis, which does not require the definition of delimited operating regions and neither trainings for each of them. However, when nonlinearities are present in processes, the application of linear external analysis impairs the quality of the monitoring. In these cases, non-linear external analysis can be used as an alternative. This solution requires complex definitions of functions and parameters that has significant influence in the quality of the model. Another option is the use of multiple linear models in external analysis. The main goal of this work is to investigate the possibility of using multiple linear models applied to the external analysis to deal with the presence of nonlinearities in the process. A methodology is proposed to detect if multiple linear models are necessary and to carry out the construction and the application of the multiple models in the process monitoring. The proposed methodology is applied in the supervision of the simulated industrial process of a Continuously Stirred Reactor Tank (CSTR) and its performance is compared with the performance of the traditional external analysis method using a single linear model. The results have shown improvement in the performance of fault detection when applied the multiple models methodology, which has also presented better sensibility to faults than the single model methodology. Thus, based on the results obtained in this work, the application of multiple linear models is a viable alternative for fault detections in industrial processes.

KEYWORDS: Industrial processes monitoring, Fault detection, External analysis, CSTR, Principal component analysis.

LISTA DE FIGURAS

Figura 1 – Gráfico de Shewhart aplicado à análise estatística univariável	20
Figura 2 – Metodologia da Análise Externa.....	26
Figura 3 - Fluxograma da etapa: obtenção dos modelos lineares.....	39
Figura 4 - Fluxograma para a etapa de obtenção do modelo PCA.....	40
Figura 5 - Fluxograma detalhado para a fase de monitoração.....	41
Figura 6 - Diagrama do CSTR.....	42
Figura 7 - Variáveis do CSTR da matriz de dados \mathbf{X}	45
Figura 8 - Variáveis do CSTR com as primeiras 2000 amostras da matriz de dados \mathbf{X} . 45	
Figura 9 – Correlações de Spearman calculadas entre as VP e as VE	46
Figura 10 - Variáveis do CSTR com as primeiras 5000 amostras da matriz de dados \mathbf{X} 47	
Figura 11 - \varphiSTAT calculado para $h14$	47
Figura 12 - Amostras 250 a 1000 do \varphiSTAT calculado para $h14$	47
Figura 13 - \varphiSTAT calculado para $h13$	48
Figura 14 - Amostras 250 a 1000 do \varphiSTAT calculado para $h13$	48
Figura 15 - \varphiSTAT calculado para $h12$	48
Figura 16 - Amostras 250 a 1000 do \varphiSTAT calculado para $h12$	48
Figura 17 - \varphiSTAT calculado para $h9$	48
Figura 18 - Amostras 250 a 1000 do \varphiSTAT calculado para $h9$	48
Figura 19 - \varphiSTAT calculado para $h8$	49
Figura 20 - Amostras 250 a 1000 do \varphiSTAT calculado para $h8$	49
Figura 21 - \varphiSTAT calculado para $h6$	49
Figura 22 - Amostras 250 a 1000 do \varphiSTAT calculado para $h6$	49
Figura 23 - \varphiSTAT calculado para $h5$	49
Figura 24 - Amostras 250 a 1000 do \varphiSTAT calculado para $h5$	49
Figura 25 – Coeficientes obtidos com o fator de esquecimento 0.8.....	50
Figura 26 – Coeficientes obtidos com o fator de esquecimento 0.9.....	51
Figura 27 – Coeficientes obtidos com o fator de esquecimento 0.95.....	51
Figura 28 – Coeficientes obtidos com o fator de esquecimento 0.97.....	52
Figura 29 – Coeficientes obtidos com o fator de esquecimento 0.99.....	52
Figura 30 – Vetor de coeficientes $\hat{\theta}_{5,1,p}$ obtidos via RLS.....	53
Figura 31 – Vetor de coeficientes $\hat{\theta}_{5,2,p}$ obtidos via RLS.....	53
Figura 32 – Vetor de coeficientes $\hat{\theta}_{6,1,p}$ obtidos via RLS.....	54
Figura 33 – Vetor de coeficientes $\hat{\theta}_{6,2,p}$ obtidos via RLS.....	54
Figura 34 – Vetor de coeficientes $\hat{\theta}_{8,1,p}$ obtidos via RLS.....	54
Figura 35 – Vetor de coeficientes $\hat{\theta}_{8,2,p}$ obtidos via RLS.....	54
Figura 36 – Vetor de coeficientes $\hat{\theta}_{8,3,p}$ obtidos via RLS.....	54
Figura 37 – Vetor de coeficientes $\hat{\theta}_{8,10,p}$ obtidos via RLS	54
Figura 38 – Vetor de coeficientes $\hat{\theta}_{9,2,p}$ obtidos via RLS.....	55
Figura 39 – Vetor de coeficientes $\hat{\theta}_{12,2,p}$ obtidos via RLS	55

Figura 40 – Vetor de coeficientes $\hat{\theta}_{13,1,p}$ obtidos via RLS	55
Figura 41 – Vetor de coeficientes $\hat{\theta}_{13,2,p}$ obtidos via RLS	55
Figura 42 – Vetor de coeficientes $\hat{\theta}_{13,3,p}$ obtidos via RLS	55
Figura 43 – Vetor de coeficientes $\hat{\theta}_{13,10,p}$ obtidos via RLS	55
Figura 44 – Vetor de coeficientes $\hat{\theta}_{13,11,p}$ obtidos via RLS	56
Figura 45 – Vetor de coeficientes $\hat{\theta}_{14,1,p}$ obtidos via RLS	56
Figura 46 – Vetor de coeficientes $\hat{\theta}_{14,2,p}$ obtidos via RLS	56
Figura 47 – Vetor de coeficientes $\hat{\theta}_{14,3,p}$ obtidos via RLS	56
Figura 48 – Vetor de coeficientes $\hat{\theta}_{14,10,p}$ obtidos via RLS	56
Figura 49 – Análise de $\hat{\theta}_{5,1,p}$ para seleção de $\theta_{5,1,k}$	57
Figura 50 – Análise de $\hat{\theta}_{5,2,p}$ para seleção de $\theta_{5,2,k}$	57
Figura 51 – Análise de $\hat{\theta}_{6,1,p}$ para seleção de $\theta_{6,1,k}$	58
Figura 52 – Análise de $\hat{\theta}_{6,2,p}$ para seleção de $\theta_{6,2,k}$	58
Figura 53 – Análise de $\hat{\theta}_{8,1,p}$ para seleção de $\theta_{8,1,k}$	58
Figura 54 – Análise de $\hat{\theta}_{8,2,p}$ para seleção de $\theta_{8,2,k}$	58
Figura 55 – Análise de $\hat{\theta}_{8,3,p}$ para seleção de $\theta_{8,3,k}$	59
Figura 56 – Análise de $\hat{\theta}_{8,10,p}$ para seleção de $\theta_{8,10,k}$	59
Figura 57 – Análise de $\hat{\theta}_{9,2,p}$ para seleção de $\theta_{9,2,k}$	59
Figura 58 – Análise de $\hat{\theta}_{12,2,p}$ para seleção de $\theta_{12,2,k}$	59
Figura 59 – Análise de $\hat{\theta}_{13,1,p}$ para seleção de $\theta_{13,1,k}$	60
Figura 60 – Análise de $\hat{\theta}_{13,2,p}$ para seleção de $\theta_{13,2,k}$	60
Figura 61 – Análise de $\hat{\theta}_{13,3,p}$ para seleção de $\theta_{13,3,k}$	60
Figura 62 – Análise de $\hat{\theta}_{13,10,p}$ para seleção de $\theta_{13,10,k}$	60
Figura 63 – Análise de $\hat{\theta}_{13,11,p}$ para seleção de $\theta_{13,11,k}$	61
Figura 64 – Análise de $\hat{\theta}_{14,1,p}$ para seleção de $\theta_{14,1,k}$	61
Figura 65 – Análise de $\hat{\theta}_{14,2,p}$ para seleção de $\theta_{14,2,k}$	61
Figura 66 – Análise de $\hat{\theta}_{14,3,p}$ para seleção de $\theta_{14,3,k}$	61
Figura 67 – Análise de $\hat{\theta}_{14,10,p}$ para seleção de $\theta_{14,10,k}$	62
Figura 68 – Dados do CSTR para obtenção do modelo PCA	64
Figura 69 – Resíduos utilizados para obter o modelo PCA	64
Figura 70 – Estatística combinada calculada para os dados utilizados para obter o modelo PCA	65
Figura 71 - Variáveis do CSTR para a simulação contendo 10 regiões de operação	66
Figura 72 - Estatística combinada calculada para a simulação contendo 10 regiões de operação	66
Figura 73 - Variáveis do CSTR para simulação com mudança das VE 2 e 10 e com ocorrência de falha 4	67
Figura 74 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 10 e com ocorrência de falha 4	68

Figura 75 - Variáveis do CSTR para simulação com mudança das VE 2 e 3 e com ocorrência de falha 5.....	68
Figura 76 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 3 e com ocorrência de falha 5	69
Figura 77 - Variáveis do CSTR para simulação com mudança das VE 2 e 3 e com ocorrência de falha 7.....	69
Figura 78 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 3 e com ocorrência de falha 7	70
Figura 79 - Variáveis do CSTR para simulação com mudança das VE 2 e 10 e falha 2	70
Figura 80 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 10 e falha 2.....	71
Figura 81 - Box plots das taxas de FN para as falhas 1, 2, 3 e 4.....	71
Figura 82 - Box plots das taxas de FN para as falhas 5, 6 e 7	72
Figura 83 - Média das taxas de FN.....	72
Figura 84 - Norma dos resíduos em MM e M1	73

LISTA DE TABELAS

Tabela 1 – Variáveis e restrições do simulador CSTR.....	43
Tabela 2 – Falhas do CSTR avaliadas	43
Tabela 3 – Faixas de operação das VE.....	44
Tabela 4 – VE selecionadas para cada VP	46
Tabela 5 – Acurácia dos classificadores.....	63

LISTA DE ABREVIATURAS E SIGLAS

CPLS	<i>Concurrent Partial Least Squares</i>
CSTR	<i>Continuously Stirred Reactor Tank</i>
CVA	<i>Canonical Variate Analysis</i>
FDA	<i>Fisher Discriminant Analysis</i>
FN	Falsos Negativos
FP	Falsos Positivos
ICA	<i>Independent Component Analysis</i>
kNN	k-vizinhos mais próximos (<i>k-nearest-neighbor</i>)
LSSVR	<i>Least-Squares Support Vector Regression</i>
M1	Modelo Único
MM	Múltiplos Modelos
MSPC	<i>Multivariate Statistical Process Control</i>
NE	Norma dos Resíduos
NN	<i>Nearest Neighbor</i>
PCA	<i>Principal Component Analysis</i>
PCR	<i>Principal Component Regression</i>
PLS	<i>Partial Least Squares</i>
RLS	<i>Recursive Least Square</i>
SPC	<i>Statistical Process Control</i>
SPE	<i>Squared Prediction Error</i>
SSID	<i>Steady-State Identification</i>
SSM	<i>Steady-State Model</i>
SVDD	<i>Support Vector Data Description</i>
VE	Variáveis Externas
VP	Variáveis Principais

LISTA DE SÍMBOLOS

\mathbf{X}	Matriz de dados
\mathbf{H}	Matriz das variáveis principais
\mathbf{G}	Matriz das variáveis externas
n	Número de amostras em \mathbf{X}
m	Número de variáveis em \mathbf{X}
mh	Número de variáveis principais
mg	Número de variáveis externas
i	Índice referente às variáveis principais
j	Índice referente às variáveis externas
$\theta_{i,j,k}$	Vetor com possíveis valores de coeficientes que relacionam a variável principal i com a variável externa j
$c_{i,j}$	Elemento do vetor $\theta_{i,j,k}$ que relaciona a variável principal i com a variável externa j
\mathbf{E}	Matriz de resíduos
$\hat{\theta}_{i,j,p}$	Coefficientes estimados via RLS entre a variável principal i e a variável externa j
φ_{CRIT}	Valor crítico do método de avaliação do comportamento estacionário proposto por Cao e Rhinehart (1995)
λ_1, λ_2 e λ_3	Valores de ponderação do método de avaliação do comportamento estacionário proposto por Cao e Rhinehart (1995)
φ_{STAT}	Valores da relação estatística do método de avaliação do comportamento estacionário proposto por Cao e Rhinehart (1995)
λ	Fator de esquecimento
$\omega_{i,j,k}$	Vetor com $L_{i,j}$ coeficientes lineares correspondentes aos $L_{i,j}$ clusters de $\hat{\theta}_{i,j,p}$
NL	Número de amostras acima do limite estatístico sem a ocorrência de falha
NT	Número de amostras total sem a ocorrência de falha
NPF	Número de amostras após a ocorrência da falha
NPD	Número de amostras após a detecção da falha

SUMÁRIO

1	INTRODUÇÃO	15
1.1	JUSTIFICATIVA E MOTIVAÇÃO	15
1.2	ESTADO DA ARTE.....	16
1.3	OBJETIVOS	18
1.4	ORGANIZAÇÃO DO TRABALHO.....	18
2	REVISÃO DA LITERATURA.....	20
2.1	CONTROLE ESTATÍSTICO DE PROCESSOS	20
2.2	CONTROLE ESTATÍSTICO MULTIVARIADO DE PROCESSOS.....	21
2.3	ANÁLISE DE COMPONENTES PRINCIPAIS	21
2.4	ESTATÍSTICAS T^2 DE HOTELLING, SPE E COMBINADA.....	23
2.5	MÚLTIPLAS REGIÕES DE OPERAÇÃO	24
2.6	ANÁLISE EXTERNA	26
2.6.1	Análise Externa Estática	26
2.6.2	Considerações sobre A Análise Externa	27
2.7	ESTIMADOR RECURSIVO DE MÍNIMOS QUADRADOS	27
2.8	SELEÇÃO DE VARIÁVEIS	28
2.9	AVALIAÇÃO DO COMPORTAMENTO ESTACIONÁRIO.....	29
2.10	MÉTODO DE AGRUPAMENTO <i>K-MEANS</i>	31
2.11	CLASSIFICADOR K-VIZINHOS MAIS PRÓXIMOS	32
3	PROPOSTA DE METODOLOGIA	34
3.1	DEFINIÇÕES GERAIS.....	34
3.2	CONSTRUÇÃO DOS MODELOS.....	35
3.2.1	Obtenção e Análise Dos Dados	35
3.2.2	Estimador Recursivo de Mínimos Quadrados	36
3.2.3	Obtenção dos Modelos Lineares.....	37
3.2.4	Treinamento dos Classificadores	39
3.2.5	Obtenção do Modelo PCA.....	40
3.3	MONITORAÇÃO	41
4	APLICAÇÃO E ANÁLISE DOS RESULTADOS	42
4.1	INTRODUÇÃO	42
4.2	TANQUE REATOR CONTINUAMENTE AGITADO	42
4.3	APLICAÇÃO DA CONSTRUÇÃO DOS MODELOS	43

4.3.1	Obtenção e análise dos dados.....	43
4.3.2	Estimador Recursivo de Mínimos Quadrados	45
4.3.3	Obtenção dos Modelos Lineares.....	57
4.3.4	Treinamento dos Classificadores	63
4.3.5	Obtenção do Modelo PCA.....	63
4.4	ANÁLISE DOS RESULTADOS	65
5	CONCLUSÕES E TRABALHOS FUTUROS	74
5.1	CONCLUSÕES.....	74
5.2	TRABALHOS FUTUROS.....	75

1 INTRODUÇÃO

1.1 JUSTIFICATIVA E MOTIVAÇÃO

A monitoração de processos industriais, quando feita de forma eficiente, pode trazer diversas vantagens para a indústria, tais como: mais segurança na produção através da previsão de falhas e correção de suas causas com antecedência; diminuição do tempo de parada da planta, trazendo o aumento da produtividade; e uma maior qualidade de produtos e processos. Dessa forma, as indústrias podem se tornar economicamente mais eficientes, sendo capazes de evitar prejuízos como desperdícios de lotes completos com má qualidade e até danos em equipamentos, garantindo, assim, maior competitividade no mercado.

No entanto, esta monitoração, em geral, não é uma tarefa simples, devido à grande e crescente complexidade dos processos industriais. Por outro lado, segundo Lee (2013), o desenvolvimento de sensores inteligentes e o aumento da capacidade de armazenamento de dados vêm disponibilizando cada vez mais informações desses processos. Dessa forma, várias técnicas baseadas em dados – visando a detecção e diagnóstico de falhas em processos industriais – têm sido desenvolvidas e vêm sendo exploradas em diversos setores já há algum tempo.

Tais técnicas trazem significativas vantagens de implementação, uma vez que não dependem de modelos matemáticos customizados – obtidos através da dedicação de elevado tempo e esforço por parte de engenheiros especialistas – para descrever o funcionamento da planta. Dependem, contudo, principalmente da extração de informações contidas em bancos de dados juntamente com conhecimentos de engenheiros e operadores da planta, tornando sua aplicabilidade bastante abrangente.

Nesse contexto, o controle estatístico multivariado de processos (*Multivariate Statistical Process Control* - MSPC) foi proposto por Jackson e Mudholkar (1979), e métodos baseados em MSPC convencionais vêm atingindo resultados satisfatórios no monitoramento de processos com um único modo de operação, como apresentado em Zhang (2000), Joe Qin (2003) e em Lee, Qin e Lee (2006). No entanto, em muitos processos, as condições de operação não são constantes por diversas causas, como: mudanças nas estratégias de produção, na composição de produtos e nas matérias primas; variações sazonais; entre outras. Dessa forma, a aplicação de técnicas convencionais – que assumem que o processo opera em um único modo de operação – a processos cujas condições de operação não são constantes resulta em altas taxas de falsos alarmes.

Vários estudos recentes propõem soluções para este problema com abordagens que englobam: a delimitação das regiões de operação; e o desenvolvimento de um conjunto de modelos locais de monitoramento, de modo que cada um deles seja válido para seu modo de operação correspondente. Porém, soluções pautadas em tal abordagem tornam-se inviáveis quando as regiões de operação não são bem delimitadas e finitas. Isto

acontece quando os modos de operação dependem de variáveis de referência que podem assumir qualquer valor real dentro de um certo intervalo, – como, por exemplo, um setpoint de temperatura – pois, a partir dessas variáveis, podem ser gerados infinitos modos de operação.

A análise externa proposta por Kano et al. (2004) chama esse tipo de variáveis de variáveis externas (VE), e permite realizar o monitoramento de processos com múltiplos modos de operação sem a necessidade de definir um número limitado de regiões. Porém, quando há a presença de não linearidades no processo, a aplicação da análise externa linear prejudica a qualidade da monitoração. Nesses casos, uma alternativa é utilizar a análise externa não linear, como alguns estudos propõem. No entanto, esta solução requer definições complexas de funções e parâmetros que influenciam consideravelmente na qualidade do modelo final. Uma alternativa à análise externa não linear é a utilização de múltiplos modelos lineares na análise externa.

1.2 ESTADO DA ARTE

Garcia-Alvarez, Fuente e Sainz (2012) propõem um método para detecção de falhas em múltiplos modos de operação. Tal método consiste no cálculo de um modelo para cada modo de operação, através do uso de análise de componentes principais (*Principal Component Analysis* - PCA), e no cálculo de modelos que monitoram os estados de transição entre os modos de operação, utilizando uma adaptação do PCA chamada U-PCA, geralmente empregada em monitoração de processos em batelada. Neste método, quando a ferramenta de supervisão detecta uma mudança no modo de operação atual, a mesma seleciona o modelo PCA adequado para este modo e monitora o estado do processo usando o modelo relacionado ao mesmo.

Yong (2015) propõe um método de monitoração adaptativo para processos não lineares que resolve o problema de múltiplos modos de operação e variância no tempo, utilizando um método de *clusterização* com a técnica *Clone Differential Evolution Harmony Search* (CloneDE- HS) para separar e delimitar os modos de operação existentes.

Vários outros diferentes métodos de *clusterização* são propostos, com a mesma finalidade, em outros estudos. Zullo (1996) identifica clusters no subespaço do PCA. Maestri et al. (2010) utilizam uma técnica de *clusterização* robusta, na qual assume-se que as observações vêm de populações com as matrizes de covariância comuns. Isso também é assumido por Hwang e Han (1999), que propõem um método baseado na *clusterização* hierárquica para as regiões de operação e na obtenção de um modelo PCA mais generalizado, – referenciado como super modelo PCA – obtido com os dados de todos os *clusters*, padronizados através da centralização e normalização, concatenados, formando uma única matriz de dados. Esse método apresenta adaptabilidade a novos modos de operação e um melhor desempenho em termos de resolução e robustez, quando comparado a um modelo PCA tradicional, seja ele desenvolvido para todos os modos de operação ou para um conjunto de modelos locais referentes a cada modo. No entanto, tal

método é válido somente para modos de operação que não apresentam não-linearidades e para processos que apresentam modos de operação que não estão longe uns dos outros.

Contudo, em todas as abordagens citadas, o número de modos de operação é finito e deve ser delimitado e, em suas implementações, é necessário realizar treinamentos segregados para cada região de operação.

Kano *et al.* (2004) propõem a metodologia de análise externa, na qual não há a necessidade de definir um número limitado de regiões de operação. Nesse método, os modos de operação dependem de um conjunto de variáveis – chamadas de variáveis externas (VE) – e, a partir delas, infinitos modos de operação podem ser monitorados, não havendo a necessidade de treinamentos particulares para cada modo de operação. Para isto, usa-se regressão linear para obter tanto as relações entre as variáveis externas quanto os resíduos entre as variáveis monitoradas e suas estimativas.

Vários estudos recentes se baseiam em Kano *et al.* (2004). Zheng, Qin e Chai (2016) propõem um método de monitoração de processos com multimodos de operação baseado em análise externa com o uso de projeção concorrente para estruturas latentes (*Concurrent PLS -CPLS*). Ge, Song e Kano (2014) propõem um sensor virtual adaptativo utilizando análise externa e obtendo modelos lineares via regressão de componentes principais (*Principal Component Regression – PCR*). Nesse último método, é introduzida uma abordagem utilizando uma janela deslizante para garantir o funcionamento *on-line* do método e é desenvolvida uma estratégia de monitoração do processo utilizando *Support Vector Data Description (SVDD)*. No entanto, as metodologias apresentadas nesses trabalhos focam somente em relações lineares entre os conjuntos das variáveis.

A linearidade entre as variáveis externas (VE) e as variáveis principais (VP) é um fator que deve ser considerado. Se a análise externa linear for aplicada em situações em que as relações das variáveis são não-lineares, a monitoração será prejudicada. Ge *et al.* (2008) afirmam que a maioria dos processos industriais modernos são não lineares. Nesse caso, Kano *et al.* (2004) afirmam que se pode utilizar a análise externa não linear, que se diferencia da análise externa linear apenas pela estrutura do modelo.

Entretanto, a aplicação de modelos não lineares pode requerer definições complexas de funções e parâmetros que influenciam consideravelmente na qualidade do modelo final. Por exemplo, um modelo de rede neural pode ser usado para descrever a relação não linear entre VE e VP. Isso é abordado com detalhes em (Patan, 2008). No entanto, a rede neural sofre de inconveniências, como por exemplo, a estratégia complexa de treinamento, – como é discutido em Ge *et al.* (2008) – fazendo-se necessário, geralmente, a definição de uma função de ativação adequada e a inicialização de parâmetros. Outro exemplo é um trabalho também apresentado por Ge *et al.* (2008), que extrai as relações não lineares entre VE e VP através da aplicação da regressão de vetor de suporte por mínimos quadrados (*Least-Squares Support Vector Regression - LSSVR*), e apresenta ainda um esquema de monitoramento *on-line*. Porém, para alcançar um alto nível de desempenho com modelos LSSVR, a função *kernel* não linear deve ser escolhida adequadamente, além da necessidade de sintonização de outros parâmetros, incluindo o parâmetro de

regularização e o parâmetro *kernel*, que influenciam consideravelmente a qualidade do modelo e afetam o desempenho da monitoração *on-line*.

1.3 OBJETIVOS

Este trabalho tem como objetivo principal investigar a possibilidade da utilização de múltiplos modelos lineares aplicados à análise externa empregada na detecção de falhas em processos industriais operando em múltiplas regiões de operação, com intuito de tratar a presença de não linearidades no processo.

Para tal, propõe-se uma metodologia cujos objetivos específicos são:

- Detectar a necessidade do uso de múltiplos modelos lineares na aplicação da metodologia de análise externa utilizada na detecção de falhas;
- Propor um método para construção dos modelos;
- Propor um método de aplicação dos múltiplos modelos na monitoração do processo;
- Avaliar o desempenho da metodologia proposta através de sua aplicação na supervisão do processo industrial simulado de um tanque reator continuamente agitado (CSTR), comparando seus resultados com os resultados obtidos através da metodologia de análise externa linear tradicional que utiliza um único modelo linear.

1.4 ORGANIZAÇÃO DO TRABALHO

Este documento encontra-se dividido em 5 (cinco) capítulos organizados de tal forma que os assuntos são apresentados com evolução gradativa, de forma que, para compreensão do trabalho como um todo, é imprescindível a leitura sequencial de todo o documento. Entretanto, cada capítulo é consistente no escopo que propõe, de forma a proporcionar compreensão do assunto abordado independente da leitura completa do trabalho.

Dessa forma, após este capítulo introdutório, tem-se o Capítulo 2 que consiste em uma revisão bibliográfica e teórica sobre as técnicas da literatura utilizadas neste trabalho. São abordados os principais métodos utilizados para controle de processos industriais, além de métodos matemáticos, estatísticos e de análise de dados que embasaram o desenvolvimento da metodologia proposta nesta dissertação.

O Capítulo 3 apresenta a metodologia elaborada para a detecção de falhas. Primeiramente, são definidas as nomenclaturas utilizadas na estruturação da metodologia e as etapas da abordagem metodológica proposta. Subsequentemente, é descrito cada uma das etapas que compõem a metodologia.

O Capítulo 4 apresenta a metodologia desenvolvida aplicada na supervisão do processo industrial simulado de um tanque reator continuamente agitado (CSTR) operando em diferentes modos de operação e com a ocorrência de diferentes falhas. A capacidade de detecção de falhas do modelo é avaliada para falhas em diferentes modos de operação, e

os resultados obtidos com a aplicação do método são comparados aos obtidos através da metodologia de análise externa linear tradicional.

Por fim, o Capítulo 5 traz as conclusões sobre a metodologia desenvolvida, apontando as possibilidades de aprimoramentos e perspectivas para trabalhos futuros.

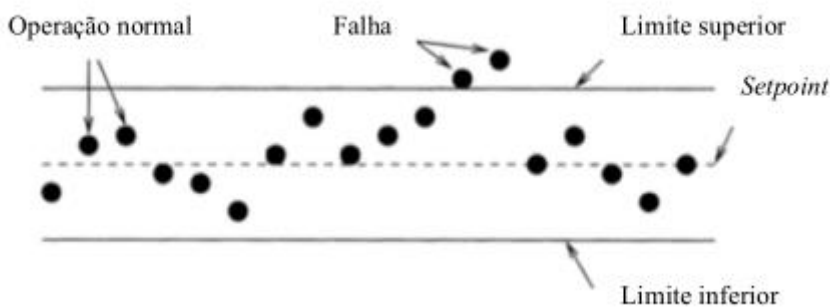
2 REVISÃO DA LITERATURA

2.1 CONTROLE ESTATÍSTICO DE PROCESSOS

Segundo Kruger e Xie (2012), métodos de controle estatístico de processos (*Statistical Process Control – SPC*) foram introduzidos no início do século XX por Walter A. Shewhart. Esses métodos fornecem meios objetivos de controle de qualidade em qualquer processo de transformação, seja na fabricação de artefatos, na prestação de serviços ou na transferência de informações. Muitas de suas ferramentas são amplamente conhecidas na literatura e têm grande aplicabilidade na indústria, sendo elas: histograma; diagrama de Pareto; diagrama Causa-efeito; diagrama de concentração de defeito; cartas de controle; diagrama de pontos; entre outras. No entanto, Oakland (2007) afirma que o SPC não é apenas um kit de ferramentas, mas sim uma estratégia para reduzir a variabilidade – causa da maioria dos problemas de qualidade – na variação nos produtos, nos tempos de entregas, nas formas de fazer coisas, nos materiais, nas atitudes das pessoas, nos equipamentos e seu uso, nas práticas de manutenção, enfim, em tudo.

Uma abordagem típica de SPC que ainda é amplamente utilizada na indústria para detecção de falhas é o gráfico de Shewhart, no qual os limiares de cada variável de observação (variável de processo observada através de uma leitura de sensor) são obtidos via análise estatística univariada. Tais limiares definem o limite para operações controladas e uma violação desses limites com dados *on-line* indicaria falha (CHIANG, RUSSELL e BRAATZ, 2001). Uma ilustração do gráfico de Shewhart pode ser observada na Figura 1. Os pontos pretos são observações de uma variável.

Figura 1 – Gráfico de Shewhart aplicado à análise estatística univariável



Fonte: Chiang, Russell e Braatz, 2001.

O erro Tipo-I se refere a falsos alarmes, que é uma indicação de uma falha quando, na realidade, uma falha não ocorreu. O erro Tipo-II se refere a uma detecção perdida, que é quando não há indicação de falha, embora uma falha tenha ocorrido. Para detecção de falhas, existe um compromisso inerente entre minimizar erros do Tipo-I e do Tipo-II, que são inversamente relacionados (CHIANG, RUSSELL e BRAATZ, 2001).

A análise estatística univariada determina os limiares para cada variável de observação sem usar qualquer informação das outras variáveis. Essa análise ignora as correlações entre as variáveis, que são muito comuns nas plantas por geralmente apresentarem sistemas realimentados, interdependência das malhas de controle e encadeamento dos processos. Portanto, apesar do sucesso generalizado da metodologia SPC, é importante notar que as correlações entre as variáveis do processo podem aumentar substancialmente o número de erros do tipo II (KRUGER e XIE, 2012). Neste contexto, surge o controle estatístico multivariado de processos.

2.2 CONTROLE ESTATÍSTICO MULTIVARIADO DE PROCESSOS

Segundo Satoyama et al. (2016), desde 1979 o controle estatístico multivariado de processos (*Multivariate Statistical Process Control* – MSPC) vem sendo aplicado com sucesso a vários processos industriais como uma ferramenta dominante de monitoramento estatístico de processos. O MSPC utiliza os dados históricos para extrair correlações das variáveis, reduzir a dimensionalidade espacial e derivar regiões operacionais normais.

Dentre as técnicas de MSPC encontradas na literatura, as mais notáveis incluem:

- Análise de componentes principais (*Principal component analysis* - PCA);
- Análise de discriminante linear ou de Fisher (*Fisher Discriminant Analysis* - FDA);
- Análise de componentes independentes (*Independent Component Analysis* - ICA);
- Análise da correlação canônica (*Canonical Variate Analysis* - CVA);
- Mínimos quadrados parciais (*Partial Least Squares* - PLS);

sendo PCA e PLS os principais métodos utilizados para monitoração.

As falhas do processo podem, então, ser detectadas e diagnosticadas usando as métricas estatísticas induzidas, como a estatística T^2 de Hotelling e a estatística residual Q (KRUGER e XIE, 2012).

2.3 ANÁLISE DE COMPONENTES PRINCIPAIS

A técnica análise de componentes principais (*Principal Component Analysis* – PCA) é um dos pilares na análise de dados e consiste de álgebra linear aplicada, sendo um método simples e não-paramétrico de extrair informações relevantes a partir de um conjunto dados. Além disso, tal técnica é capaz de reduzir um complexo conjunto de dados para uma dimensão inferior a fim de revelar a dinâmica ou característica presente nos dados, por vezes escondidas (SHLENS, 2014).

Considerando a matriz de dados normalizados $\mathbf{X} \in \mathfrak{R}^{k \times m}$, em que k é o número de amostras e m é o número de variáveis, as componentes principais – também chamadas de *loading vectors* e representadas pela matriz \mathbf{P} – são obtidas selecionando os l autovetores

correspondentes aos l maiores autovalores (λ) da matriz de covariância de \mathbf{X} (WOLD, ESBENSEN e GELADI, 1987; MA, HU e SHI, 2012; JOE e QIN, 2003).

A projeção de \mathbf{X} no espaço de menor dimensão é contida na matriz de *scores* \mathbf{T} , obtida através de

$$\mathbf{T} = \mathbf{X}\mathbf{P} \quad (2.1)$$

e a projeção de \mathbf{T} de volta no espaço de observações inicial com m dimensões é obtido por

$$\hat{\mathbf{X}} = \mathbf{T}\mathbf{P}^T \quad (2.2)$$

A diferença entre \mathbf{X} e $\hat{\mathbf{X}}$ é a matriz de resíduos dada por

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}} \quad (2.3)$$

A matriz de resíduos captura as variações do espaço das observações das componentes principais relativas aos $(m - l)$ menores autovalores (CHIANG, RUSSELL e BRAATZ, 2001).

A normalização das variáveis é um passo importante neste método, uma vez que o mesmo é sensível à ordem de grandeza das variáveis utilizadas. Caso a normalização não seja realizada, há a possibilidade de algumas variáveis sem importância dominarem a variância na matriz \mathbf{X} (dependendo de sua ordem de grandeza), escondendo a variância de variáveis de menor grandeza. Dessa forma, é comum que os dados sejam padronizados a fim de assumirem média zero e variância unitária (ZHANG, 2000).

O número de componentes principais l pode ser selecionado via diferentes abordagens. Dois exemplos são a percentagem de variância cumulativa e a validação cruzada, descritas a seguir conforme Zhang (2000):

1. Percentagem de variância cumulativa: essa abordagem mede a variância percentual capturada pelas primeiras a componentes principais, medida de acordo com a Equação (2.4):

$$CPV(l) = \frac{\sum_{j=1}^a \lambda_j}{\sum_{j=1}^m \lambda_j} 100\% \quad (2.4)$$

l componentes principais são escolhidas se a $CPV(l)$ explicar uma variância predeterminada dos dados (η). Ou seja, l é escolhido tal que a $CPV(l)$ seja no mínimo igual a η .

2. Validação cruzada: nessa abordagem, um conjunto de dados é dividido em vários blocos. Cada vez que um bloco de dados é deixado de fora, uma análise de PCA é realizada no restante dos dados e uma predição da soma do erro quadrático é calculada baseada no bloco de dados que foi deixado de fora. O procedimento é repetido até que

cada bloco de dados tenha sido deixado de fora uma vez. Somando todos os erros quadráticos resultantes, é obtido o erro quadrático acumulado. A ordem ideal do PCA por essa abordagem é a ordem que minimiza este erro acumulado. Embora este método seja longo e não tenha base estatística, é efetivo em uso prático.

2.4 ESTATÍSTICAS T^2 DE HOTELLING, SPE E COMBINADA

A estatística T^2 de Hotelling é uma medida de variação das componentes principais e tem seu limiar definido por uma distribuição F, pois considera que os dados seguem uma distribuição normal multivariada (JOE QIN, 2003).

A estatística residual Q , ou SPE (*Squared Prediction Error*), é uma medida que se relaciona com os resíduos do modelo e mede quão bem a representação dimensional reduzida de dados descreve os mesmos. Considerando um cenário extremo em que duas variáveis são perfeitamente correlacionadas, o vetor residual seria nulo, já que os valores de ambas as variáveis podem ser recuperadas sem erros através da projeção da amostra associada. No caso de alto grau de correlação, os valores projetados de cada amostra estariam próximos das amostras de observação. As variáveis residuais, suas variâncias e o limite de controle da estatística de monitoramento residual descrevem estatisticamente essa proximidade (KRUGER e XIE, 2012).

Para monitorar processos, essas estatísticas são empregadas de acordo com

$$T^2 = x^T P \Lambda^{-1} P^T x \quad (2.5)$$

$$SPE = r^T r, \quad r = (I - PP^T)x \quad (2.6)$$

em que Λ^{-1} é a inversa da matriz diagonal dos l maiores autovalores. Os seus limites estatísticos são dados por

$$T_\alpha^2 = \frac{l(k-1)}{k-l} F_\alpha(l, k-l) \quad (2.7)$$

$$SPE_\alpha \sim g\chi_h^2 \quad (2.8)$$

em que $F_\alpha(l, k-l)$ é a distribuição F com nível de significância α , com l e $(k-l)$ graus de liberdade. O limite de controle do SPE é obtido de acordo com uma distribuição χ^2 para a distribuição gerada pelos dados normais, e g e h são calculadas por $g = b/2a$ e $h = 2a^2/b$, em que a e b são a média e a variância estimada do SPE, respectivamente.

A ocorrência de uma falha pode influenciar separadamente tanto a estatística SPE como a T^2 , ou ambas simultaneamente. Essas estatísticas se comportam de forma complementar. Nesse contexto, a Estatística Combinada foi proposta por Yue e Qin (2001) como índice combinado para a detecção de falhas, combinando SPE e T^2 de acordo com a Equação (2.9).

$$\varphi = \frac{SPE(\mathbf{x})}{\delta_\alpha^2} + \frac{T^2(\mathbf{x})}{\chi_{l;\alpha}^2} = \mathbf{x}^T \boldsymbol{\phi} \mathbf{x} \quad (2.9)$$

sendo

$$\boldsymbol{\phi} = \frac{\mathbf{P}\boldsymbol{\Lambda}^{-1}\mathbf{P}^T}{\chi_{l;\alpha}^2} + \frac{\mathbf{I} - \mathbf{P}\mathbf{P}^T}{\delta_\alpha^2} \quad (2.10)$$

O limite de confiança é estabelecido usando a distribuição de formas quadráticas, dessa forma $\boldsymbol{\phi}$ segue aproximadamente

$$\varphi = \mathbf{x}^T \boldsymbol{\phi} \mathbf{x} \sim g\chi_h^2 \quad (2.11)$$

com o coeficiente

$$g = \frac{tr(\mathbf{S}\boldsymbol{\phi})^2}{tr(\mathbf{S}\boldsymbol{\phi})} \quad (2.12)$$

O número de graus de liberdade da distribuição χ^2 é

$$h = \frac{[tr(\mathbf{S}\boldsymbol{\phi})]^2}{tr(\mathbf{S}\boldsymbol{\phi})^2} \quad (2.13)$$

sendo

$$tr(\mathbf{S}\boldsymbol{\phi}) = \frac{l}{\chi_{l;\alpha}^2} + \frac{\sum_{i=l+1}^m \lambda_i}{\delta_\alpha^2} \quad (2.14)$$

e

$$tr(\mathbf{S}\boldsymbol{\phi})^2 = \frac{l}{\chi_{l;\alpha}^4} + \frac{\sum_{i=l+1}^m \lambda_i^2}{\delta_\alpha^4} \quad (2.15)$$

Então, pode-se obter o limite estatístico de φ com o nível de significância α , e assim uma falha é detectada se

$$\varphi > g\chi_{h;\alpha}^2 \quad (2.16)$$

2.5 MÚLTIPLAS REGIÕES DE OPERAÇÃO

Métodos baseados em MSPC convencionais – como PCA e PLS – vêm tendo resultados satisfatórios no monitoramento de processos com um único modo de operação, como apresentado em Westerhuis, Gurden e Smilde (2000) e em Joe Qin (2003). Tais métodos assumem que o processo opera em um estado estacionário específico, ou seja, em um único modo de operação. Porém, em muitos processos, as condições de operação não são constantes por diversas causas, como variações sazonais, mudanças nas estratégias de produção, na composição de produtos, nas matérias primas, entre outras. Além disso, os

setpoints das malhas de controle são frequentemente alterados por otimizadores de processos, causando mudanças de estado estacionário (KANO *et al.*, 2004; GE e SONG, 2008).

Situações desse tipo podem produzir mudanças na estrutura da covariância usada para construir os modelos baseados nos dados. Quando essas técnicas convencionais são empregadas, tais mudanças são detectadas como falhas, gerando altas taxas de falsos alarmes (ZHAO, ZHANG e XU, 2004; GARCIA-ALVAREZ, FUENTE e SAINZ, 2012). Portanto, são necessários métodos de monitoração que consigam diferenciar mudanças normais nas condições de operação de mudanças provocadas por falhas.

Segundo Hwang e Han (1999), três tipos de possíveis soluções estão disponíveis na literatura para este problema, sendo elas:

- (a) Desenvolver um conjunto de modelos locais de monitoramento, de modo que cada um deles seja válido para seu modo de operação correspondente;
- (b) Atualizar de forma adaptável o modelo para refletir as mudanças dos modos de operação;
- (c) Desenvolver um único modelo de monitoramento PCA tradicional para todos os diferentes modos de operação.

A opção (c) torna o modelo muito conservador, levando a uma menor sensibilidade no monitoramento e um monitoramento errôneo, ou seja, ocorrência de erros Tipo II. A opção (a) tem uma maior sensibilidade no monitoramento, no entanto, o modelo local pode perder uma parte da estrutura de covariância verdadeira durante o desenvolvimento do modelo, devido a dados operacionais incompletos que não são suficientes para representar as características do processo.

Algumas abordagens da literatura apresentam a característica da opção (a), entre elas estão os trabalhos de Yong, Xin e Zhenlei (2015), Zullo (1996) e Maestri, Farall, *et al.* (2010), Hwang e Han (1999), Garcia-Alvarez, Fuente e Sainz (2012), entre outros. Porém, estas abordagens ficam inviáveis quando as regiões de operação não são bem delimitadas e finitas. Isto acontece quando os modos de operação dependem de variáveis de referência que podem assumir qualquer valor real dentro de um certo intervalo – como um *setpoint* de temperatura ou a vazão de alimentação – pois, a partir dessas variáveis, podem ser gerados infinitos modos de operação.

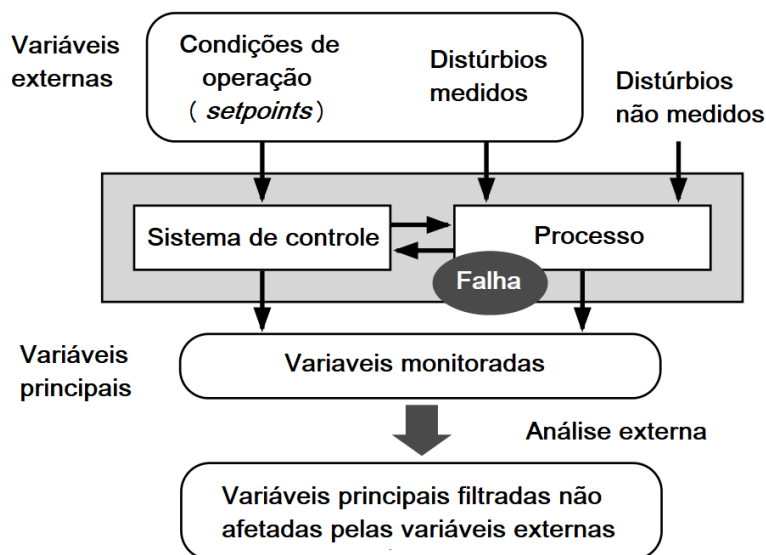
No entanto, a análise externa proposta em Kano *et al.* (2004) não se enquadra em nenhuma das três opções de abordagens listadas por Hwang e Han (1999). Nesse método, a influência das variáveis relacionadas aos modos de operação é retirada das variáveis principais do processo. Portanto, um único modelo de monitoramento pode ser construído com uma boa sensibilidade, sem a necessidade de definir um número limitado de regiões. Tal metodologia será abordada em detalhes na Seção 2.6.

2.6 ANÁLISE EXTERNA

No método da análise externa, as variáveis responsáveis por alterar a região de operação – como vazão de alimentação e *setpoints* – são classificadas como variáveis externas (VE). As variáveis que são afetadas tanto pelas VE como também por outros distúrbios no processo são denominadas variáveis principais (VP). Dessa forma, mudanças nas VE geram nas VP efeitos que não devem ser interpretados como falhas.

Para diferenciar esses efeitos e as falhas, a matriz de VP é decomposta em duas partes: uma parte que é explicada pela influência das VE e outra parte que não. Isto pode ser feito através da análise de regressões lineares múltiplas, – como é realizado em Kano *et al.* (2004) – utilizando as VE como entradas e as VP como saídas. Uma matriz de resíduos é então gerada, representando a parte das VP que não sofre influência das VE. Logo, é possível utilizar qualquer método de SPC na matriz de resíduos para realizar a monitoração do processo. O diagrama da Figura 2 ilustra o método da análise externa.

Figura 2 – Metodologia da Análise Externa



Fonte: Kano, 2004.

2.6.1 ANÁLISE EXTERNA ESTÁTICA

Seja a matriz de dados $X \in \mathfrak{R}^{k \times m}$ descrita por

$$X = [H \ G] \quad (2.17)$$

em que k é o número de amostras, m é o número de variáveis, H é a matriz das VP e G é a matriz das VE. Seja $H \in \mathfrak{R}^{k \times mh}$ e $G \in \mathfrak{R}^{k \times mg}$, sendo mh o número de VP e $mg = (m - mh)$ o número de VE. A matriz de coeficientes C é, então, obtida por mínimos quadrados conforme Equação (2.18).

$$\mathbf{C} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{H} \quad (2.18)$$

A matriz de resíduos, que é a parte decomposta das VP que não depende das VE, é obtida através da Equação (2.19), podendo ser utilizada para monitoração de falhas.

$$\mathbf{E} = \mathbf{H} - \mathbf{GC} \quad (2.19)$$

2.6.2 CONSIDERAÇÕES SOBRE A ANÁLISE EXTERNA

No método de análise externa proposto por Kano *et al.* (2004), as VE podem gerar infinitos modos de operação, logo, não há a necessidade nem de definir um número específico de regiões de operação, nem de realizar treinamento para cada região definida, podendo ser aplicado desde que as VE sejam conhecidas.

A definição das VE do processo pode, por vezes, ter maior complexidade, dificultando a aplicação da análise externa. Embora em muitos casos elas possam ser facilmente identificadas, – através de mudanças de *setpoints*, mudanças de variáveis de entrada do processo, dentre outros fatores – em sistemas complexos, as VE podem ser difíceis de serem localizadas. Se alteradas de forma irracional, há casos em que algumas VE podem não apenas causar mudanças no processo, como também resultar em falhas. Além disso, em muitos processos, somente quando várias VE sofrem alteração, que uma mudança de modo de operação é considerada. No caso de apenas uma ou duas delas mudarem, este evento pode ser considerado como distúrbio ou falha do processo. Portanto, a escolha das VE é de grande relevância para a análise externa. Quando escolhidas adequadamente, há aprimoramento na eficiência desse método e melhoria de desempenho do monitoramento. No entanto, no que se sabe, a definição das VE é principalmente implementada pela experiência e pelo conhecimento do processo, e escolhê-las automaticamente é um desafio na área (GE *et al.*, 2008).

2.7 ESTIMADOR RECURSIVO DE MÍNIMOS QUADRADOS

Obter modelos matemáticos a partir de dados observados é um processo fundamental na ciência e na engenharia. Na área de controle, esse processo foi denominado identificação de sistemas, e seu objetivo é encontrar modelos dinâmicos a partir de sinais de entrada e saída observados (VAHIDI, STEFANOPOULOU e PENG, 2005). A identificação de sistemas tem sido uma área de pesquisa ativa, e muitas das técnicas tornaram-se ferramentas padrão de engenharia de controle e processamento de sinais (LJUNG, 1998).

No estimador de mínimos quadrados, os parâmetros desconhecidos de um modelo linear são obtidos de tal modo que a soma dos quadrados da diferença entre os valores observados e os valores calculados seja mínimo. A maior parte de suas aplicações busca estimar parâmetros em tempo real e, para isso, pode-se utilizar o estimador recursivo de mínimos quadrados (*Recursive Least Square* – RLS), cuja estrutura é similar a maioria dos algoritmos de estimação recursiva. Uma descrição mais detalhada desse método pode

ser encontrada em livros sobre identificação de sistemas – como em Ljung (1987) – onde o mesmo pode ser resumido por

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \psi^T(t)\hat{\theta}(t-1)] \quad (2.20)$$

$$L(t) = \frac{P(t-1)\psi(t)}{\lambda(t) + \psi^T(t)P(t-1)\psi(t)} \quad (2.21)$$

$$P(t) = \frac{1}{\lambda(t)} \left[P(t-1) - \frac{P(t-1)\psi(t)\psi^T(t)P(t-1)}{\lambda(t) + \psi^T(t)P(t-1)\psi(t)} \right] \quad (2.22)$$

sendo $\hat{\theta}(t)$ o vetor de parâmetros estimados no momento t , $L(t)$ um pré-filtro para os erros de predição, $y(t)$ a variável de saída no momento t , $\psi(t)$ o vetor de regressão no momento t , $P(t)$ a matriz de covariância de θ e $\lambda(t)$ o fator de esquecimento.

A Equação (2.20) atualiza as estimativas de $\hat{\theta}$ em cada etapa com base no erro medido. O fator de esquecimento λ , por sua vez, é um escalar entre 0 e 1. Definir $\lambda = 1$ corresponde a não descartar dados para a estimação, enquanto definir $\lambda < 1$ implica em gradualmente descartar medidas passadas a favor de informações mais recentes.

Na metodologia proposta neste trabalho, o RLS é aplicado com o intuito de estimar coeficientes dos modelos entre variáveis externas e principais, em diferentes regiões de operação.

2.8 SELEÇÃO DE VARIÁVEIS

A análise externa aplica a regressão linear múltipla, utilizando todas as VE como entradas nos modelos que preveem todas as VP. Porém, na maioria dos casos, nem todas as VE influenciam todas as VP, sendo assim, os modelos resultantes contêm preditores (VE) desnecessários na sua estrutura. Dessa forma, segundo Guyon e Elisseeff (2003), realizar uma seleção de variáveis, ou seja, escolher os melhores preditores para o modelo, pode gerar as seguintes vantagens:

- Melhora o desempenho de predição dos preditores;
- Proporciona preditores mais rápidos e mais eficazes economicamente;
- Proporciona um melhor entendimento do processo que gera os dados;
- Facilita a visualização e a compreensão de dados;
- Reduz requisitos de medição e armazenamento;
- Redução dos tempos de treinamento e utilização.

Allen (1974) afirma que a adição de uma variável em uma equação estimada via mínimos quadrados quase sempre aumenta (no entanto, nunca diminui) a variância da resposta estimada. Portanto, utilizar somente os preditores necessários resulta em modelos menores, nos quais graus de liberdade não são desperdiçados e ruídos não são adicionados

desnecessariamente na estimação via tais preditores. Além disso, uma menor quantidade de variáveis requer um menor tempo de processamento para o cálculo dos modelos e para sua utilização, ou seja, o custo computacional é menor.

Para seleção de variáveis, uma opção é avaliar o grau de relacionamento entre as mesmas. É comum a utilização do coeficiente de correlação de Pearson como critério de avaliação desta relação, porém, para utilizá-lo, considera-se que as variáveis seguem uma distribuição normal. Nos casos em que as mesmas não seguem uma distribuição teórica conhecida, outras alternativas de coeficientes deverão ser aplicadas, como, por exemplo, o coeficiente de Spearman (BONETT e WRIGHT, 2000).

O coeficiente de correlação de Spearman (r_s) é um tipo de medida estatística não-paramétrica calculado por

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n} \quad (2.23)$$

Com isso, pode-se utilizar a definição de um limiar mínimo para um coeficiente de correlação como forma de seleção de variáveis. No entanto, há na literatura outras soluções mais elaboradas que podem ser utilizadas para tal seleção, como, por exemplo, o teste de significância para o coeficiente de correlação de Spearman, ou métodos de seleção *forward*, de eliminação *backward*, ou de seleção *stepwise*, como descreve Derksen e Keselman (1992). É possível, ainda, considerar soluções mais complexas que visam também eliminar preditores com informações redundantes, como algumas soluções descritas em Guyon e Elisseeff (2003).

2.9 AVALIAÇÃO DO COMPORTAMENTO ESTACIONÁRIO

Segundo Mejía *et al.* (2008), a identificação do estado estacionário (*Steady-State Identification* – SSID) é uma importante tarefa para o controle satisfatório de muitos processos. Os modelos de estado estacionário (*Steady-State Model* – SSM) são amplamente usados nas funções de controle do processo, na identificação de modelos, na detecção de falhas, na otimização e avaliação do desempenho do processo, na análise do estado dos sensores e no controle de dados.

Mejía *et al.* (2008) cita o exemplo de processos químicos e de manufatura que são inerentemente não-estacionários, e, nesses casos, determinados parâmetros de seus modelos têm que ser frequentemente ajustados para que os mesmos sejam mantidos adequados. Entretanto, o ajuste de parâmetros em modelos de estado estacionário deve ser realizado somente utilizando as mudanças ocorridas em dados próximos ao estado estacionário, caso contrário, as mudanças registradas ao longo do processo conduzem a erro de modelagem. Para evitar que a dinâmica do regime transitório do sistema afete a construção dos modelos, optou-se por somente utilizar dados do processo em regime estacionário. Para isso, pode-se compor os dados escolhendo-os manualmente. No entanto, o presente trabalho utiliza um método de avaliação do comportamento estacionário.

O método proposto por Cao e Rhinehart (1995) consiste no estudo comparativo entre o valor da relação estatística das variâncias de ruído estimadas de uma série temporal (φ_{STAT}) e um predeterminado valor crítico de projeto (φ_{CRIT}). Quando a relação φ_{STAT} excede seu valor crítico, a condição de estado não-estacionário é identificada. O método foi projetado para a identificação *on-line*, é capaz de adaptar-se a mudanças no nível de ruído e é independente da escala de valores.

Para a implementação prática desse método, deve-se, primeiramente, inicializar os ajustes dos parâmetros: φ_{CRIT} , λ_1 , λ_2 e λ_3 . A escolha dos valores λ_i e φ_{CRIT} deve equilibrar o compromisso da redução dos erros de Tipo-I e os de Tipo-II. O uso dos valores $\lambda_1 = 0.2$ e $\lambda_2 = \lambda_3 = 0.1$ segundo Cao e Rhinehart (1995; 1997) proporciona um adequado balanço dos compromissos. Em seguida, deve-se calcular as variâncias de ruído estimadas de acordo com

$$y_{eF}(k) = \lambda_1 y_e(k) + (1 - \lambda_1) y_{eF}(k - 1) \quad (2.24)$$

$$\sigma_{F1}^2(k) = \lambda_2 (y_e(k) - y_{eF}(k - 1))^2 + (1 - \lambda_2) \sigma_{F1}^2(k - 1) \quad (2.25)$$

$$s_1^2(k) = \frac{2 - \lambda_1}{2} \sigma_{F1}^2(k) \quad (2.26)$$

$$\sigma_{F2}^2(k) = \lambda_3 (y_e(k) - y_e(k - 1))^2 + (1 - \lambda_3) \sigma_{F2}^2(k - 1) \quad (2.27)$$

$$s_2^2(k) = \frac{\sigma_{F2}^2(k)}{2} \quad (2.28)$$

A relação estatística φ_{STAT} é definida como sendo a relação adimensional entre as estimativas de variâncias de ruído, e pode ser calculada por

$$\varphi_{STAT} = \frac{s_1^2(k)}{s_2^2(k)} = \frac{(2 - \lambda_1) \sigma_{F1}^2(k)}{\sigma_{F2}^2(k)} \quad (2.29)$$

A identificação do estado estacionário é realizada de acordo com

$$\varphi_{STAT} = \begin{cases} > \varphi_{CRIT}(\alpha) \rightarrow SS = 0, & \text{em estado não-estacionário} \\ \leq \varphi_{CRIT}(\alpha) \rightarrow SS = 1, & \text{em estado estacionário} \end{cases} \quad (2.30)$$

Se o valor da relação φ_{STAT} for maior que o valor $\varphi_{CRIT}(\alpha)$, então existe uma porcentagem de confiabilidade de $100(1 - \alpha)$ do processo não se encontrar em estado estacionário. Consequentemente, um valor de φ_{STAT} menor ou igual ao valor crítico $\varphi_{CRIT}(\alpha)$ indica que o processo pode estar em estado estacionário.

Segundo Mejía *et al.* (2008), quando comparado com técnicas convencionais, esse método se destaca pela sua simplicidade de implementação, baixo esforço computacional, baixa armazenagem de dados e facilidade de ajuste dos parâmetros.

2.10 MÉTODO DE AGRUPAMENTO *K-MEANS*

A análise de *clusters* é o estudo formal de métodos e algoritmos para agrupar – ou *clusterizar* – objetos de acordo com características intrínsecas ou similaridades medidas ou percebidas. Algoritmos de *clusterização* são ferramentas úteis para mineração e compactação de dados, estimativa de densidade de probabilidade e muitas outras tarefas importantes. A análise de *clusters* não usa etiquetas de categoria (*category labels*) que rotulam objetos com identificadores anteriores, ou seja, rótulos de classe (*class labels*). A ausência desses rótulos distingue métodos de *clusterização* (aprendizagem não-supervisionada) de métodos de classificação (aprendizagem supervisionada).

Nesse contexto, o método *k-means* é um dos algoritmos de agrupamento mais populares e simples e, apesar de ter sido proposto há mais de 50 anos e centenas de algoritmos de agrupamento terem sido publicados desde então, esse método ainda é amplamente utilizado. Os principais motivos de sua popularidade são a facilidade de implementação, a simplicidade, a eficiência e seu sucesso empírico (HAMERLY e ELKAN, 2004; JAIN, 2010). Tal método visa minimizar a soma de distâncias quadradas entre todos os pontos e o centro de cada *cluster* (RAY e TURI, 1999). Esta soma é descrita pela Equação (2.31).

$$J(C) = \sum_{i=1}^K \sum_{x \in C_i} \|x - z_i\|^2 \quad (2.31)$$

sendo x as amostras, K o número de clusters e C_i o conjunto de amostras cujo centro de *cluster* é z_i .

Uma vez que a soma $J(C)$ sempre diminui com o aumento do número de *clusters* K , – sendo nula quando K é igual ao número de amostras – ela pode ser minimizada apenas para um número fixo de *clusters* (JAIN, 2010). O método *k-means* é descrito com detalhes em Tou e Gonzalez (1974) e consiste nas seguintes etapas:

1. Escolha de K centros de *clusters* iniciais $z_1(1), z_2(1), \dots, z_K(1)$
2. Na etapa t -iterativa, distribua as amostras $\{x\}$ entre os *clusters* K usando a relação da Equação (2.32), para todos $i = 1, 2, \dots, K$ com $i \neq j$; $C_j(t)$ denota o conjunto de amostras cujo centro de cluster é $z_j(t)$.

$$x \in C_j(t) \text{ se } \|x - z_j(t)\| < \|x - z_i(t)\| \quad (2.32)$$

3. Compute os novos centros dos *clusters* $z_j(t + 1)$, $j = 1, 2, \dots, K$, de modo que a soma $J(C)$ seja minimizada de acordo com os novos centros de cada *cluster*. A média das amostras de $C_j(t)$ é a medida que minimiza tal soma. Portanto, o novo centro de cada *cluster* é dado por

$$z_j(t + 1) = \frac{1}{N_j} \sum_{x \in C_j(t)} x, j = 1, 2, \dots, K \quad (2.33)$$

sendo N_j o número de amostras em $C_j(t)$

4. O algoritmo converge quando $z_j(t + 1) = z_j(t)$ para $j = 1, 2, \dots, K$, e o procedimento pode ser encerrado. Caso contrário, deve-se voltar a segunda etapa.

Neste trabalho, o método de agrupamento *k-means* será aplicado aos coeficientes (obtidos via RLS) que serão utilizados para analisar a necessidade do uso de múltiplos modelos lineares.

2.11 CLASSIFICADOR K-VIZINHOS MAIS PRÓXIMOS

As técnicas de classificação podem ser usadas para prever rótulos de classes categóricas e classificar os dados com base em conjuntos de treinamento e em rótulos de classe. O termo classificação pode abranger qualquer contexto em que alguma decisão ou previsão seja feita com base em informações atuais disponíveis. Seu procedimento é um método reconhecido para fazer tais decisões repetidamente em novas situações. A classificação de um conjunto de dados para o qual as classes exatas são conhecidas antecipadamente é denominado reconhecimento de padrões ou aprendizagem supervisionada (NIKAM, 2015).

Conforme Peterson (2009), o classificador k-vizinhos mais próximos (*k-Nearest Neighbor* – kNN) é um dos métodos de classificação mais fundamentais e simples, e deve ser uma das primeiras escolhas para um estudo de classificação quando há pouco ou nenhum conhecimento prévio sobre a distribuição dos dados. Apesar de ser simples e fácil de implementar, esse método é capaz de produzir resultados competitivos, mesmo em comparação com métodos de aprendizado de máquina mais sofisticados (SONG *et al.*, 2007).

A regra do vizinho mais próximo (*Nearest Neighbor* - NN) realiza a classificação da amostra desconhecida com base na amostra vizinha mais próxima, cuja classe já é conhecida (NIKAM, 2015). O método kNN é geralmente baseado na distância euclidiana entre uma amostra de teste e as amostras de treinamento especificadas. Seja x_i amostras de entrada com p características $(x_{i1}, x_{i2}, \dots, x_{ip})$, n o número total de amostras de entrada ($i = 1, 2, \dots, n$), e p o número total de características ($j = 1, 2, \dots, p$). A distância euclidiana entre as amostras x_i e x_l ($l = 1, 2, \dots, n$) é obtida de acordo com a Equação (2.34).

$$d(x_i, x_l) = \sqrt{(x_{i1} - x_{l1})^2 + (x_{i2} - x_{l2})^2 + \dots + (x_{ip} - x_{lp})^2} \quad (2.34)$$

Sejam x_i amostras de treinamento, x as amostras de teste, ω a classe verdadeira de uma amostra de treinamento, $\hat{\omega}$ a classe prevista para uma amostra de teste e Ω o número total de classes. Durante o processo de treinamento, usa-se apenas a classe verdadeira ω de

cada amostra de treino para treinar o classificador. No teste – ou na utilização do classificador – a classe $\hat{\omega}$ de cada amostra de teste é prevista. Com NN, ou kNN com $k=1$, a classe prevista da amostra de teste \mathbf{x} é definida igual à classe verdadeira ω do seu vizinho mais próximo, onde \mathbf{m}_i é o vizinho mais próximo de \mathbf{x} de acordo com a Equação (2.35).

$$d(\mathbf{m}_i, \mathbf{x}) = \operatorname{argmin}_j d(\mathbf{m}_j, \mathbf{x}) \quad (2.35)$$

Para kNN, a classe prevista da amostra de teste \mathbf{x} é definida como a classe verdadeira mais frequente entre as k amostras de treinamento mais próximas.

Na metodologia proposta neste trabalho, utiliza-se o classificador kNN para selecionar os coeficientes dos modelos a serem utilizados na monitoração de acordo com os valores das VE.

3 PROPOSTA DE METODOLOGIA

3.1 DEFINIÇÕES GERAIS

Seja a matriz de dados $X \in \mathfrak{R}^{n \times m}$ descrita por

$$X = [H \ G] \quad (3.1)$$

sendo $H \in \mathfrak{R}^{n \times mh}$ a matriz das VP, $G \in \mathfrak{R}^{n \times mg}$ a matriz das VE, n o número de amostras, m o número de variáveis, mh o número de variáveis principais e $mg = (m - mh)$ o número de variáveis externas.

A matriz H é composta por mh vetores $h_i \in \mathfrak{R}^{n \times 1}$, e a matriz G por mg vetores $g_j \in \mathfrak{R}^{n \times 1}$, descritas nas equações (3.2) e (3.3).

$$H = [h_1 \ h_2 \ \dots \ h_i \ \dots \ h_{mh}] \quad (3.2)$$

$$G = [g_1 \ g_2 \ \dots \ g_j \ \dots \ g_{mg}] \quad (3.3)$$

Quando há a presença de não linearidades no processo e se aplica a Análise Externa Linear para detecção de falhas, um modelo de regressão linear pode não modelar adequadamente todas as regiões de operação consideradas, gerando resíduos altos, e diminuindo, assim, a sensibilidade do algoritmo para monitoração de falhas. Nesses casos, alguns estudos utilizam a análise externa não linear, porém, para tal, há a necessidade de definições complexas comuns para modelos não lineares.

Considerando esses fatores, o principal objetivo deste trabalho é propor, implementar e avaliar uma metodologia que utilize a análise externa linear para realizar detecção de falhas em processos industriais, relacionando, através de modelos lineares, as variáveis externas e principais em todas as regiões de operação, utilizando, se necessário, mais de um modelo linear para tratar a presença de não linearidades nas relações das variáveis.

Para isso, na metodologia proposta, os coeficientes $c_{i,j}$ relacionam h_i com g_j de acordo com

$$\begin{bmatrix} h_1^T \\ h_2^T \\ \dots \\ h_i^T \\ \dots \\ h_{mh}^T \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,mg} \\ \dots & \dots & c_{i,j} & \dots \\ c_{mh,1} & c_{mh,2} & \dots & c_{mh,mg} \end{bmatrix} \begin{bmatrix} g_1^T \\ g_2^T \\ \dots \\ g_j^T \\ \dots \\ g_{mg}^T \end{bmatrix} + E \quad (3.4)$$

em que $c_{i,j}$ é um elemento do vetor $\theta_{i,j,k}$, $k = 1, \dots, L_{i,j}$, o qual conterà possíveis valores de coeficientes do modelo para a variável principal i . O elemento $c_{i,j}$ será selecionado por

um classificador, que usará as amostras do processo para esta classificação. A matriz E é a matriz de resíduos, composta por mh vetores $e_i \in \mathfrak{R}^{n \times 1}$, descrita por

$$E = [e_1 \ e_2 \ \dots \ e_i \ \dots \ e_{mh}] \quad (3.5)$$

A matriz E é a parte decomposta das variáveis principais que não depende das variáveis externas, portanto, pode ser utilizada para monitoração de falhas em múltiplas regiões de operação.

O tamanho $L_{i,j}$ do vetor $\theta_{i,j,k}$ será igual à quantidade de coeficientes lineares considerados necessários para modelar as relações entre a variável principal i e a variável externa j . Quando um único modelo for considerado necessário, $L_{i,j}$ será igual a um, quando houver a presença de não linearidades nas relações entre a variável principal i e a variável externa j mais de um elemento deve existir no vetor $\theta_{i,j,k}$.

A abordagem proposta neste trabalho se divide em duas fases:

1. Construção dos modelos;
2. Monitoração.

3.2 CONSTRUÇÃO DOS MODELOS

Nesta fase, o RLS é utilizado para acompanhar variações nos coeficientes dos modelos que relacionam as variáveis externas e principais. Estes coeficientes são agrupados de forma que os múltiplos modelos gerados possam reduzir os resíduos. Classificadores são, então, utilizados para associar as variáveis externas já rotuladas durante a fase de agrupamento aos grupos de modelos. Os resíduos gerados são utilizados para monitoramento das falhas.

As seções subsequentes destinam-se à explicação detalhada das etapas que compõem a fase de construção de modelos.

3.2.1 OBTENÇÃO E ANÁLISE DOS DADOS

Nesta etapa, identifica-se as variáveis externas e principais do processo, de acordo com suas definições feitas na seção 2.6.

As seguintes considerações são feitas em relação aos dados adquiridos do processo para o treinamento e a validação dos modelos:

- As faixas de operação de cada VE do processo devem ser identificadas, de modo que os dados sejam obtidos contendo as regiões de operação existentes no processo. Para isso, as VE devem ter variações ricas dentro das suas faixas de operação, dessa forma, os modelos obtidos através desses dados irão abranger todas as regiões.
- Considera-se que as VE variam entre estados estacionários, ou seja, sistemas que contenham VE variando de forma contínua no sistema não foram considerados, pois,

como será explicado na seção 3.2.2, as amostras do regime transitório serão desconsideradas na fase da construção de modelos.

- Considera-se, ainda, que as VE não possuem o comportamento de um *setpoint* que varia entre regiões de valores constantes, pois isso geraria singularidade na matriz de dados.
- Por último, considera-se que as VE não variam simultaneamente, pois, na metodologia proposta neste trabalho, identifica-se a influência de cada VE separadamente nas VP.

Os dados obtidos do processo devem ser normalizados considerando as médias e os desvios padrões das variáveis obtidas no estado estacionário de referência. A normalização remove os efeitos causados pelo uso de variáveis com diferentes escalas de medição. Evita-se, assim, que a escala da variável atribua à mesma uma influência maior sobre o método (PETERSON, 2009). Seja \mathbf{X} a matriz de dados composta com dados normalizados do processo.

3.2.2 ESTIMADOR RECURSIVO DE MÍNIMOS QUADRADOS

Nesta etapa, o RLS é aplicado para estimar os possíveis coeficientes entre a variável principal i e a variável externa j em diferentes regiões de operação formando o vetor $\hat{\theta}_{i,j,p}$, que será posteriormente avaliado na etapa de agrupamento para verificar a necessidade do uso de múltiplos modelos.

Como na maioria dos processos, nem todas as VE influenciam todas as VP e, como foi visto na seção 2.8, há algumas vantagens geradas pela realização de uma seleção de variáveis. Para cada VP, será realizada uma seleção das VE que farão parte do seu modelo, fazendo com que modelos não sejam gerados contendo preditores (VE) desnecessários na sua estrutura. Para isso, pode-se selecionar as variáveis com base em um valor mínimo definido para o coeficiente de correlação de Spearman ou realizar uma análise de gráficos de dispersão entre variáveis externas e principais.

Com isto, o RLS será aplicado no vetor \mathbf{h}_i de cada variável principal i , saída do modelo a ser estimado, tendo como entrada o vetor \mathbf{g}_j de cada variável externa j que a influencia, para gerar os elementos para os vetores $\hat{\theta}_{i,j,p}$ correspondentes.

Como visto na seção 2.9, para evitar que a dinâmica do regime transitório do sistema afete a construção dos modelos, optou-se, nesta fase, por somente utilizar dados do processo em regime estacionário. Para isso, pode-se escolher manualmente segmentos de dados do processo para compor a matriz de dados \mathbf{X} , ou pode-se utilizar algum método de avaliação do comportamento estacionário para desabilitar a obtenção de coeficientes via RLS durante o regime transitório das variáveis, como o método proposto por Cao e Rhinehart (1995), por exemplo.

Esse método de avaliação do comportamento estacionário proposto por Cao e Rhinehart (1995) requer a definição de alguns parâmetros, sendo eles: o valor crítico φ_{CRIT} e os

valores de ponderação λ_1 , λ_2 e λ_3 . A definição de φ_{CRIT} pode ser realizada através da observação dos valores da relação estatística φ_{STAT} obtidos via um conjunto de dados no qual as amostras dos regimes estacionário e transitório são conhecidas, definindo, então, φ_{CRIT} a partir dos valores do limiar de φ_{STAT} na região estacionária. Para os valores de ponderação λ_1 , λ_2 e λ_3 , os mesmos podem ser definidos como $\lambda_1 = 0.2$ e $\lambda_2 = \lambda_3 = 0.1$ de acordo com a indicação de Cao e Rhinehart (1995; 1997), que afirma que esses valores proporcionam um adequado balanço dos compromissos da redução dos erros do Tipo-I e do Tipo-II.

Para aplicação do RLS, deve-se definir um fator de esquecimento (FE) adequado. Saeed, Zerguine e Zummo (2014) realizam uma análise do impacto do fator de esquecimento sobre o desempenho dos algoritmos propostos, comparando os resultados para os valores de $\lambda = \{0.9, 0.95, 0.99\}$. Chalam (1987) sugere que um fator de esquecimento de 0.95 seja usado para variações rápidas. Landau *et al.* (2011) afirma que para um fator de esquecimento constante, valores típicos vão de 0.95 a 0.99.

Na metodologia proposta neste trabalho, para definir um fator de esquecimento adequado, deve-se considerar que mudanças nas variáveis externas podem levar o sistema para regiões com as mesmas relações lineares entre VE e VP que a região anterior, ou para regiões com relações lineares diferentes. Isso deve ser identificado corretamente pelo estimador, ou seja, o RLS deve estimar novos valores de coeficientes para regiões de operação diferentes da região anterior, e não deve variar significativamente os coeficientes estimados em uma mesma região.

Observa-se que é possível variar entre estados estacionários e não variar a relação linear entre as variáveis, no entanto, considerando que em estado estacionário os dados sofrem somente pequenas variações causadas apenas por ruídos e não por mudanças de VE, não é possível variar entre relações lineares em um mesmo estado estacionário. Portanto, os coeficientes de um único estado estacionário nunca devem ter variações significativas. Variações indevidas desses coeficientes podem ser causadas por oscilações nos dados devido às características da recursividade influenciadas pela definição do fator de esquecimento.

De acordo com esses aspectos e com as características intrínsecas de cada sistema, é possível realizar uma análise do impacto do fator de esquecimento na obtenção dos coeficientes dos modelos, para, assim, definir um fator de esquecimento adequado.

3.2.3 OBTENÇÃO DOS MODELOS LINEARES

Nesta etapa, os vetores de coeficientes $\hat{\theta}_{i,j,p}$ obtidos pelo RLS serão analisados para formar os vetores de coeficientes $\theta_{i,j,k}$, que conterão apenas os elementos associados aos coeficientes necessários em cada modelo.

Para avaliar a quantidade $L_{i,j}$ de coeficientes lineares necessários, ou seja, o tamanho do vetor $\theta_{i,j,k}$, propõe-se comparar a norma dos resíduos provenientes da aplicação de diferentes números de coeficientes.

Dessa forma, para cada $\hat{\theta}_{i,j,p}$, primeiramente faz-se $L_{i,j} = 1$ e um único coeficiente linear deve ser obtido –via Mínimos Quadrados – através de h_i e g_j . Os resíduos relativos a esse coeficiente são obtidos e com eles calcula-se sua norma. A seguir, aplica-se a *clusterização* via *k-means* em $\hat{\theta}_{i,j,p}$ para gerar $L_{i,j}$ clusters, iniciando com $L_{i,j} = 2$. Essa *clusterização* gerará rótulos para cada elemento de $\hat{\theta}_{i,j,p}$, bem como para as amostras em h_i e g_j . Com isso, $L_{i,j}$ coeficientes lineares são gerados – via mínimos quadrados – com as amostras de h_i e g_j correspondentes a cada *cluster* de $\hat{\theta}_{i,j,p}$, gerando, assim, o vetor $\omega_{i,j,k}$, $k = 1, \dots, L_{i,j}$, contendo tais coeficientes. Em seguida, os resíduos relativos a cada coeficiente devem ser obtidos, e com eles calcula-se a norma dos resíduos. Repete-se o mesmo procedimento aumentando o número de *clusters* $L_{i,j}$.

A norma dos resíduos pode diminuir indefinidamente com o aumento do número de coeficientes utilizados. Por esse motivo, optou-se por:

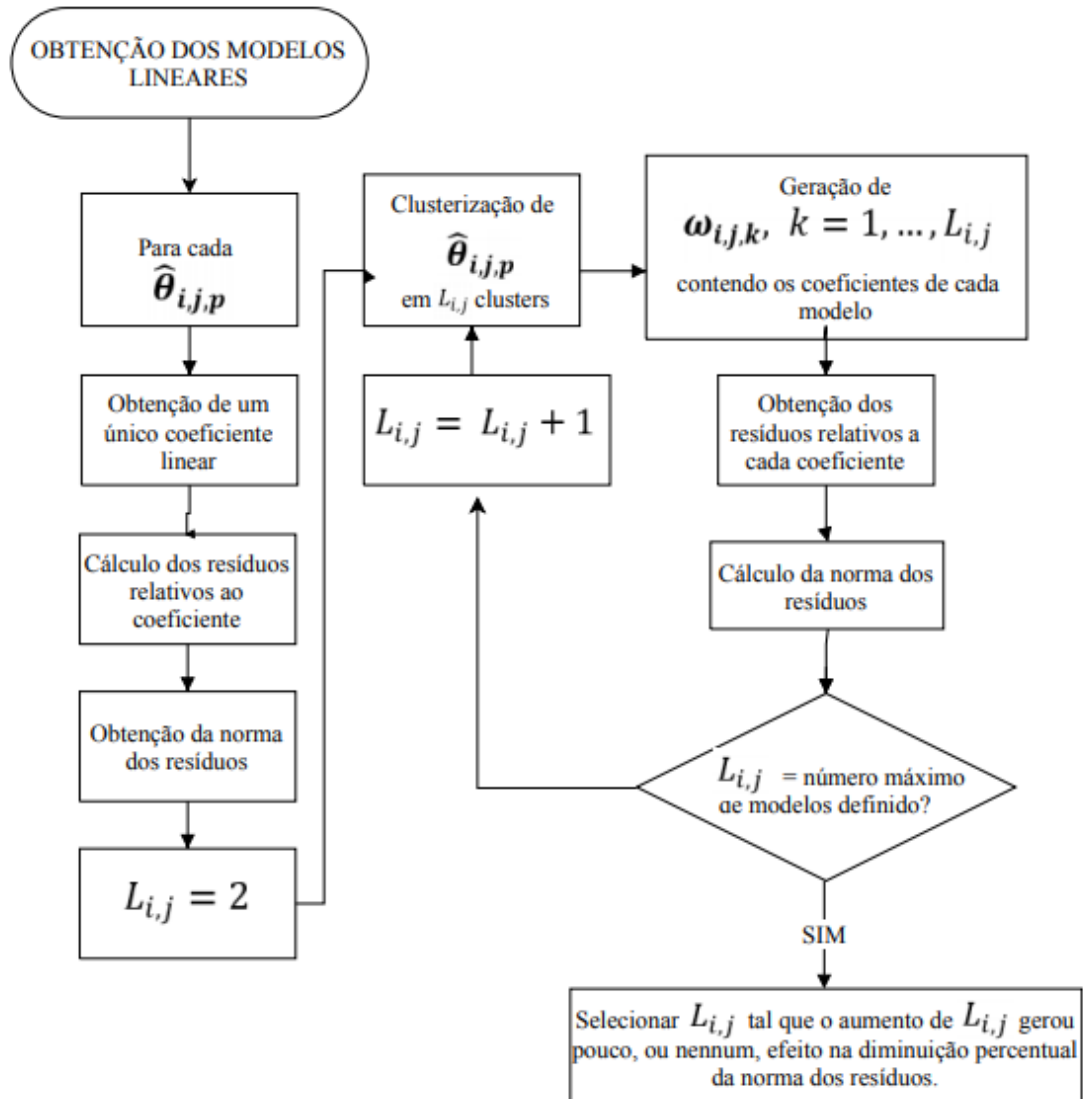
- Definir um número máximo para $L_{i,j}$;
- Selecionar um $L_{i,j}$ tal que um aumento do seu valor gere pouco – ou nenhum – efeito na diminuição percentual da norma dos resíduos. Para isso, definiu-se um limite mínimo que será considerado para essa diminuição.

Dessa forma, o vetor $\theta_{i,j,k}$ escolhido será o $\omega_{i,j,k}$ referente ao $L_{i,j}$ selecionado a partir dos critérios relatados.

Optou-se por definir tal limite para que a decisão da escolha do aumento do número de coeficientes considere o efeito negativo deste aumento, que é a exigência de mais memória e processamento dos computadores que executarão o algoritmo. Com esse limite, pode-se avaliar se o aumento do número de coeficientes gera uma melhoria no resultado da previsão dos modelos que compensa o efeito negativo gerado pelo mesmo. No entanto, isso deve ser avaliado de acordo com a precisão requerida na monitoração de cada processo e com a capacidade de memória e processamento que se tem para implementar o algoritmo de monitoração.

A etapa descrita nesta seção está resumida no fluxograma da Figura 3.

Figura 3 - Fluxograma da etapa: obtenção dos modelos lineares



Fonte: Próprio autor

3.2.4 TREINAMENTO DOS CLASSIFICADORES

Para cada $\theta_{i,j,k}$ que contenha mais de um elemento, um classificador será necessário para, na etapa de monitoração, classificar qual coeficiente $c_{i,j}$ utilizar para cada amostra dos dados monitorados. A necessidade de alterar os coeficientes lineares acontece quando as regiões de operação mudam para novas regiões nas quais a relação linear usada anteriormente não é ideal. Como a mudança de regiões de operação é provocada por mudanças nas VE, os valores destas serão as características utilizadas na classificação.

O classificador kNN foi escolhido pois – como visto na Seção 2.10 – o kNN é um método de classificação simples e, apesar disto, pode produzir resultados competitivos, mesmo em comparação com outros métodos de classificação mais sofisticados.

O treinamento de cada classificador será realizado com 70% das amostras de cada g_j relativas a $\theta_{i,j,k}$, rotuladas através do *k-means*. Seu desempenho será medido através da medição da acurácia, utilizando o restante das amostras de g_j relativas a $\theta_{i,j,k}$.

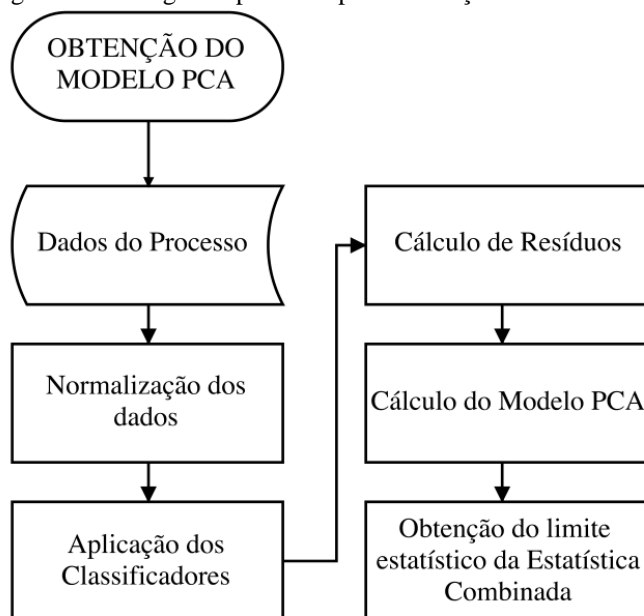
3.2.5 OBTENÇÃO DO MODELO PCA

Kano et al. (2004) explica que qualquer método de SPC pode ser utilizado para monitorar as variáveis dos resíduos. O PCA é um dos principais métodos de MSPC utilizados para monitoração, e será o método utilizado neste trabalho. Para isso, primeiramente é preciso calcular seu modelo através das variáveis de resíduos.

Para obtenção das variáveis de resíduos, deve-se aplicar os classificadores em dados do processo para selecionar os coeficientes $c_{i,j}$, e então calcular os seus resíduos através da Equação (3.4). Para isso, deve-se utilizar dados que não foram utilizados para a obtenção de $\theta_{i,j,k}$ e para o treinamento dos classificadores. Os dados utilizados na obtenção das variáveis de resíduos devem conter as diferentes regiões de operação existentes no sistema e devem, também, ser normalizados, da mesma forma que a normalização realizada na Seção 3.2.1. Então, pode-se obter a matriz de resíduos desses dados.

O modelo PCA deve ser obtido de acordo com a Seção 2.3, considerando a matriz de resíduos como a matriz de dados. O limite estatístico da estatística combinada deve ser obtido de acordo com a Seção 2.4. A Figura 4 apresenta um fluxograma que resume esta etapa.

Figura 4 - Fluxograma para a etapa de obtenção do modelo PCA

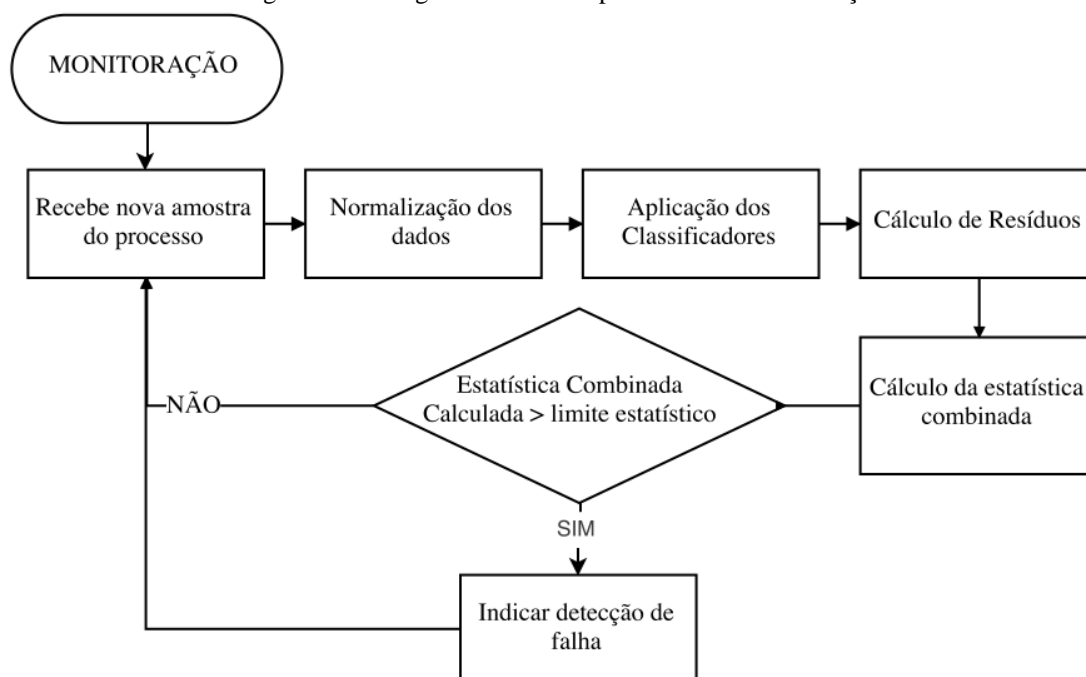


Fonte: Próprio autor

3.3 MONITORAÇÃO

Na fase de monitoração, cada nova amostra deve ser normalizada da mesma forma que os dados foram tratados na Seção 3.2.1. Em seguida, os classificadores devem ser aplicados em cada amostra para selecionar os coeficientes $c_{i,j}$ que serão utilizados para calcular os resíduos da mesma. Então, com os resíduos, deve-se calcular a estatística combinada de acordo com a Seção 2.4. Caso a estatística combinada calculada para essa amostra seja maior que o seu limite estatístico, um estado de falha é detectado, caso contrário, uma condição normal do processo é indicada. Esta etapa está resumida no fluxograma da Figura 5.

Figura 5 - Fluxograma detalhado para a fase de monitoração



Fonte: Próprio autor

4 APLICAÇÃO E ANÁLISE DOS RESULTADOS

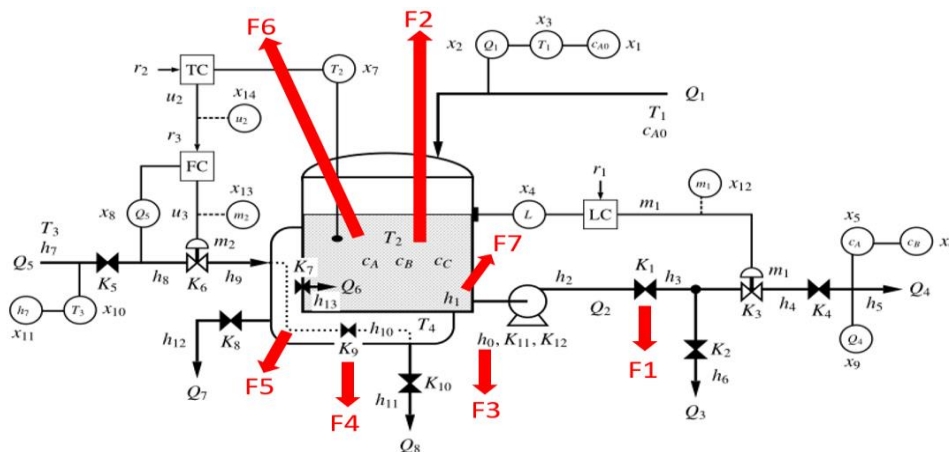
4.1 INTRODUÇÃO

A metodologia proposta neste trabalho é aplicada ao CSTR, um simulador de um reator químico enjaquetado onde ocorre uma reação exotérmica. Primeiramente, será realizada a construção dos modelos desse sistema de acordo com o Capítulo 3. Em seguida, será verificada se a taxa de falsos positivos (FP) obtida com a aplicação é consideravelmente similar ao nível de significância definido. O desempenho do método proposto com múltiplos modelos (MM) será comparado ao desempenho do método da análise externa linear com um modelo único (M1) através da taxa de falsos negativos (FN). Por fim, é avaliada a independência das abordagens via MM e M1 em relação à região de operação e as características dos resíduos obtidos.

4.2 TANQUE REATOR CONTINUAMENTE AGITADO

O processo simulado pelo tanque reator continuamente agitado (*Continuously Stirred Reactor Tank – CSTR*) apresenta não linearidades e malhas de controle interagindo entre si, e permite simular falhas simples e complexas. A descrição analítica do simulador, ilustrado na Figura 6, encontra-se em (Finch, 1989).

Figura 6 - Diagrama do CSTR



Fonte: Finch (1989)

As variáveis do CSTR estão descritas na Tabela 1, com seus respectivos acrônimos apresentados na Figura 6. As falhas do CSTR que serão avaliadas neste trabalho são apresentadas na Tabela 2, na qual indica-se o parâmetro afetado que gera a falha, assim como seu valor nominal, ou seja, o valor do parâmetro no estado normal do processo. É possível definir o instante de ocorrência das falhas, assim como as suas intensidades, através da inclusão de um valor no parâmetro afetado que seja diferente do seu valor nominal.

Tabela 1 – Variáveis e restrições do simulador CSTR

#	Nome da Variável/Restrição	Acrônimo	Valor nominal	Unidade
1	Concentração do líquido de alimentação	cA0	20	Mol/m ³
2	Vazão do líquido de alimentação	F1	0.25	m ³ /s
3	Temperatura do líquido de alimentação	T1	30	°C
4	Nível do reator	L	2	m
5	Concentração do produto A	cA	2.85	Mol/m ³
6	Concentração do produto B	cB	17.11	Mol/m ³
7	Temperatura do Reator	T2	80	°C
8	Vazão do refrigerante	F5	0.9	m ³ /s
9	Vazão do produto	F4	0.25	m ³ /s
10	Temp. de entrada do refrigerante	T3	20	°C
11	Pressão de entrada do refrigerante	PCW	56250	Pa
12	Saída do Controlador de Nível	CNT1	74.7	----
13	Saída do Controlador de Resfriamento	CNT2	0.9	----
14	Setpoint do Refrigerante	CNT3	59.3	----
15	Estoque	r_1	0.0	----
16	Balanço de Mol	r_2	0.0	----
17	Perda de Pressão de Água de Refrigeração	r_3	0.0	----
18	Perda de Pressão do Efluente	r_4	0.0	----

Fonte: Finch (1989)

Tabela 2 – Falhas do CSTR avaliadas

#	Nome da Falha	Parâmetro Afetado	Valor nominal do parâmetro afetado
1	Bloqueio na saída do tanque	R_1	100
2	Reação secundária da energia de ativação	β_2	45000
3	Perda de pressão na bomba	PP	48000
4	Bloqueio na jaqueta	R_9	0
5	Incrustação da área de resfriamento da jaqueta	UA	1901
6	Reação primária da energia de ativação	β_1	25000
7	Pressão anormal dos efluentes do reator	REP	0

Fonte: Finch (1989)

4.3 APLICAÇÃO DA CONSTRUÇÃO DOS MODELOS

4.3.1 OBTENÇÃO E ANÁLISE DOS DADOS

De acordo com a metodologia proposta, primeiramente, deve-se definir as variáveis externas e principais do processo tratado. As VE são aquelas responsáveis por alterar a região de operação da planta, escolhidas com o objetivo de modelar as possíveis

mudanças dessas regiões, sendo, ainda, as variáveis de entrada no modelo da análise externa. As VP são as variáveis de saída do modelo, e podem sofrer variações causadas pelas VE sem que isto represente uma falha, ou podem sofrer variações devido a distúrbios do processo caracterizando falha. A definição das VE dependerá das regiões de operação de cada planta, que, por sua vez, dependem das características e funcionalidades do processo no qual a planta opera.

Para a aplicação da metodologia proposta neste trabalho, será considerado que a planta do CSTR opera em diferentes regiões de operação através de mudanças das variáveis 1, 2, 3, 10 e 11, que são variáveis de entrada do processo e independem umas das outras. Portanto, essas serão as VE.

Inicialmente, as variáveis 4, 5, 6, 7, 8, 9, 12, 13 e 14 serão consideradas VP, portanto, as alterações por elas sofridas que não possam ser estimadas pelas VE devem ser interpretadas como falhas.

As variáveis 15, 16, 17 e 18 são variáveis de restrições do processo, calculadas a partir das outras variáveis já citadas. Portanto, para evitar informações redundantes no algoritmo de monitoração, essas variáveis não serão consideradas. Os limites operacionais considerados para cada VE se encontram na Tabela 3.

Tabela 3 – Faixas de operação das VE

Índice da variável	Limite operacional inferior	Limite operacional superior
1	16	21.5
2	0.215	0.285
3	22	38
10	13	27
11	30000	90000

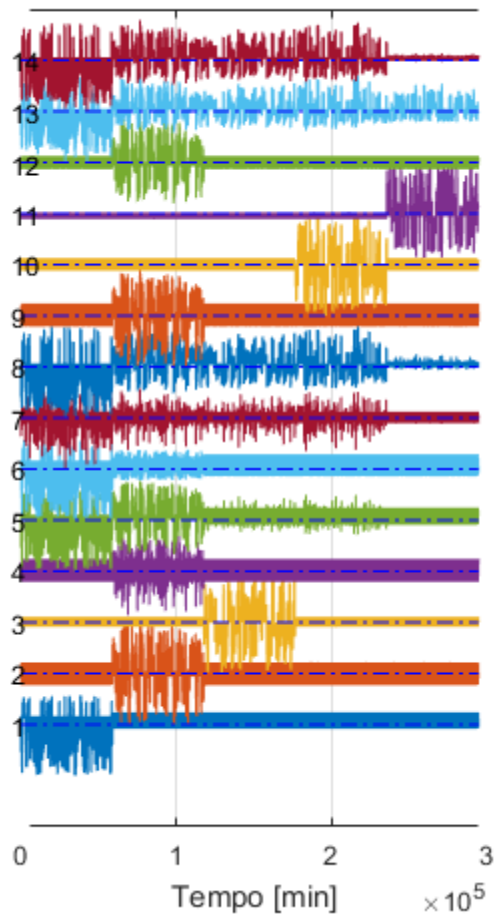
Fonte: Próprio autor

Para obter os dados do CSTR de forma a abranger todas as faixas de operação consideradas, os dados foram obtidos variando cada VE em 150 estados estacionários, nos quais seus valores médios foram escolhidos aleatoriamente dentro dos limites operacionais definidos para cada VE. O tempo de amostragem definido no CSTR foi de 1 minuto. Cada estado estacionário apresenta 400 amostras. As variações de cada VE foram realizadas na sequência crescente dos seus índices, ou seja, os dados obtidos contêm, primeiramente, dados originados de variações da VE 1, em seguida da VE 2, e assim por diante.

O estado estacionário inicial do processo foi selecionado como estado estacionário de referência e os valores nominais de suas variáveis encontram-se na Tabela 1. A média e o desvio padrão desse estado foram calculados a partir de uma simulação com 10000 amostras desse estado e, então, os dados foram normalizados.

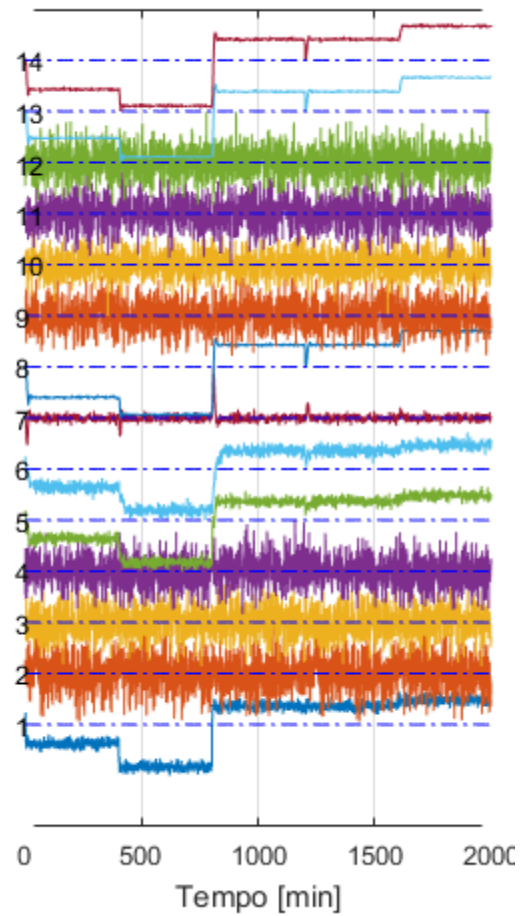
Por fim, as variáveis geraram a matriz de dados X . A Figura 7 apresenta os dados referentes a cada uma das variáveis. Na Figura 8 são apresentadas as 2000 primeiras amostras de X , podendo-se observar com mais detalhes os dados.

Figura 7 - Variáveis do CSTR da matriz de dados X



Fonte: Próprio autor

Figura 8 - Variáveis do CSTR com as primeiras 2000 amostras da matriz de dados X



Fonte: Próprio autor

Para que a identificação das variáveis externas e principais seja feita utilizando os mesmos índices das variáveis do CSTR apresentados na Tabela 1, as matrizes H e G serão compostas por

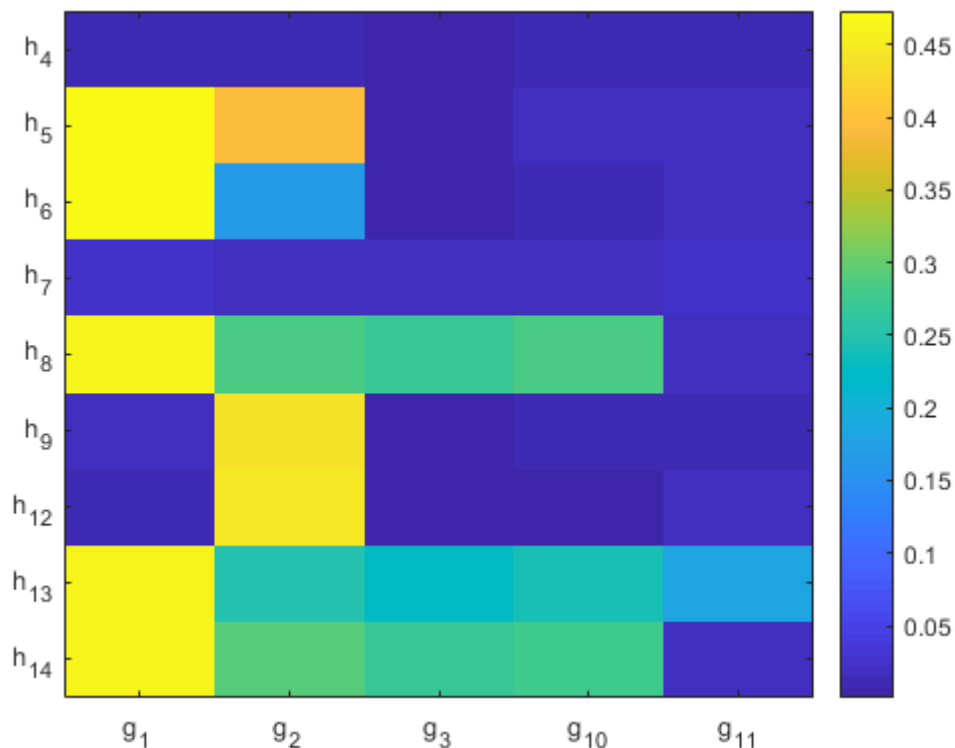
$$H = [h_4 \ h_5 \ h_6 \ h_7 \ h_8 \ h_9 \ h_{12} \ h_{13} \ h_{14}] \quad (4.1)$$

$$G = [g_1 \ g_2 \ g_3 \ g_{10} \ g_{11}] \quad (4.2)$$

4.3.2 ESTIMADOR RECURSIVO DE MÍNIMOS QUADRADOS

Para realizar uma seleção de variáveis com objetivo de escolher as melhores VE para o modelo de cada VP, primeiramente, os coeficientes de correlação de Spearman entre VE e VP foram calculados e estão ilustrados na Figura 9.

Figura 9 – Correlações de Spearman calculadas entre as VP e as VE



Fonte: Próprio autor

O limiar mínimo para a correlação – definido como referência para selecionar as VE a serem utilizadas no modelo de cada VP – foi 10%. A Tabela 4 indica as VE selecionadas para cada VP.

Tabela 4 – VE selecionadas para cada VP

VP	VE
h_4	-
h_5	$[g_1, g_2]$
h_6	$[g_1, g_2]$
h_7	-
h_8	$[g_1, g_2, g_3, g_{10}]$
h_9	g_2
h_{12}	g_2
h_{13}	$[g_1, g_2, g_3, g_{10}, g_{11}]$
h_{14}	$[g_1, g_2, g_3, g_{10}]$

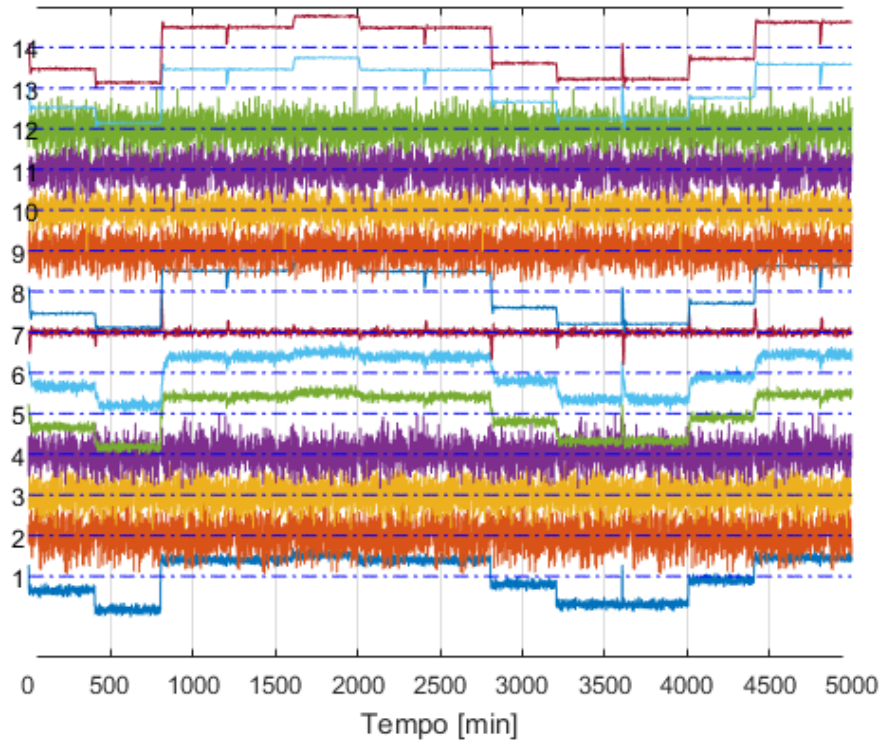
Fonte: Próprio autor

Observa-se que h_4 e h_7 não apresentaram nenhuma VE correlacionada mais do que 10%, portanto serão desconsideradas.

Para evitar que a dinâmica do regime transitório do sistema afete a construção dos modelos, optou-se por desabilitar a obtenção de coeficientes via RLS durante o regime transitório, utilizando o método de avaliação do comportamento estacionário proposto por Cao e Rhinehart (1995). Para sua aplicação, os valores de ponderação λ_1 , λ_2 e λ_3 foram definidos como sendo $\lambda_1 = 0.2$ e $\lambda_2 = \lambda_3 = 0.1$, de acordo com a indicação de

Cao e Rhinehart (1995; 1997). Para a definição do φ_{CRIT} , o φ_{STAT} foi analisado para as primeiras 5000 amostras da matriz de dados X , apresentadas na Figura 10.

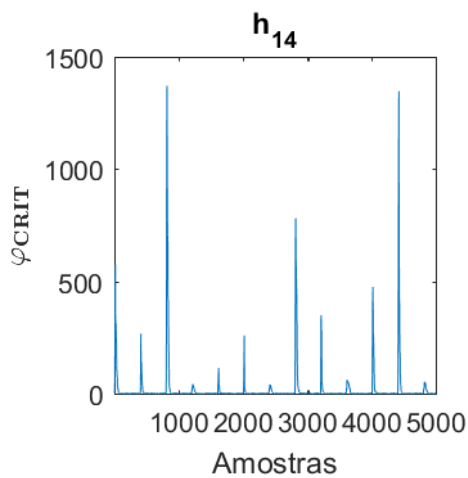
Figura 10 - Variáveis do CSTR com as primeiras 5000 amostras da matriz de dados X



Fonte: Próprio autor

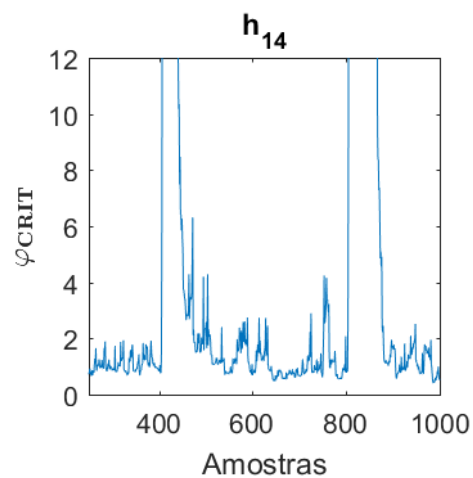
Os valores de φ_{STAT} obtidos para as VP podem ser observados nas Figuras 11 a 24 que contêm, para cada VP, uma figura com o φ_{STAT} calculado para as 5000 amostras dos dados da Figura 10, e uma figura contendo apenas o φ_{STAT} das amostras 250 a 1000, para que se possa observar mais detalhadamente o seu comportamento.

Figura 11 - φ_{STAT} calculado para h_{14}

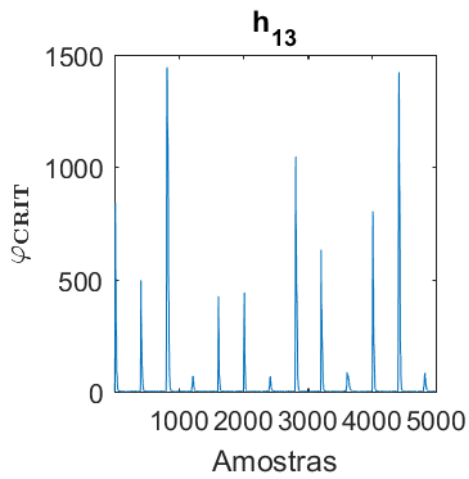


Fonte: Próprio autor

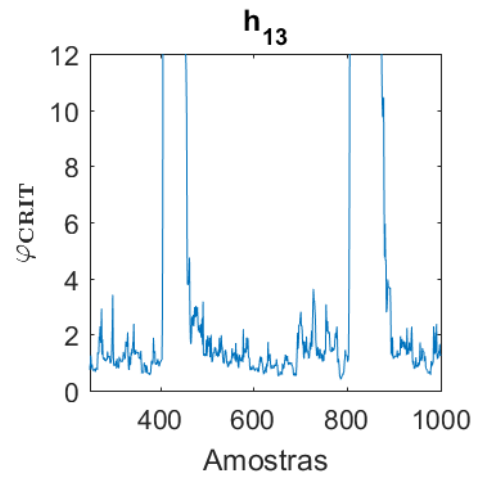
Figura 12 - Amostras 250 a 1000 do φ_{STAT} calculado para h_{14}



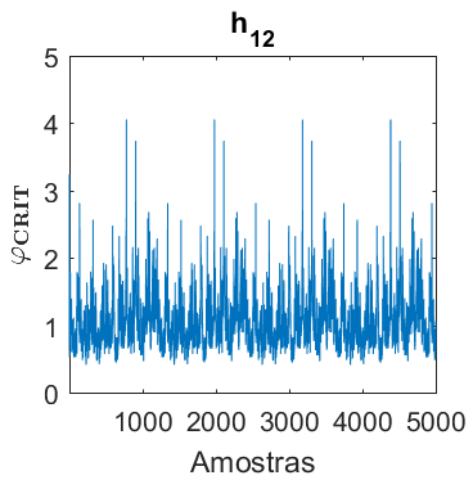
Fonte: Próprio autor

Figura 13 - φ_{STAT} calculado para h_{13} 

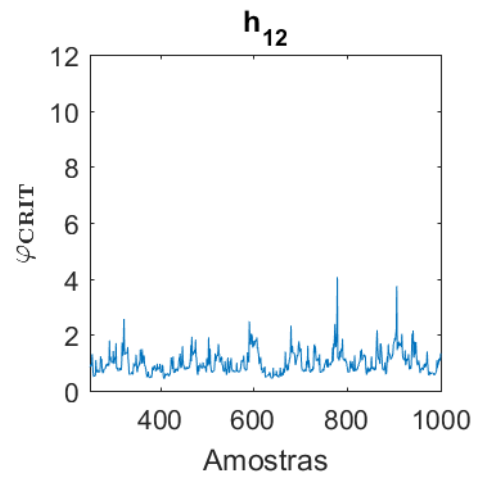
Fonte: Próprio autor

Figura 14 - Amostras 250 a 1000 do φ_{STAT} calculado para h_{13} 

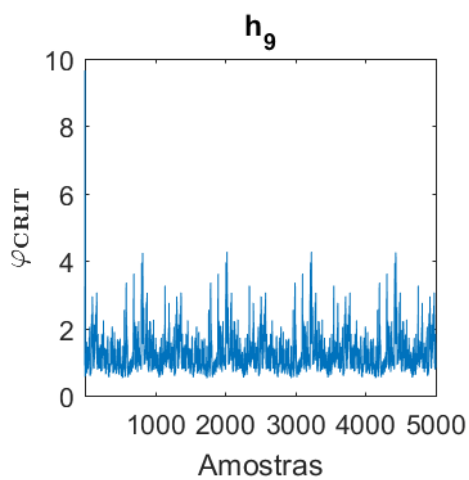
Fonte: Próprio autor

Figura 15 - φ_{STAT} calculado para h_{12} 

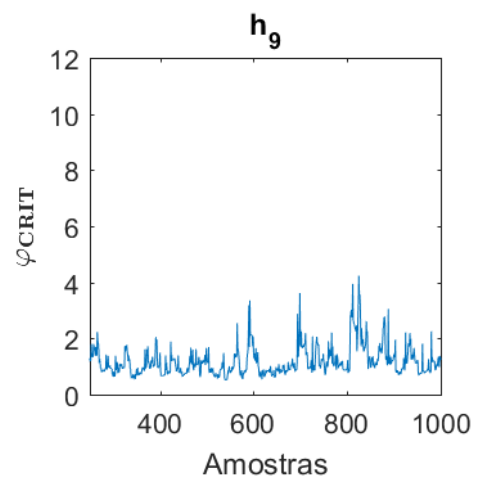
Fonte: Próprio autor

Figura 16 - Amostras 250 a 1000 do φ_{STAT} calculado para h_{12} 

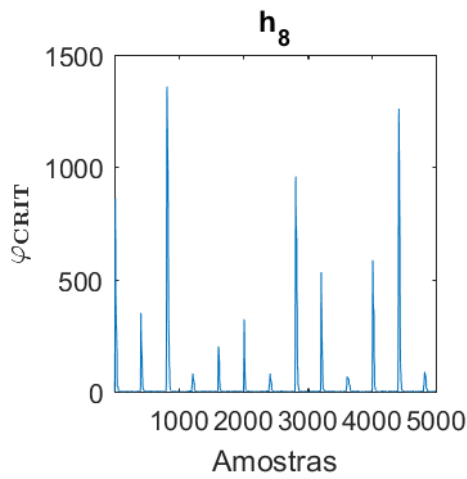
Fonte: Próprio autor

Figura 17 - φ_{STAT} calculado para h_9 

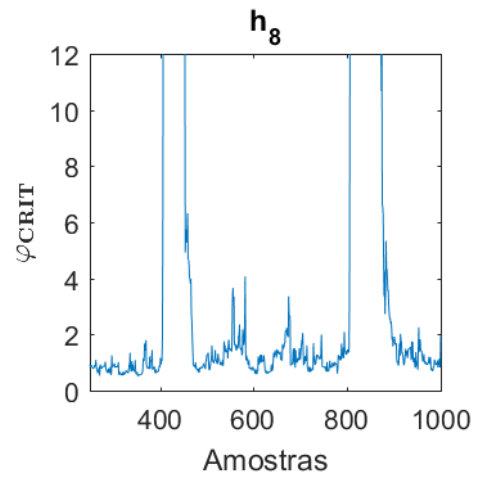
Fonte: Próprio autor

Figura 18 - Amostras 250 a 1000 do φ_{STAT} calculado para h_9 

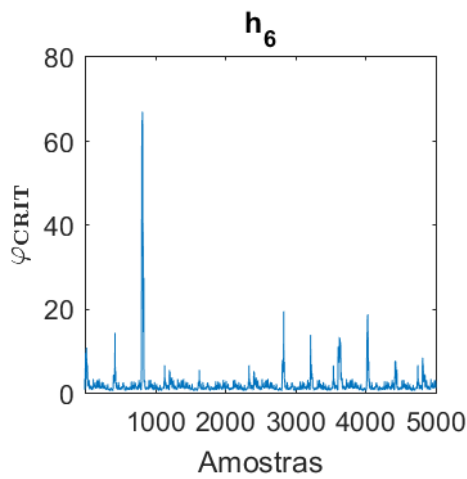
Fonte: Próprio autor

Figura 19 - φ_{STAT} calculado para h_8 

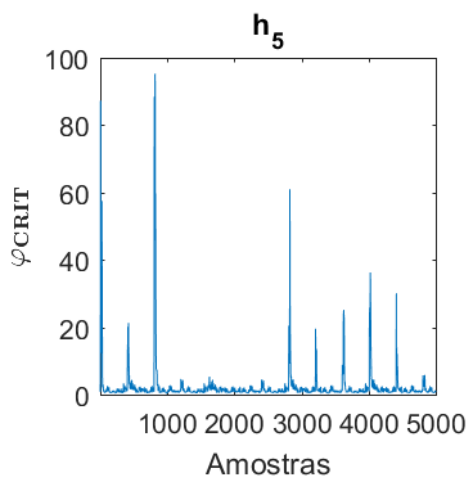
Fonte: Próprio autor

Figura 20 - Amostras 250 a 1000 do φ_{STAT} calculado para h_8 

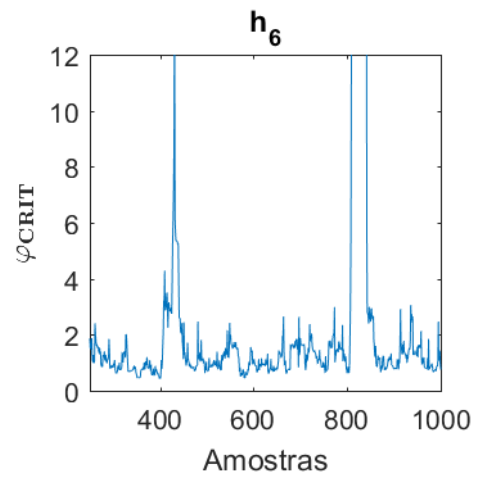
Fonte: Próprio autor

Figura 21 - φ_{STAT} calculado para h_6 

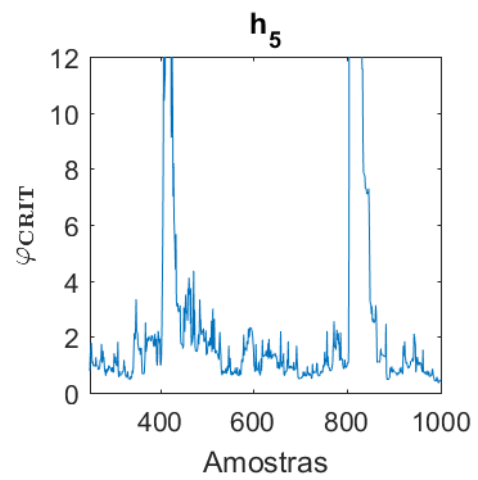
Fonte: Próprio autor

Figura 23 - φ_{STAT} calculado para h_5 

Fonte: Próprio autor

Figura 22 - Amostras 250 a 1000 do φ_{STAT} calculado para h_6 

Fonte: Próprio autor

Figura 24 - Amostras 250 a 1000 do φ_{STAT} calculado para h_5 

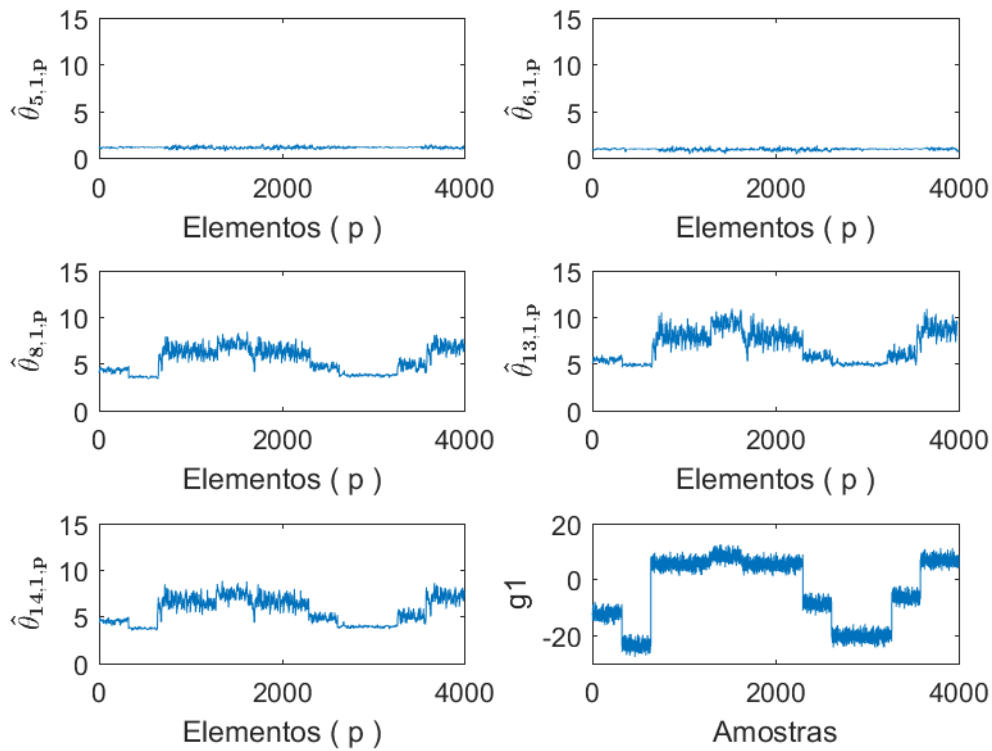
Fonte: Próprio autor

De acordo com os resultados, pode-se observar que φ_{STAT} se mantém, aproximadamente, abaixo de 6 durante os estados estacionários, portanto o φ_{CRIT} utilizado será 6.

Para definir um fator de esquecimento adequado, o impacto do mesmo na obtenção dos coeficientes dos modelos também será analisado através dos dados da Figura 10. Nela, pode-se observar que a única VE que está variando entre estados estacionários é a g_1 , que afeta as variáveis h_5, h_6, h_8, h_{13} e h_{14} , corroborando o resultado da seleção de variáveis que selecionou a g_1 para essas variáveis principais.

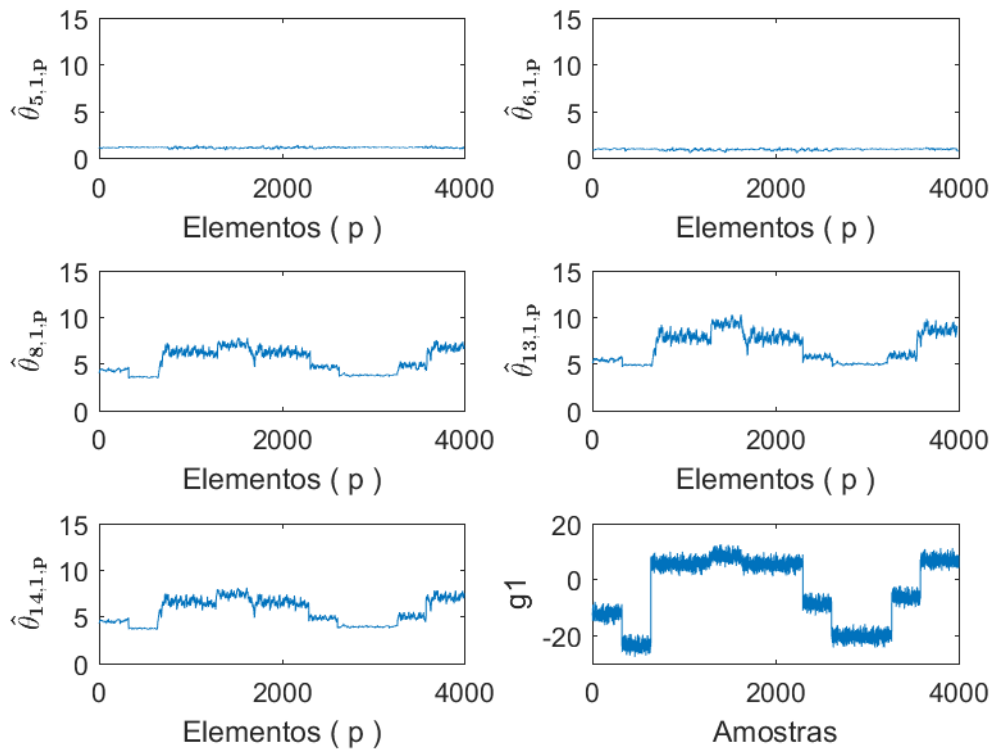
Em seguida, os coeficientes para esses dados serão calculados de acordo com a notação do Capítulo 3. A metodologia proposta na seção 3.2.2 será aplicada cinco vezes, uma para cada um dos seguintes valores do fator de esquecimento: 0.8, 0.9, 0.95, 0.97, 0.99. Os elementos de $\hat{\theta}_{5,1,p}, \hat{\theta}_{6,1,p}, \hat{\theta}_{8,1,p}, \hat{\theta}_{13,1,p}, \hat{\theta}_{14,1,p}$ obtidos com a aplicação de cada fator de esquecimento testado podem ser visualizados da Figura 25 até a Figura 29.

Figura 25 – Coeficientes obtidos com o fator de esquecimento 0.8



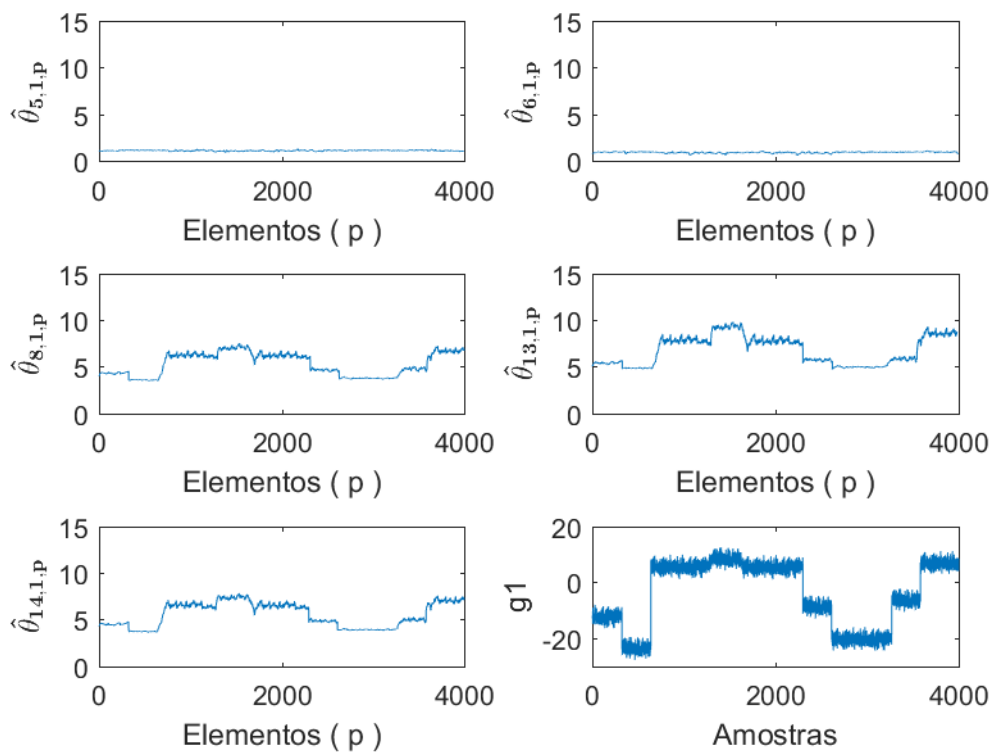
Fonte: Próprio autor

Figura 26 – Coeficientes obtidos com o fator de esquecimento 0.9



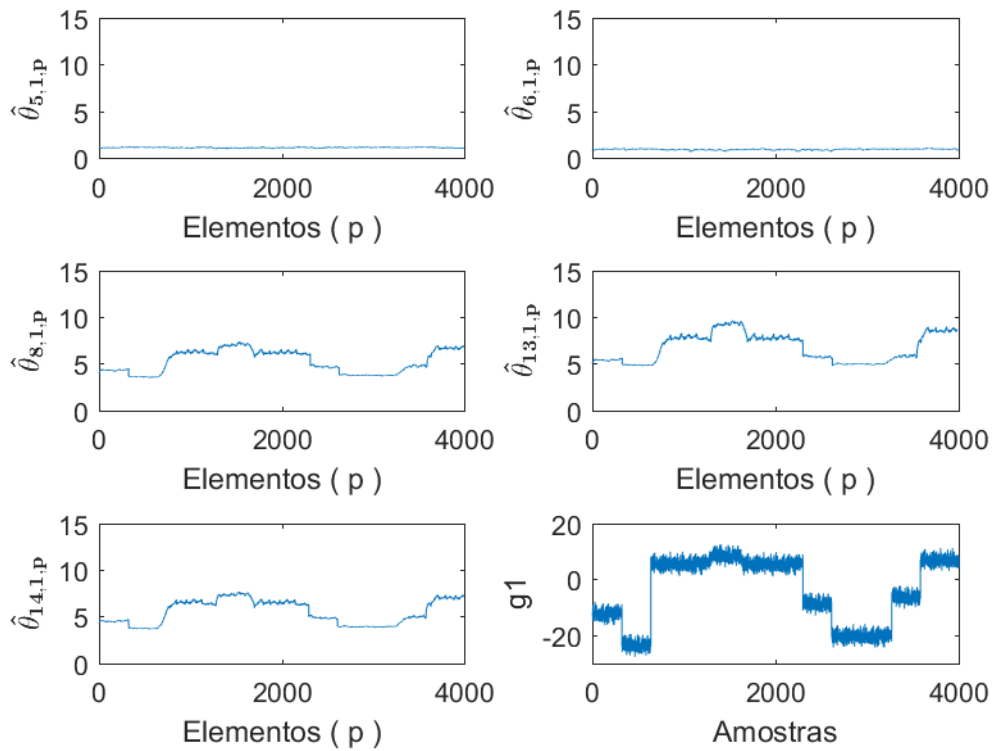
Fonte: Próprio autor

Figura 27 – Coeficientes obtidos com o fator de esquecimento 0.95



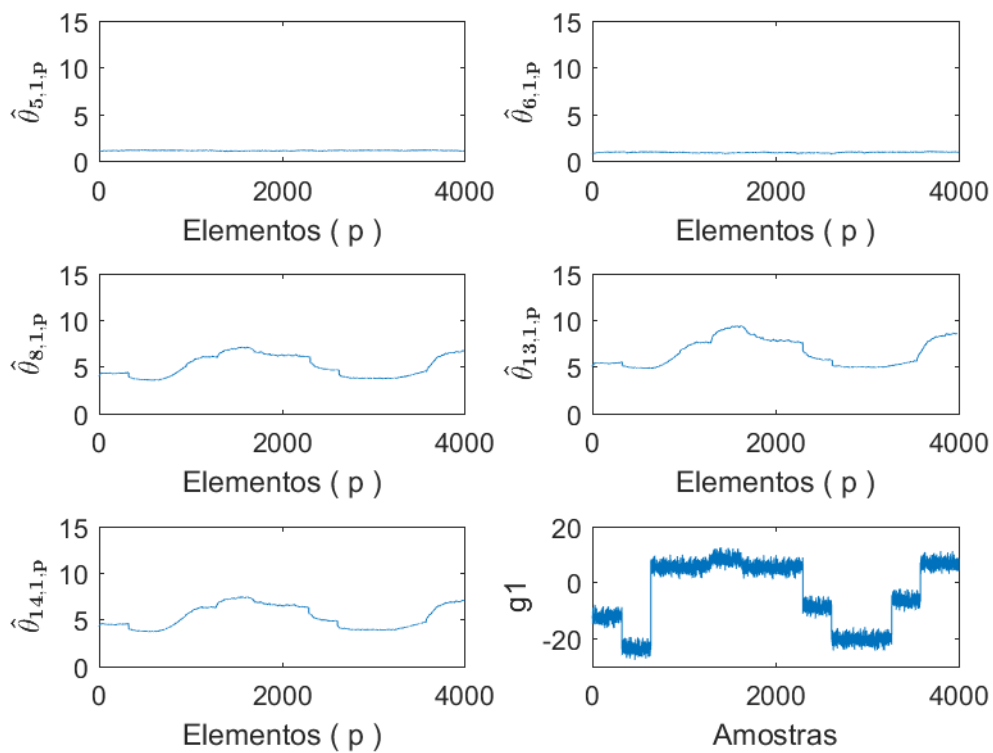
Fonte: Próprio autor

Figura 28 – Coeficientes obtidos com o fator de esquecimento 0.97



Fonte: Próprio autor

Figura 29 – Coeficientes obtidos com o fator de esquecimento 0.99



Fonte: Próprio autor

Na Figura 25, pode-se observar os coeficientes resultantes da aplicação do fator de esquecimento 0.8. Os elementos de $\hat{\theta}_{5,1,p}$ e $\hat{\theta}_{6,1,p}$ mantiveram-se constantes, o que indica que a relação linear é a mesma em todas as regiões de operação testadas com esses dados. Resultados similares foram obtidos com os outros valores dos fatores de esquecimento testados. Já os elementos de $\hat{\theta}_{8,1,p}$, $\hat{\theta}_{13,1,p}$ e $\hat{\theta}_{14,1,p}$ foram para outros patamares com variações de g_1 . No entanto, observa-se variações consideráveis nos elementos em um mesmo estado estacionário.

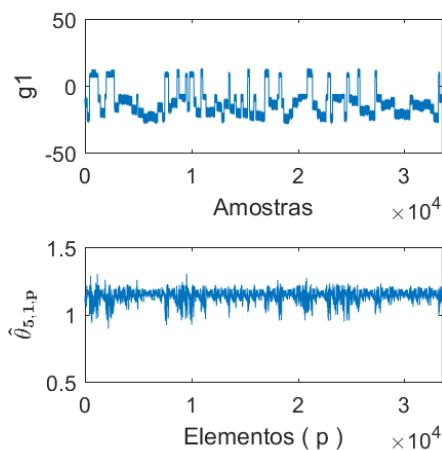
Na Figura 26, pode-se observar os coeficientes resultantes da aplicação do fator de esquecimento 0.9 apresentando um resultado similar ao da Figura 25, no entanto, neste caso, as variações dos coeficientes obtidos dentro de um mesmo estado estacionário diminuíram.

A Figura 27 apresenta os coeficientes resultantes da aplicação do fator de esquecimento 0.95. Vê-se que houveram variações bem delimitadas dos elementos de $\hat{\theta}_{8,1,p}$, $\hat{\theta}_{13,1,p}$ e $\hat{\theta}_{14,1,p}$ em concordância com variâncias de g_1 , e uma variação menor em um mesmo estado estacionário, se comparado ao resultado utilizando fator de esquecimento 0.9.

Na Figuras 28 e 29 pode-se observar os resultados para o fator de esquecimento 0.97 e 0.99, respectivamente. Observa-se que os elementos obtidos em $\hat{\theta}_{8,1,p}$, $\hat{\theta}_{13,1,p}$ e $\hat{\theta}_{14,1,p}$ variaram mais lentamente com as mudanças de g_1 , principalmente para 0.99. Este comportamento não é desejável para a metodologia proposta, uma vez que o comportamento das variações dos coeficientes devem estar em concordância com as variâncias de g_1 .

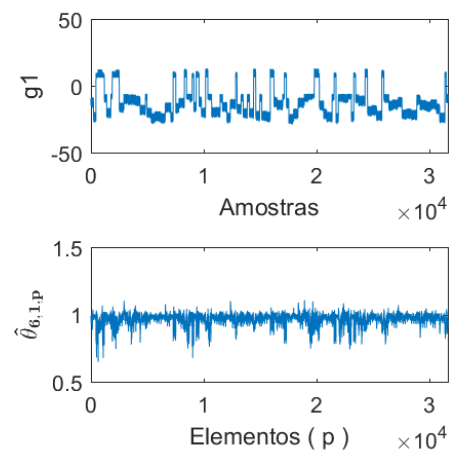
Portanto, o fator de esquecimento 0.95 mostrou-se o mais adequado para esta aplicação e será utilizado para obter os coeficientes utilizando a matriz de dados X . As Figuras 30 até 48 apresentam os coeficientes $\hat{\theta}_{i,j,p}$ resultantes desta aplicação, assim como os gráficos das respectivas amostras de g_j utilizadas pelo estimador.

Figura 30 – Vetor de coeficientes $\hat{\theta}_{5,1,p}$ obtidos via RLS



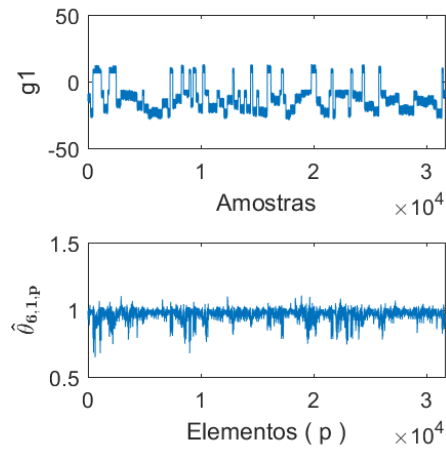
Fonte: Próprio autor

Figura 31 – Vetor de coeficientes $\hat{\theta}_{5,2,p}$ obtidos via RLS



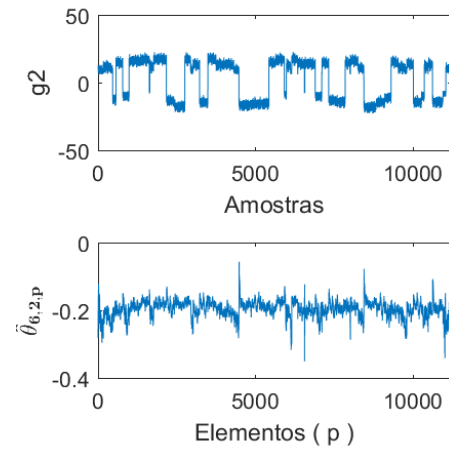
Fonte: Próprio autor

Figura 32 – Vetor de coeficientes $\hat{\theta}_{6,1,p}$ obtidos via RLS



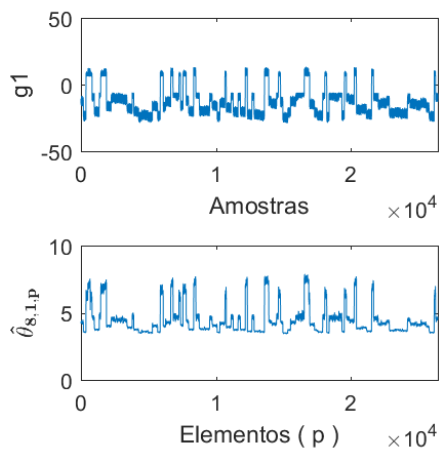
Fonte: Próprio autor

Figura 33 – Vetor de coeficientes $\hat{\theta}_{6,2,p}$ obtidos via RLS



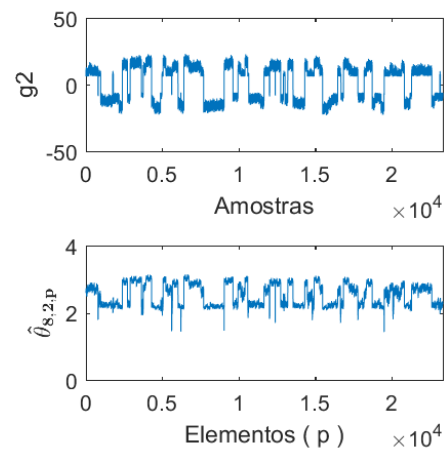
Fonte: Próprio autor

Figura 34 – Vetor de coeficientes $\hat{\theta}_{8,1,p}$ obtidos via RLS



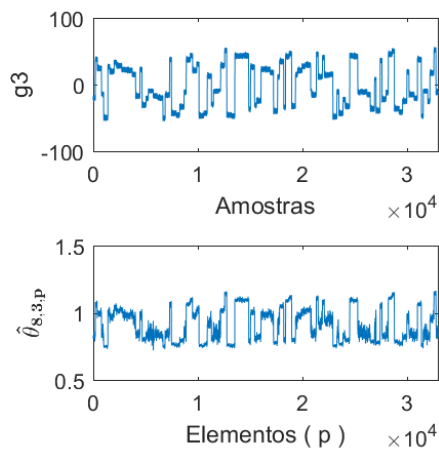
Fonte: Próprio autor

Figura 35 – Vetor de coeficientes $\hat{\theta}_{8,2,p}$ obtidos via RLS



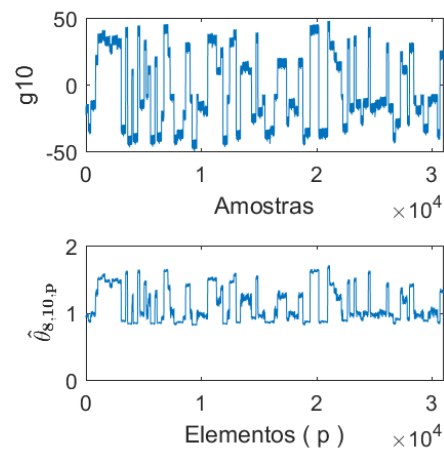
Fonte: Próprio autor

Figura 36 – Vetor de coeficientes $\hat{\theta}_{8,3,p}$ obtidos via RLS



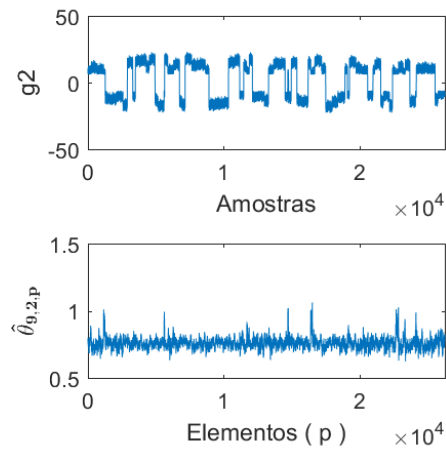
Fonte: Próprio autor

Figura 37 – Vetor de coeficientes $\hat{\theta}_{8,10,p}$ obtidos via RLS



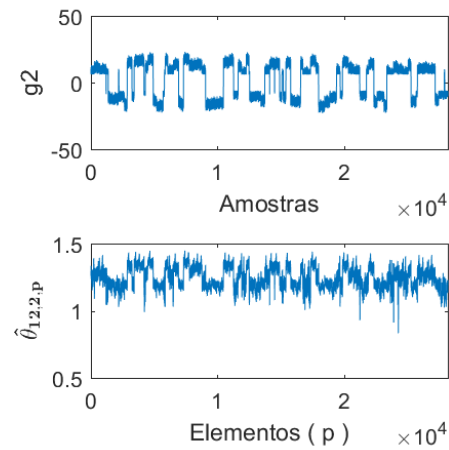
Fonte: Próprio autor

Figura 38 – Vetor de coeficientes $\hat{\theta}_{9,2,p}$ obtidos via RLS



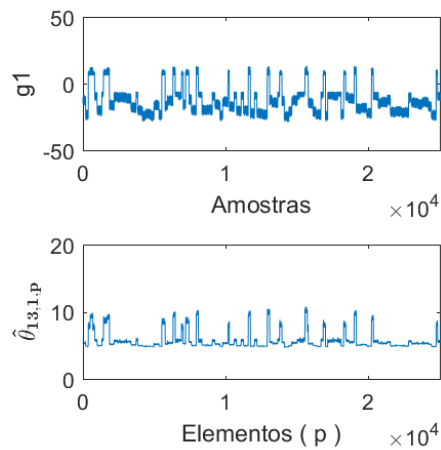
Fonte: Próprio autor

Figura 39 – Vetor de coeficientes $\hat{\theta}_{12,2,p}$ obtidos via RLS



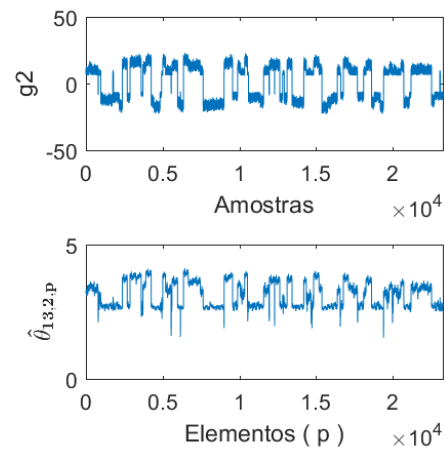
Fonte: Próprio autor

Figura 40 – Vetor de coeficientes $\hat{\theta}_{13,1,p}$ obtidos via RLS



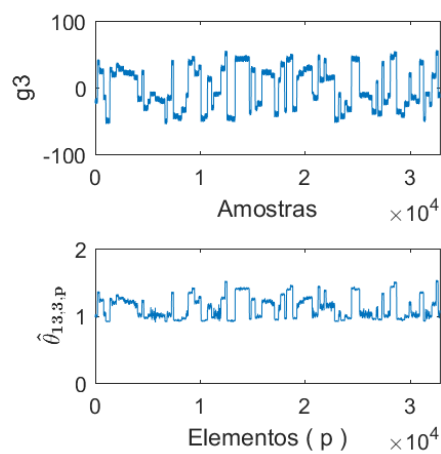
Fonte: Próprio autor

Figura 41 – Vetor de coeficientes $\hat{\theta}_{13,2,p}$ obtidos via RLS



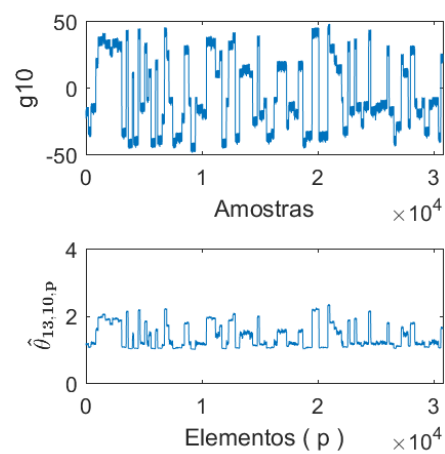
Fonte: Próprio autor

Figura 42 – Vetor de coeficientes $\hat{\theta}_{13,3,p}$ obtidos via RLS



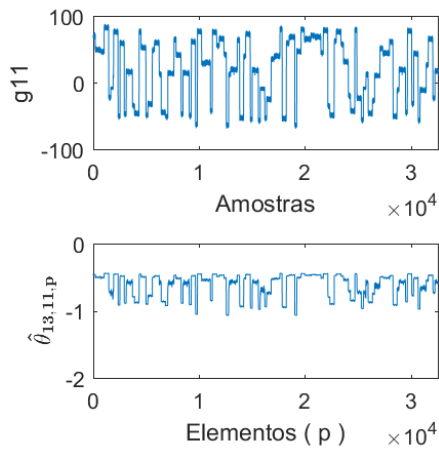
Fonte: Próprio autor

Figura 43 – Vetor de coeficientes $\hat{\theta}_{13,10,p}$ obtidos via RLS



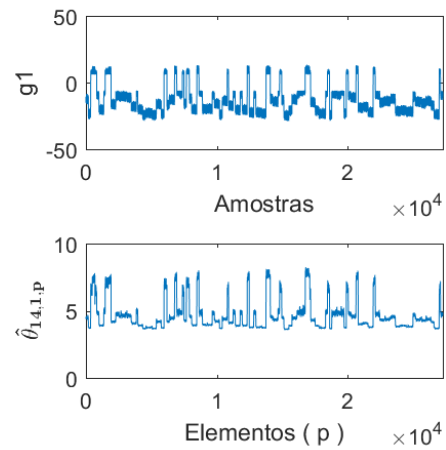
Fonte: Próprio autor

Figura 44 – Vetor de coeficientes $\hat{\theta}_{13,11,p}$ obtidos via RLS



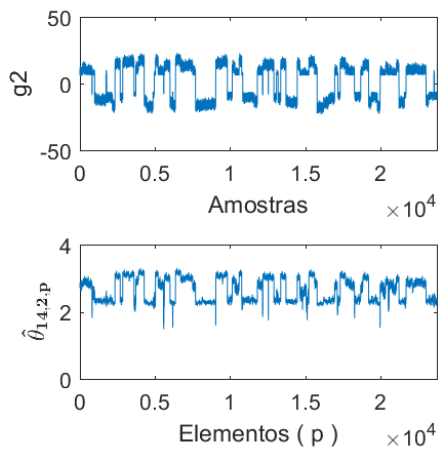
Fonte: Próprio autor

Figura 45 – Vetor de coeficientes $\hat{\theta}_{14,1,p}$ obtidos via RLS



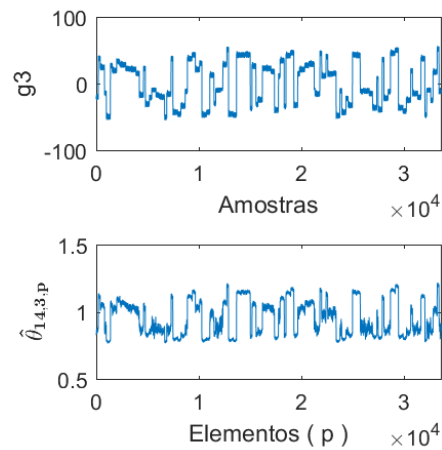
Fonte: Próprio autor

Figura 46 – Vetor de coeficientes $\hat{\theta}_{14,2,p}$ obtidos via RLS



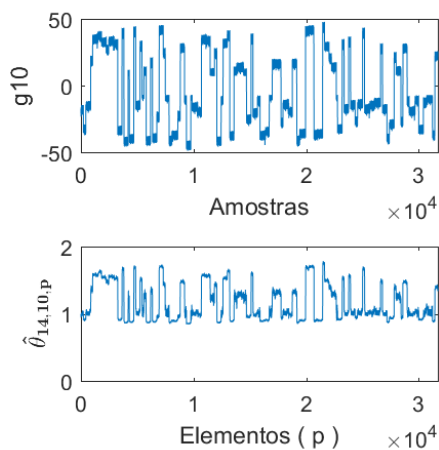
Fonte: Próprio autor

Figura 47 – Vetor de coeficientes $\hat{\theta}_{14,3,p}$ obtidos via RLS



Fonte: Próprio autor

Figura 48 – Vetor de coeficientes $\hat{\theta}_{14,10,p}$ obtidos via RLS



Fonte: Próprio autor

4.3.3 OBTENÇÃO DOS MODELOS LINEARES

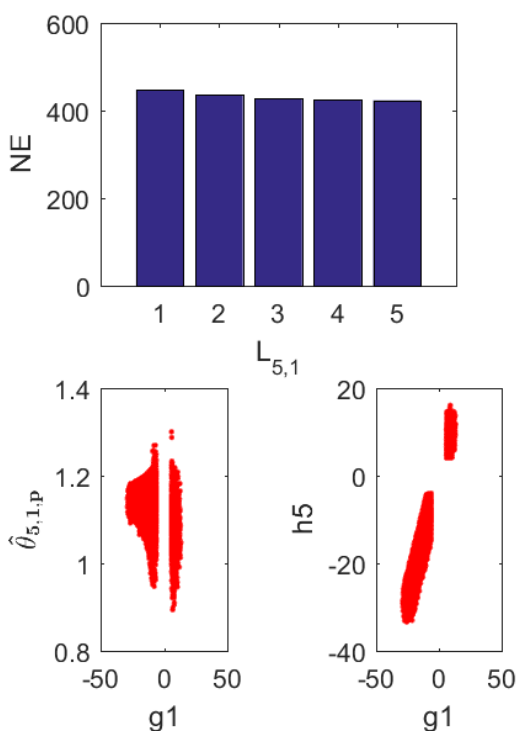
A metodologia explicada na seção 3.2.3 foi aplicada para cada vetor de coeficientes $\hat{\theta}_{i,j,p}$ obtido via RLS, e o número máximo de modelos considerados nesta aplicação foi 5. A mínima redução percentual da norma dos resíduos que foi considerada para aumentar o número de modelos foi 20%.

Nos resultados apresentados nas Figuras 49 a 67, pode-se observar para cada $\hat{\theta}_{i,j,p}$:

- As normas dos resíduos (NE) obtidas para $L_{i,j}$ de 1 até 5;
- Um gráfico de dispersão entre os coeficientes $\hat{\theta}_{i,j,p}$ obtidos via RLS e os valores de g_j ;
- Um gráfico de dispersão entre as amostras consideradas de g_j e h_i .

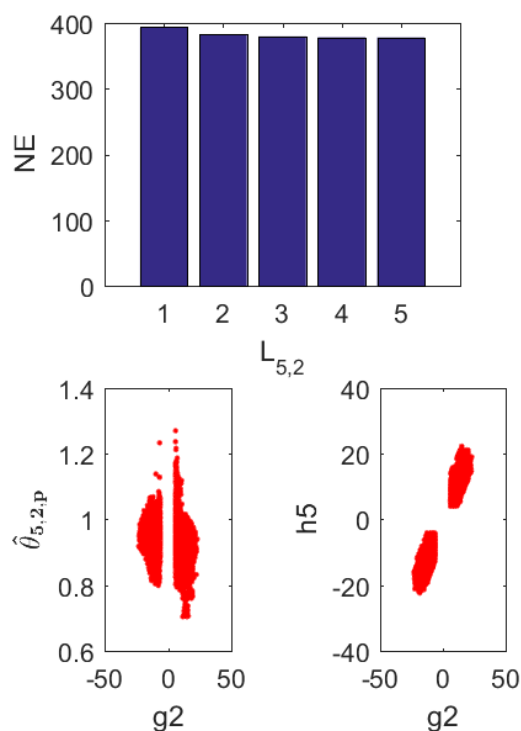
Nos gráficos de dispersão, as amostras foram coloridas de acordo com os rótulos obtidos via a *clusterização* relativa ao $L_{i,j}$ escolhido, ou seja, de acordo com o número de coeficientes selecionado para formar $\theta_{i,j,k}$.

Figura 49 – Análise de $\hat{\theta}_{5,1,p}$ para seleção de $\theta_{5,1,k}$



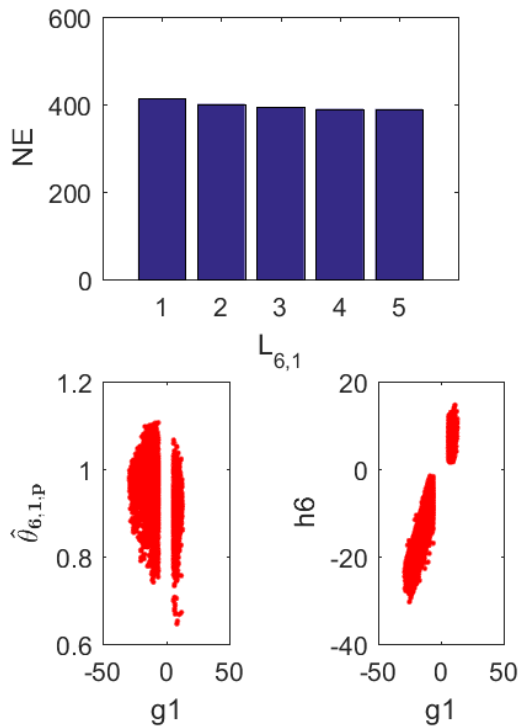
Fonte: Próprio autor

Figura 50 – Análise de $\hat{\theta}_{5,2,p}$ para seleção de $\theta_{5,2,k}$



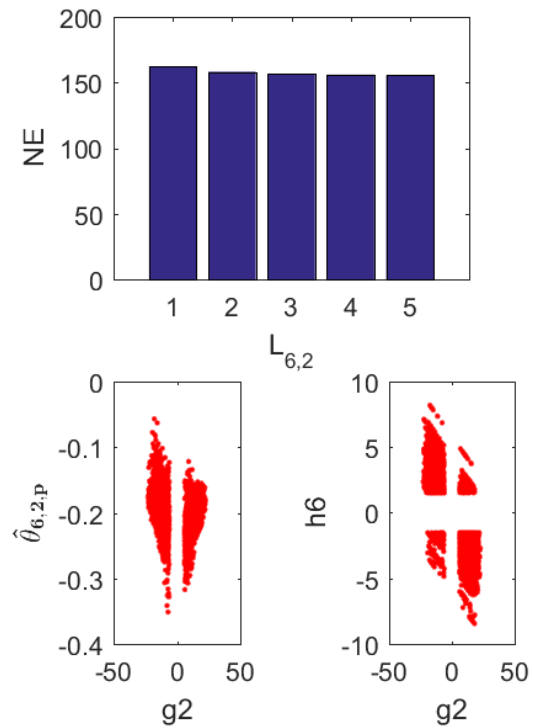
Fonte: Próprio autor

Figura 51 – Análise de $\hat{\theta}_{6,1,p}$ para seleção de $\theta_{6,1,k}$



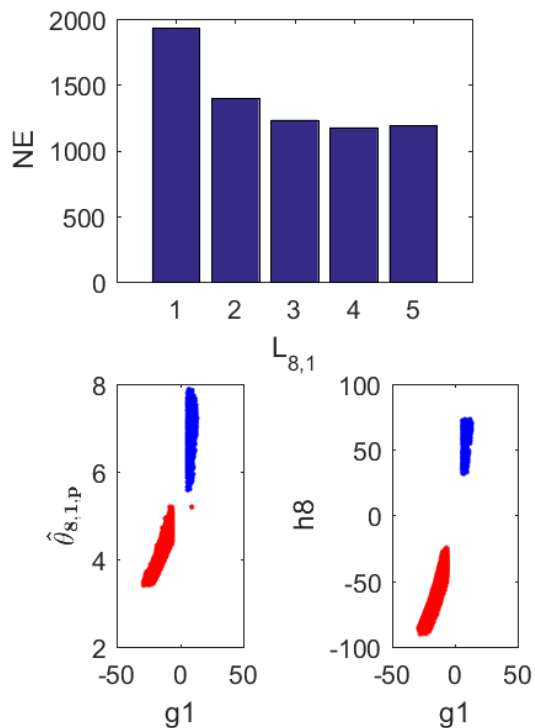
Fonte: Próprio autor

Figura 52 – Análise de $\hat{\theta}_{6,2,p}$ para seleção de $\theta_{6,2,k}$



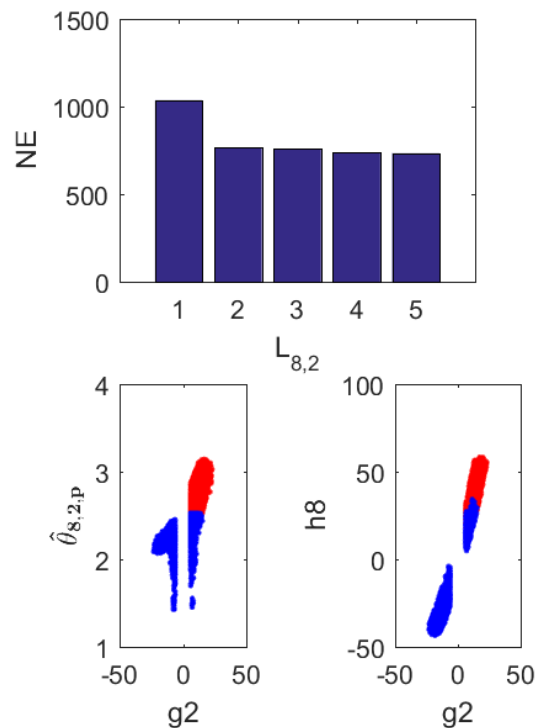
Fonte: Próprio autor

Figura 53 – Análise de $\hat{\theta}_{8,1,p}$ para seleção de $\theta_{8,1,k}$



Fonte: Próprio autor

Figura 54 – Análise de $\hat{\theta}_{8,2,p}$ para seleção de $\theta_{8,2,k}$



Fonte: Próprio autor

Figura 55 – Análise de $\hat{\theta}_{8,3,p}$ para seleção de $\theta_{8,3,k}$

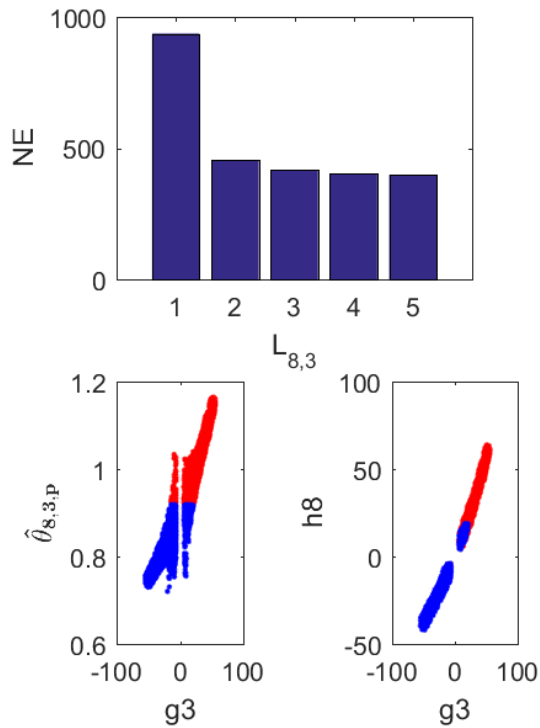


Figura 56 – Análise de $\hat{\theta}_{8,10,p}$ para seleção de $\theta_{8,10,k}$

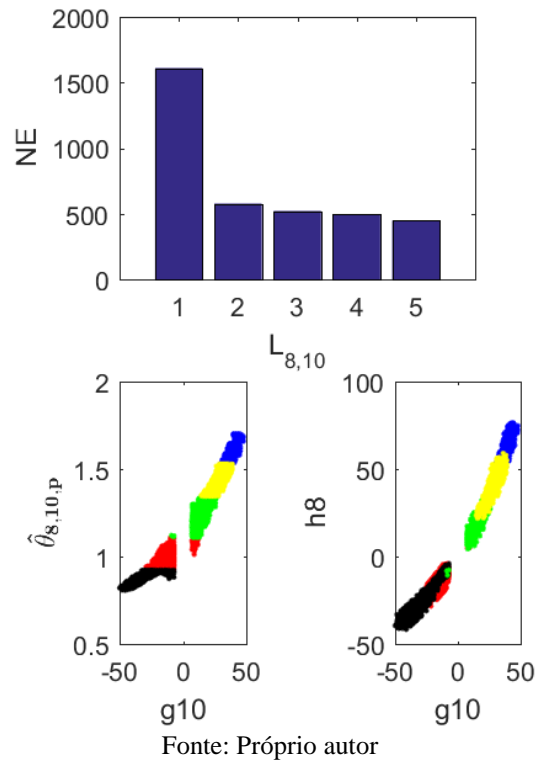


Figura 57 – Análise de $\hat{\theta}_{9,2,p}$ para seleção de $\theta_{9,2,k}$

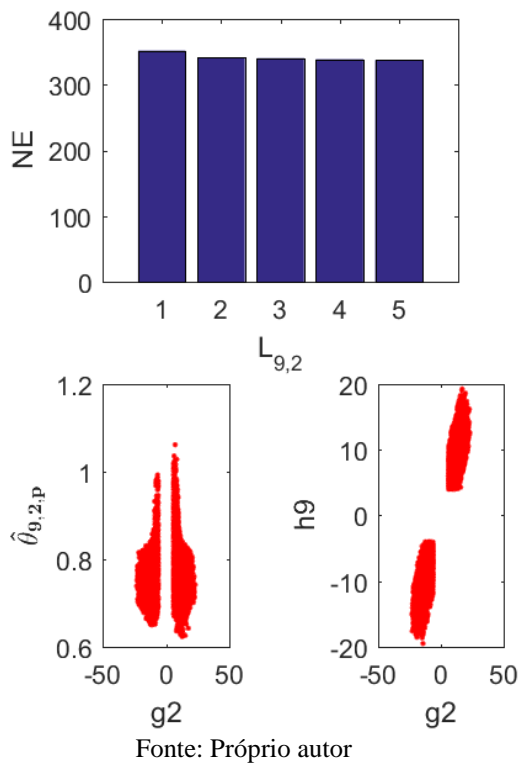


Figura 58 – Análise de $\hat{\theta}_{12,2,p}$ para seleção de $\theta_{12,2,k}$

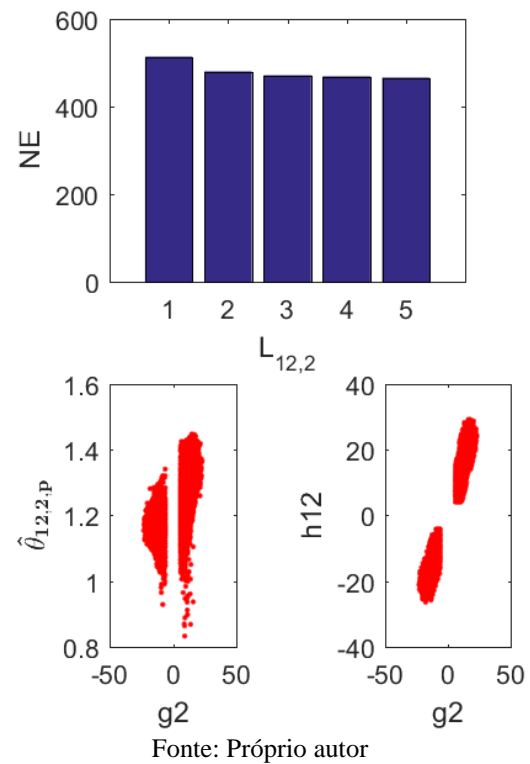


Figura 59 – Análise de $\hat{\theta}_{13,1,p}$ para seleção de $\theta_{13,1,k}$

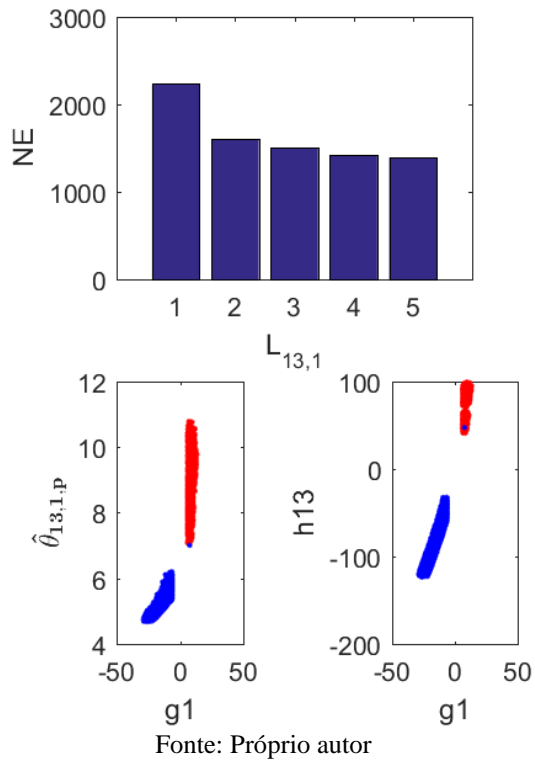


Figura 60 – Análise de $\hat{\theta}_{13,2,p}$ para seleção de $\theta_{13,2,k}$

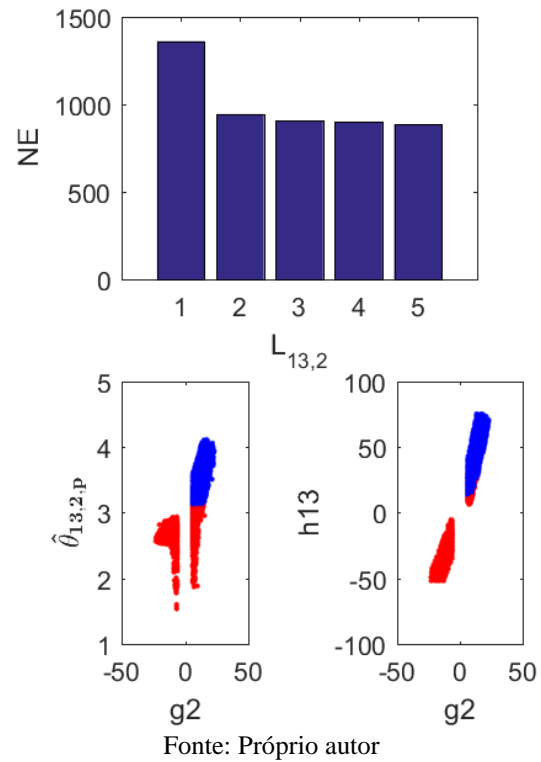


Figura 61 – Análise de $\hat{\theta}_{13,3,p}$ para seleção de $\theta_{13,3,k}$

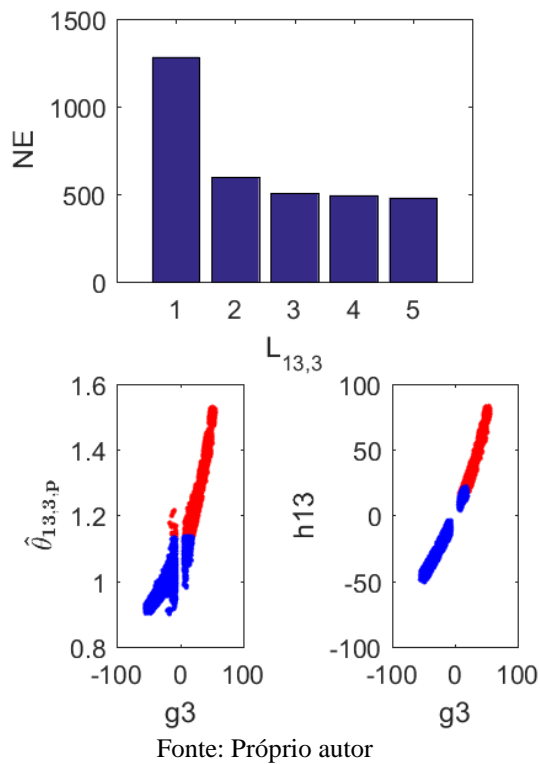


Figura 62 – Análise de $\hat{\theta}_{13,10,p}$ para seleção de $\theta_{13,10,k}$

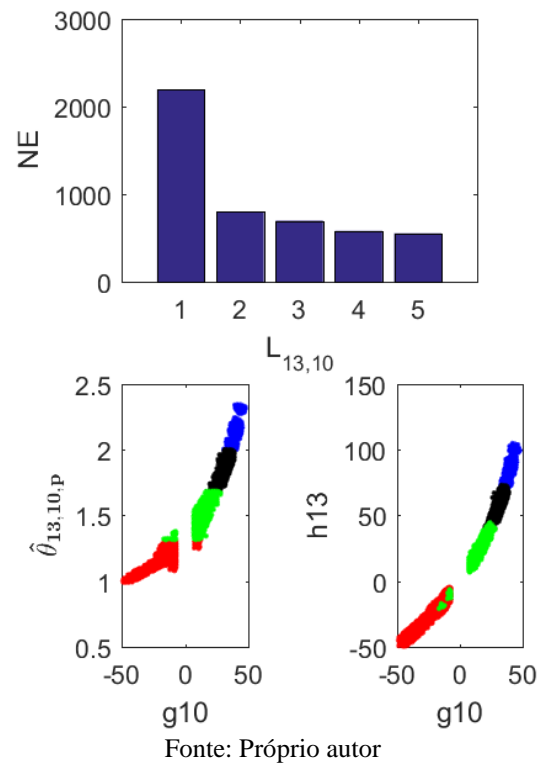
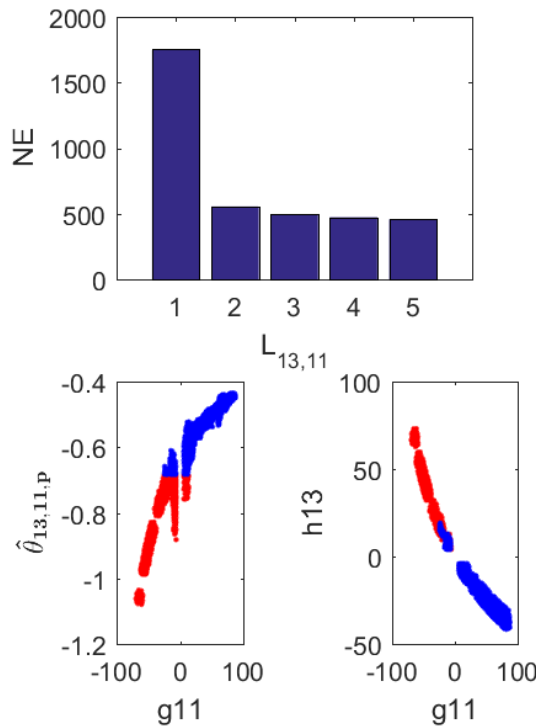
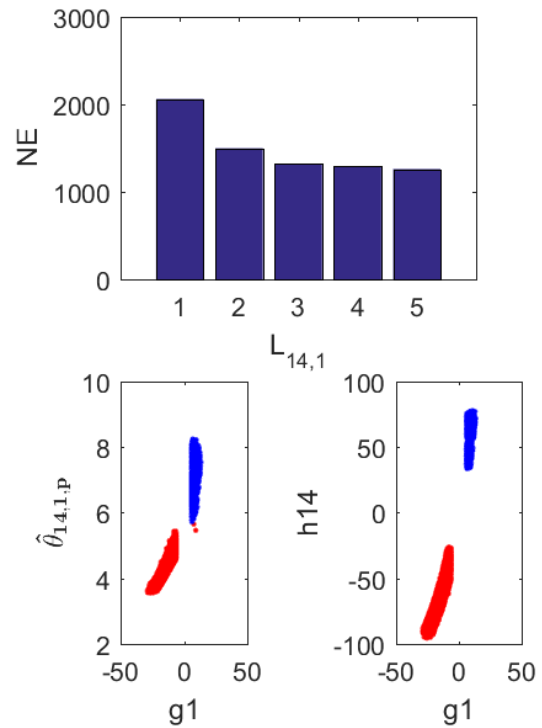


Figura 63 – Análise de $\hat{\theta}_{13,11,p}$ para seleção de $\theta_{13,11,k}$



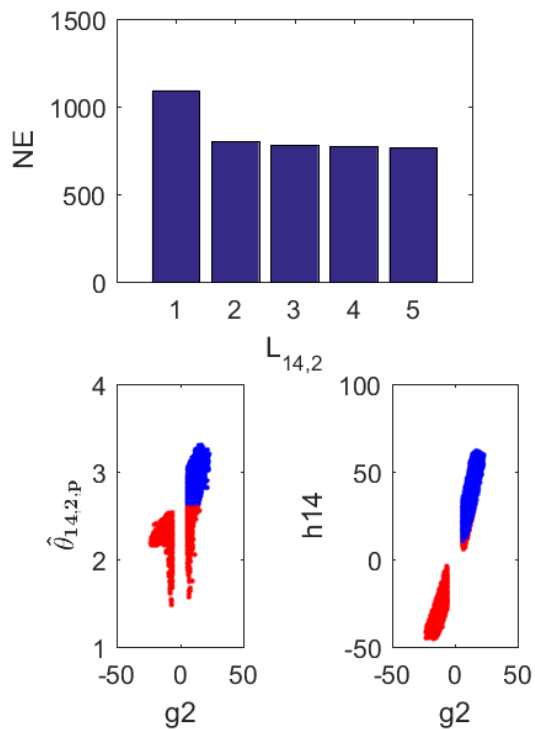
Fonte: Próprio autor

Figura 64 – Análise de $\hat{\theta}_{14,1,p}$ para seleção de $\theta_{14,1,k}$



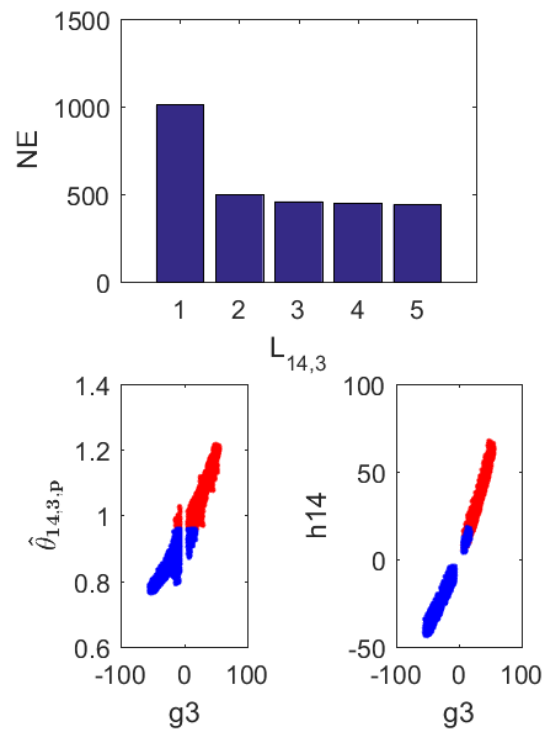
Fonte: Próprio autor

Figura 65 – Análise de $\hat{\theta}_{14,2,p}$ para seleção de $\theta_{14,2,k}$

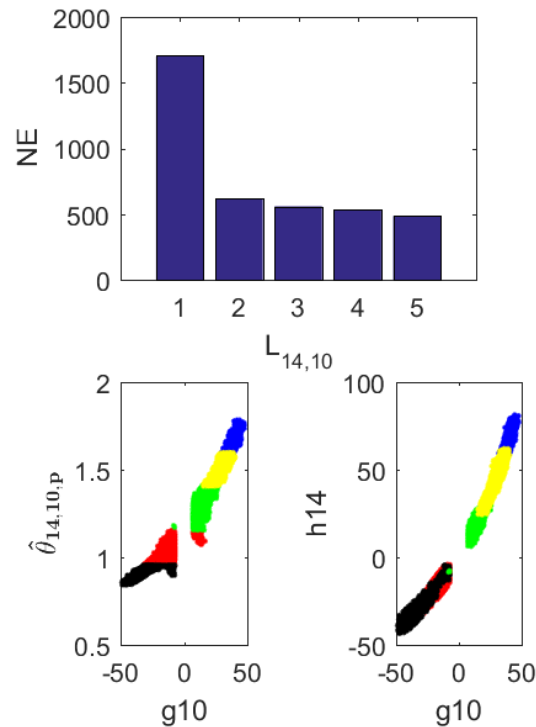


Fonte: Próprio autor

Figura 66 – Análise de $\hat{\theta}_{14,3,p}$ para seleção de $\theta_{14,3,k}$



Fonte: Próprio autor

Figura 67 – Análise de $\hat{\theta}_{14,10,p}$ para seleção de $\theta_{14,10,k}$ 

Fonte: Próprio autor

Na Figura 49, pode-se observar que a norma dos resíduos sofreu apenas pequenas diminuições com o aumento de $L_{i,j}$. Observa-se, ainda, que os elementos de $\hat{\theta}_{5,1,p}$ se concentram entre 1.1 e 1.2 ao longo do domínio de \mathbf{g}_1 , corroborando o que pode ser visto no gráfico de dispersão entre \mathbf{h}_5 e \mathbf{g}_1 . Portanto, o modelo único foi selecionado, ou seja, $L_{5,1} = 1$. Características similares podem ser observadas nas Figuras 50, 51, 52, 57 e 58, logo, $L_{5,2}, L_{6,1}, L_{6,2}, L_{9,2}$ e $L_{12,2}$ também serão 1, com isto os vetores de coeficientes correspondentes $\theta_{i,j,k}$ conterão apenas um elemento referente ao seu modelo linear único.

Por outro lado, na Figura 53 observa-se que houve uma grande diminuição da norma dos resíduos na mudança de $L_{8,1}$ de 1 para 2, e pequenas diminuições para $L_{8,1}$ maiores que 2. Na mesma figura pode ser visto que os valores de $\hat{\theta}_{8,1,p}$ variam entre aproximadamente 4 e 8 ao longo do domínio de \mathbf{g}_1 . Através da comparação entre as normas dos resíduos calculadas, selecionou-se $L_{8,1}=2$. Logo, $\theta_{8,1,k}$ conterá dois elementos. Características similares podem ser observadas nas Figuras 54, 55, 59, 60, 61, 64, 65 e 66. Logo, $\theta_{8,2,k}, \theta_{8,3,k}, \theta_{13,1,k}, \theta_{13,2,k}, \theta_{13,3,k}, \theta_{14,1,k}, \theta_{14,2,k}$ e $\theta_{14,3,k}$ também conterão dois elementos referentes aos seus modelos lineares calculados.

As Figuras 62, 63 e 67 mostram o resultado da análise para a seleção de $\theta_{13,10,k}, \theta_{13,11,k}$ e $\theta_{14,10,k}$. Observa-se que $\theta_{13,10,k}$ foi selecionado com 4 elementos. Já $\theta_{13,11,k}$ e $\theta_{14,10,k}$ foram selecionados com 5 elementos. Nesses casos, pode-se observar nos gráficos de dispersão entre as respectivas \mathbf{h}_i e \mathbf{g}_j uma não-linearidade mais aparente.

4.3.4 TREINAMENTO DOS CLASSIFICADORES

Para cada vetor de coeficientes $\theta_{i,j,k}$ obtido com mais de um elemento, um classificador kNN foi treinado considerando 4 (quatro) vizinhos mais próximos. O treinamento dos classificadores foi realizado com 70% das amostras de cada g_j relativas a $\theta_{i,j,k}$, rotuladas através de *k-means*. Seu desempenho foi medido através da acurácia utilizando o restante das amostras de g_j relativas a $\theta_{i,j,k}$. Na Tabela 5 apresenta a acurácia dos classificadores.

Tabela 5 – Acurácia dos classificadores

	g_1	g_2	g_3	g_{10}	g_{11}
h_5	-	-	-	-	-
h_6	-	-	-	-	-
h_8	1	0,9859	0,9919	0,9849	-
h_9	-	-	-	-	-
h_{12}	-	-	-	-	-
h_{13}	1	0,988	0,994	0,9956	0,9868
h_{14}	1	0,986	0,9935	0,9852	-

Fonte: Próprio autor

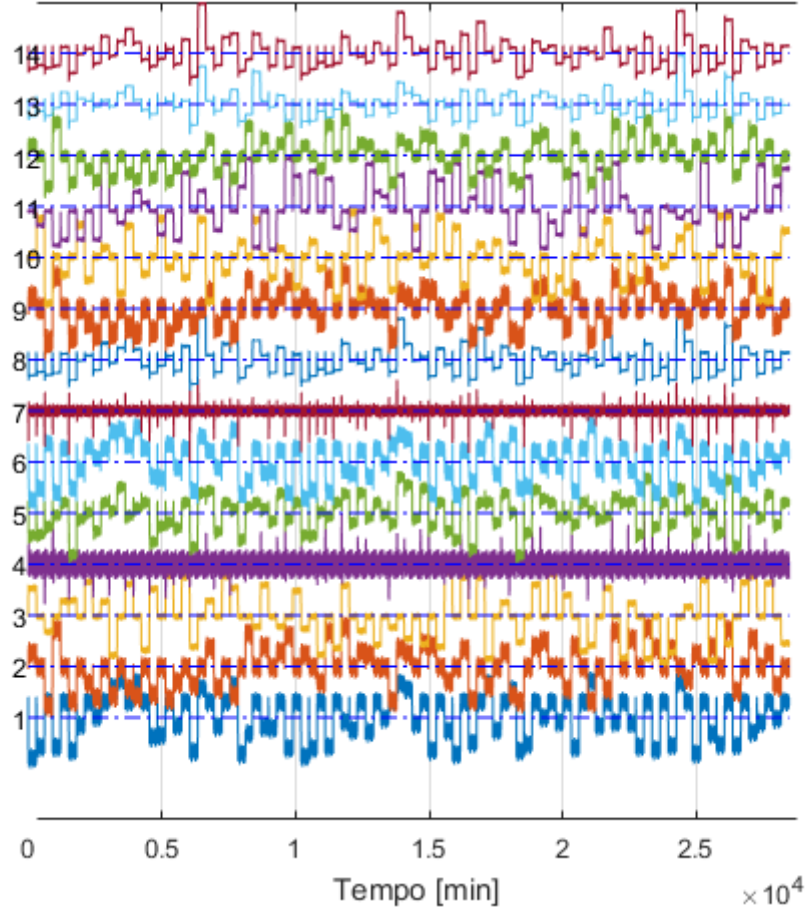
4.3.5 OBTENÇÃO DO MODELO PCA

Para obter o modelo do PCA, novos dados do CSTR foram obtidos contendo as mesmas faixas de operação da Tabela 3 e foram, então, normalizados. A Figura 68 apresenta esses dados.

Os classificadores foram aplicados a esses dados selecionando os coeficientes $c_{i,j}$ a serem utilizados para cada amostra, e, com isso, a matriz de resíduos foi obtida. As variáveis de resíduos dessa matriz estão apresentadas na Figura 69.

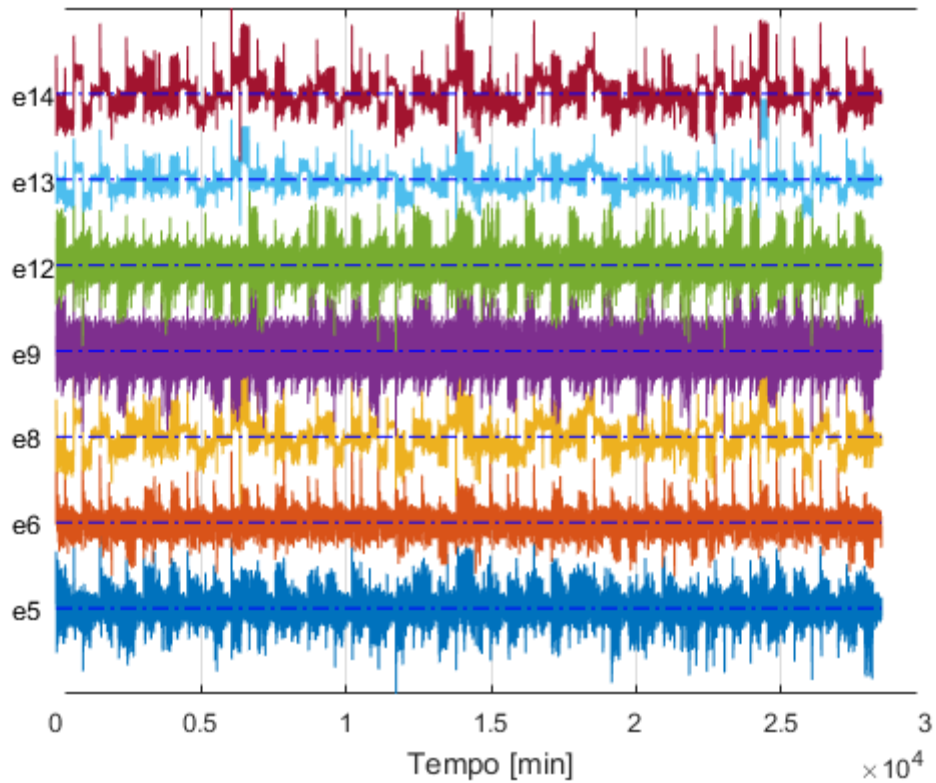
O modelo PCA foi obtido de acordo com a seção 2.3. O número de componentes principais foi selecionado via o método da percentagem de variância cumulativa, utilizando um limite de 90% da variância. O nível de significância foi definido como 5%. O limite estatístico da estatística combinada foi obtido de acordo com a seção 2.4. A Figura 70 mostra a estatística combinada calculada para os dados utilizados para obter o modelo PCA e seu limiar obtido.

Figura 68 – Dados do CSTR para obtenção do modelo PCA



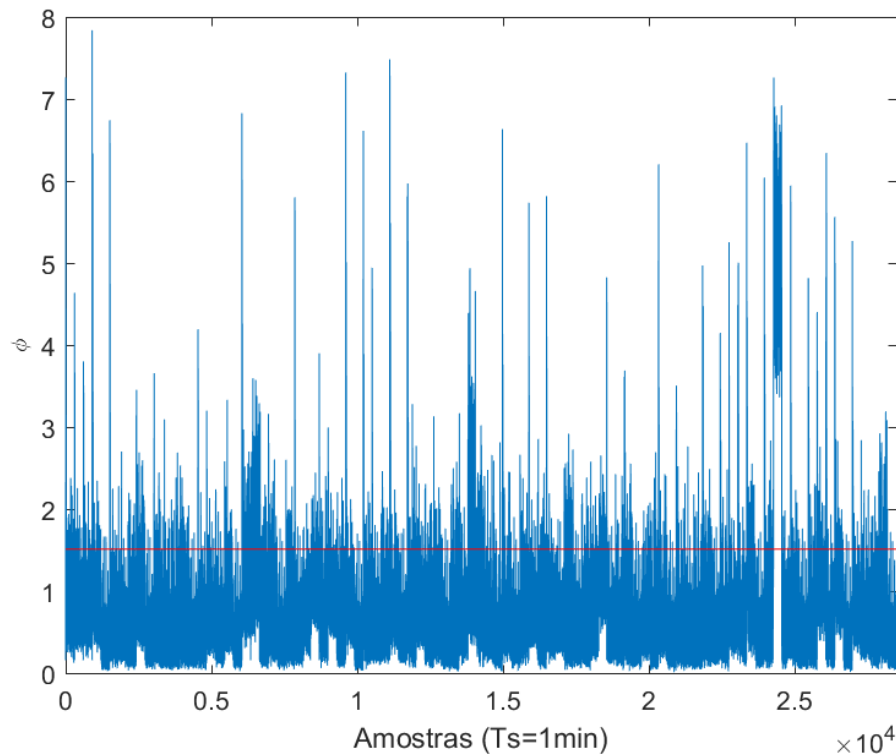
Fonte: Próprio autor

Figura 69 – Resíduos utilizados para obter o modelo PCA



Fonte: Próprio autor

Figura 70 – Estatística combinada calculada para os dados utilizados para obter o modelo PCA



Fonte: Próprio autor

4.4 ANÁLISE DOS RESULTADOS

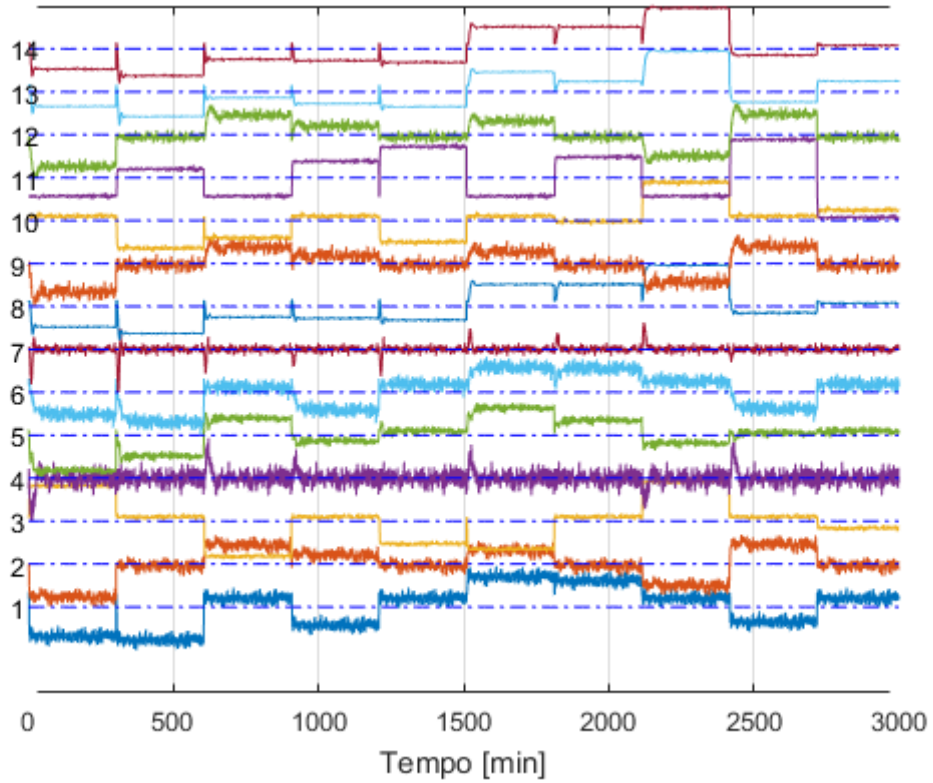
Uma simulação foi feita no CSTR abrangendo 10 regiões de operação, com variações a cada 300 minutos de três VE em valores aleatórios dentro das suas faixas de operação. Objetivou-se, com essa simulação, verificar que, com a aplicação do método proposto, variações nas regiões de operação geram uma taxa FP consideravelmente similar ao nível de significância definido. A taxa de FP é calculada de acordo com a equação (4.3)

$$FP = \frac{NL}{NT} \quad (4.3)$$

sendo NL o número de amostras acima do limite estatístico sem a ocorrência de falha e NT o número de amostras total sem a ocorrência de falha.

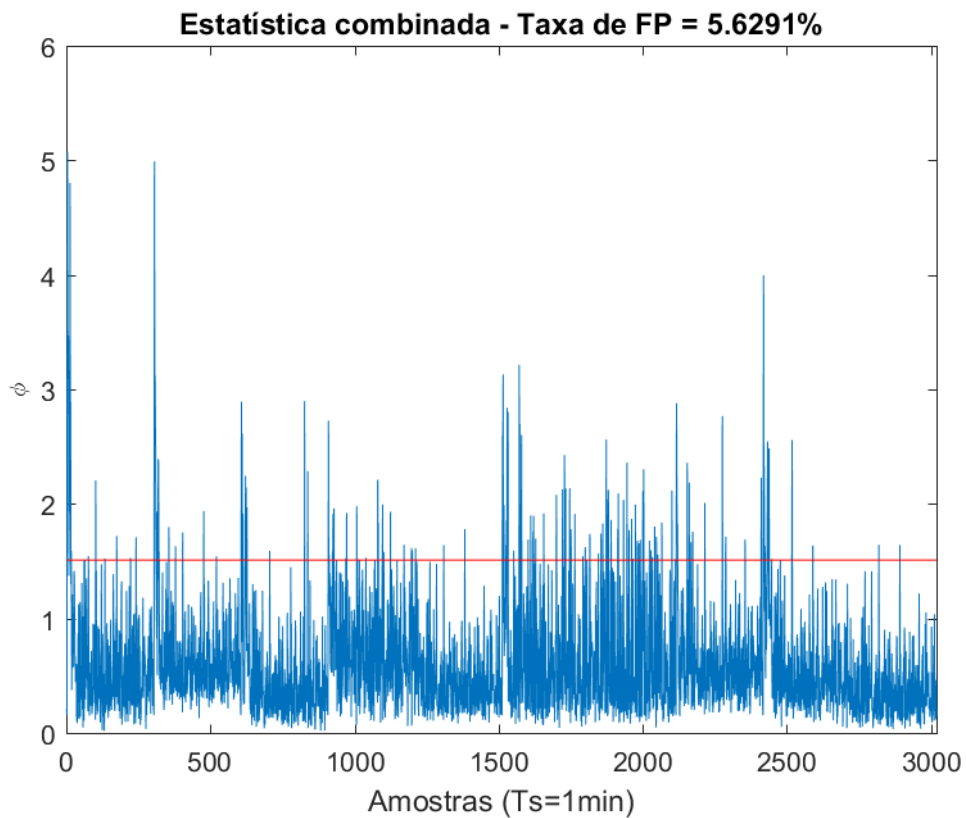
A Figura 71 mostra as variáveis do CSTR no tempo, e a Figura 72 mostra a estatística combinada calculada para essa simulação.

Figura 71 - Variáveis do CSTR para a simulação contendo 10 regiões de operação



Fonte: Próprio autor

Figura 72 - Estatística combinada calculada para a simulação contendo 10 regiões de operação



Fonte: Próprio autor

Pode-se observar que a taxa de FP ficou próxima ao nível de significância definido, que foi 5%.

Em seguida, os desempenhos de MM e M1 foram avaliados através da comparação das taxas de FN calculadas de acordo com

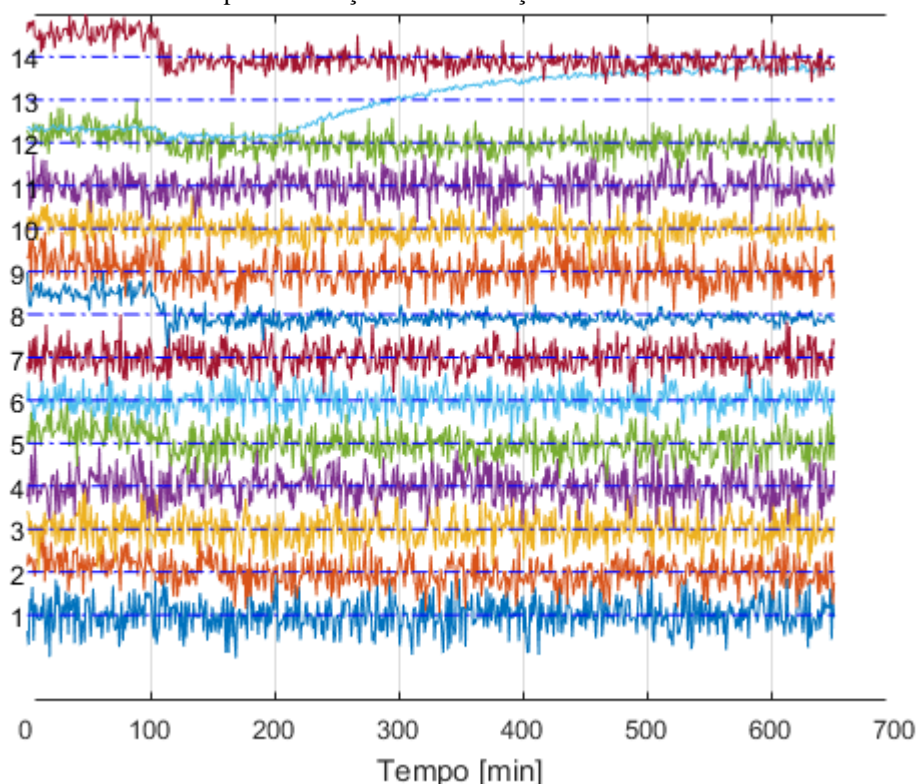
$$FN = \frac{NPF - NPD}{NPF} \quad (4.4)$$

em que NPF se refere ao número de amostras após a ocorrência da falha e NPD ao número de amostras após a detecção da falha. Uma baixa taxa de FN indica sensibilidade do método, ou seja, capacidade do método de detectar falhas incipientes.

Em seguida, as falhas apresentadas na Tabela 2 foram avaliadas. Suas intensidades foram aumentadas gradualmente durante as simulações, uma vez que afetam consideravelmente o desempenho dos métodos de detecção.

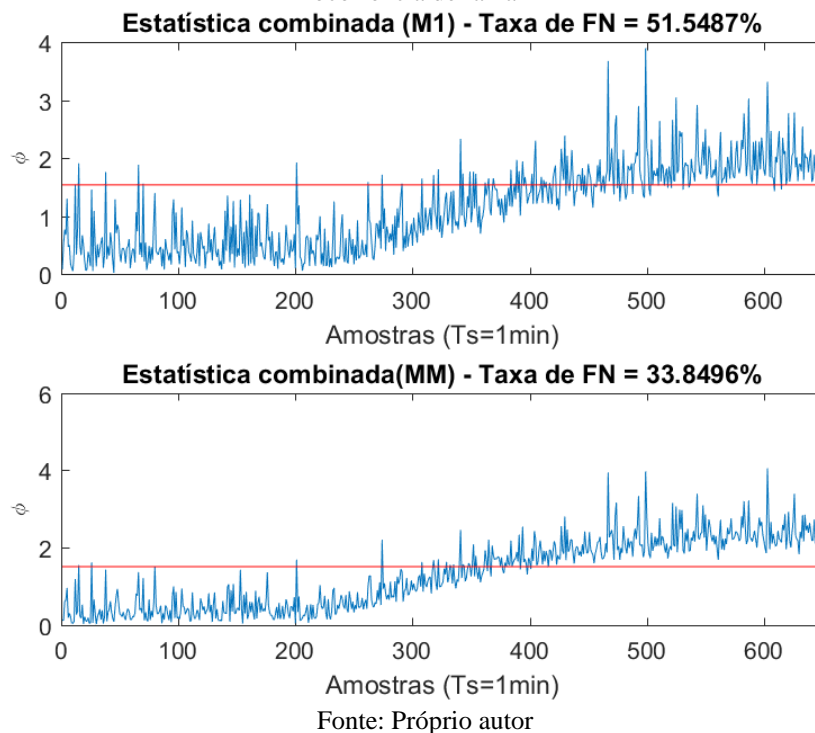
A Figura 73 mostra as variáveis do CSTR no tempo de uma das simulações realizadas, na qual ocorre mudança de região de operação na amostra 100 através das variáveis externas 2 e 10, e falha 4 na amostra 200. A Figura 74 mostra a estatística combinada calculada via MM e M1. Observa-se que MM obteve uma taxa de FN aproximadamente 20% menor que M1.

Figura 73 - Variáveis do CSTR para simulação com mudança das VE 2 e 10 e com ocorrência de falha 4



Fonte: Próprio autor

Figura 74 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 10 e com ocorrência de falha 4



A Figura 75 mostra as variáveis do CSTR no tempo de outra simulação que foi realizada simulando mudança de região de operação na amostra 100 através das variáveis externas 2 e 3, e com ocorrência da falha 5 na amostra 200. A Figura 76 mostra a estatística combinada calculada via os métodos MM e M1. Observa-se que o método MM obteve uma taxa de FN aproximadamente 11% menor que o método M1.

Figura 75 - Variáveis do CSTR para simulação com mudança das VE 2 e 3 e com ocorrência de falha 5

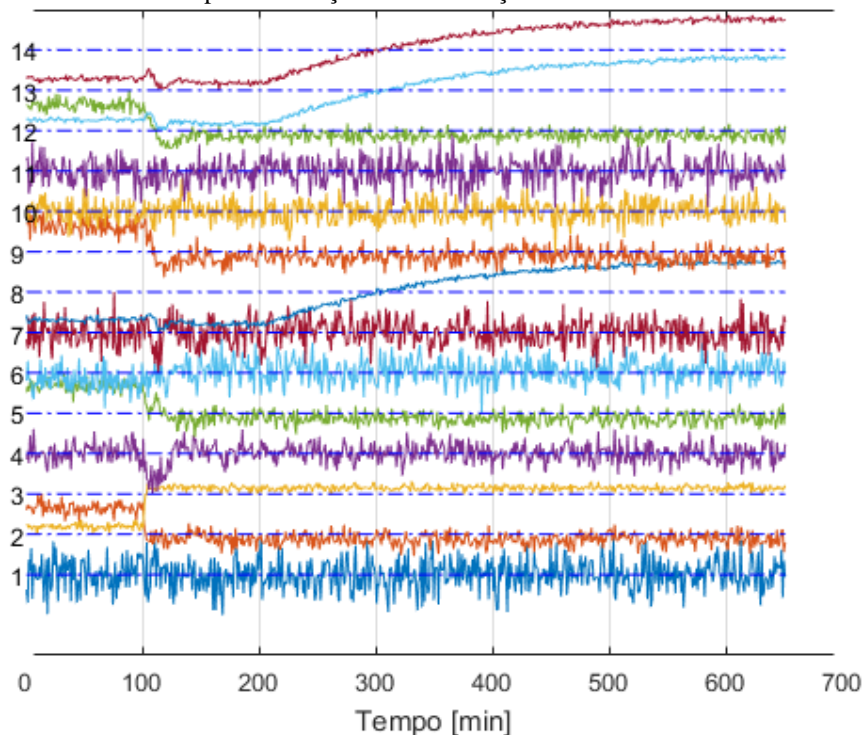
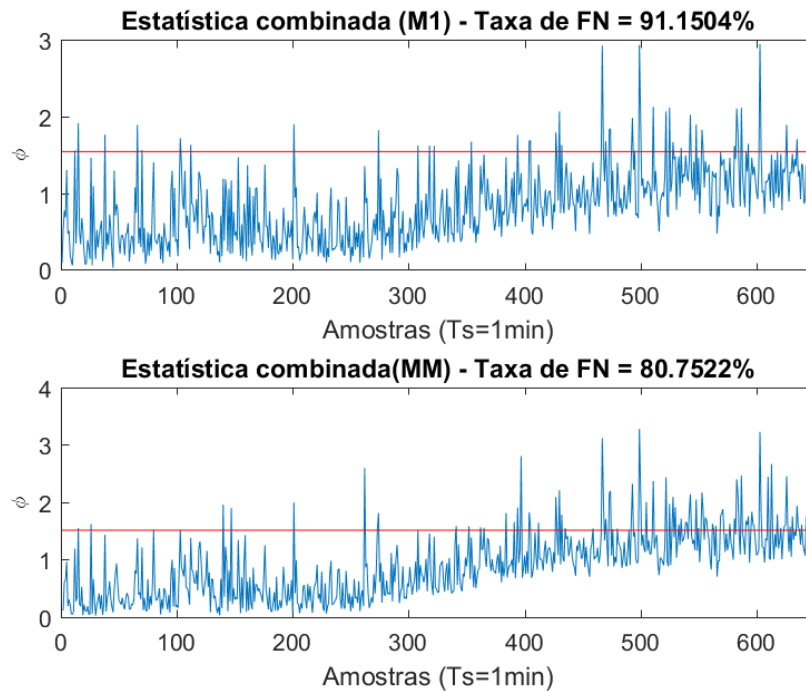


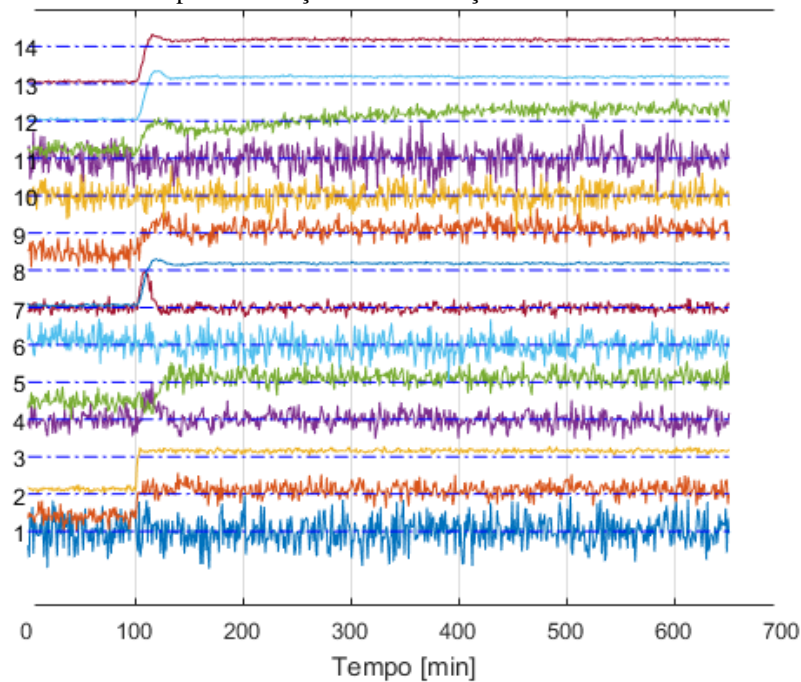
Figura 76 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 3 e com ocorrência de falha 5



Fonte: Próprio autor

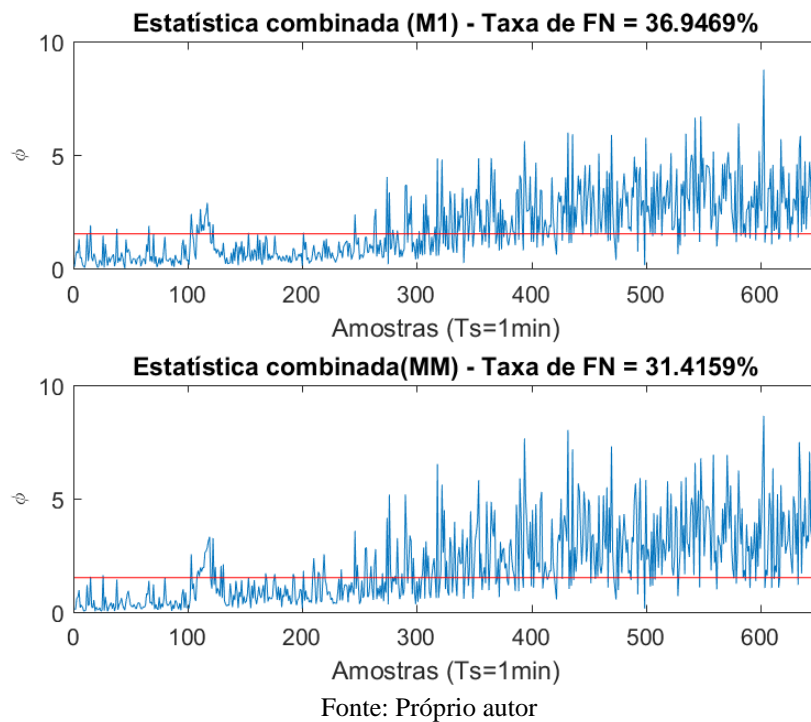
A Figura 77 apresenta as variáveis do CSTR no tempo de uma simulação na qual ocorre a mudança da região de operação na amostra 100 através das variáveis externas 2 e 3, e a ocorrência da falha 7 na amostra 200. A Figura 78 mostra a estatística combinada calculada via MM e M1. Observa-se que MM obteve uma taxa de FN aproximadamente 5% menor que M1.

Figura 77 - Variáveis do CSTR para simulação com mudança das VE 2 e 3 e com ocorrência de falha 7



Fonte: Próprio autor

Figura 78 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 3 e com ocorrência de falha 7



A Figura 79 mostra as variáveis do CSTR no tempo de mais uma simulação, na qual ocorre uma mudança das variáveis externas 2 e 10 na amostra 100, e a falha 2 na amostra 200. A Figura 80 apresenta a estatística combinada calculada via MM e M1. Observa-se que as taxas de FN obtidas via MM e M1 foram aproximadamente iguais.

Figura 79 - Variáveis do CSTR para simulação com mudança das VE 2 e 10 e falha 2

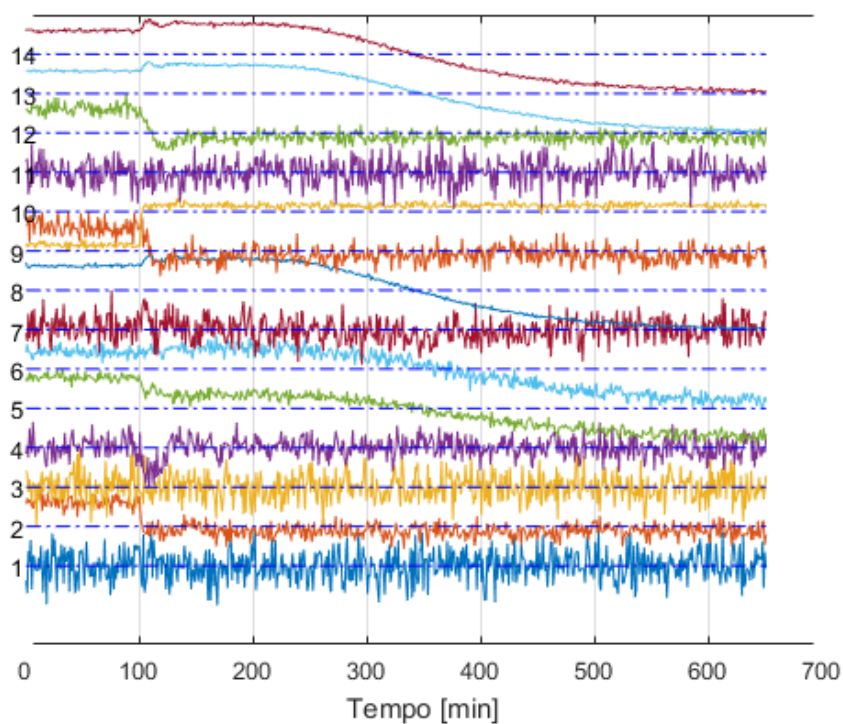
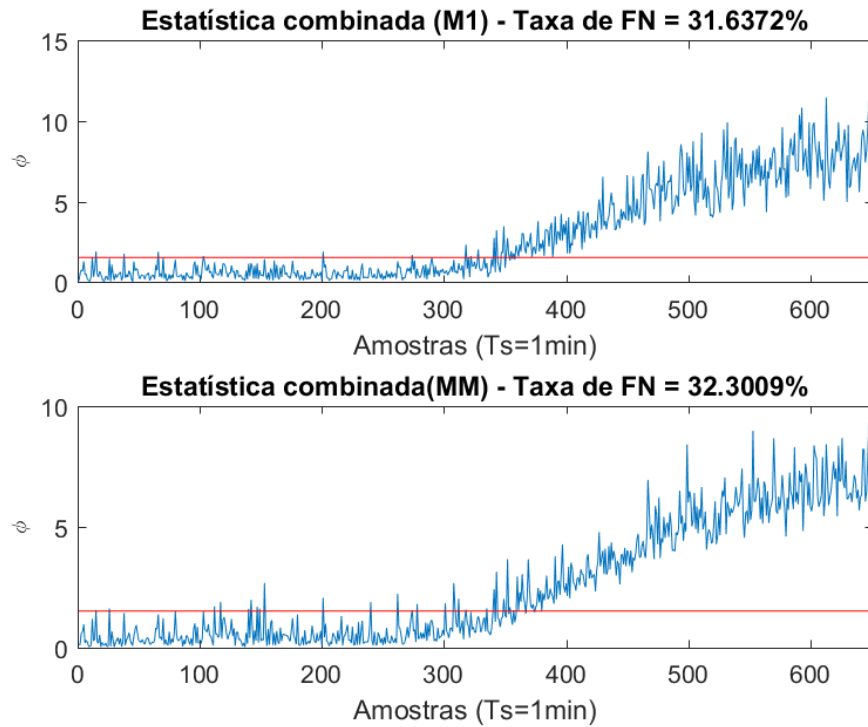


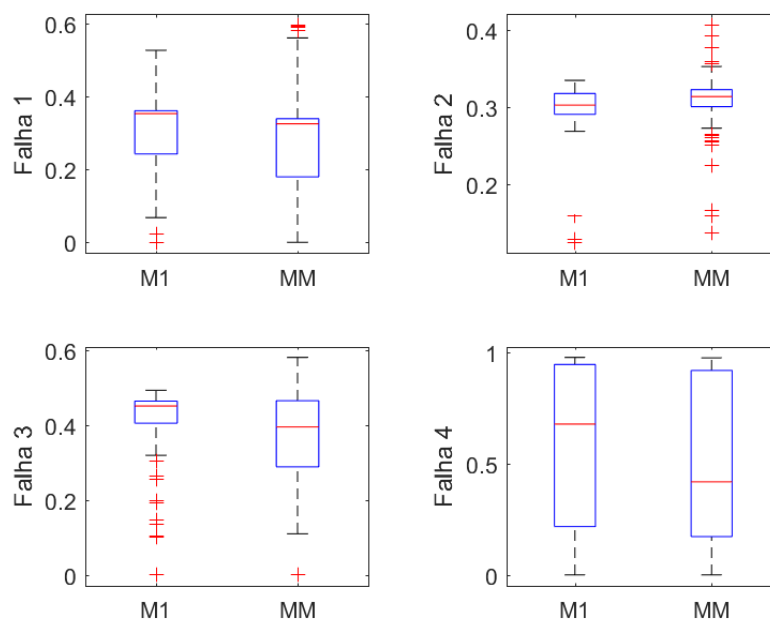
Figura 80 - Estatística combinada de M1 e MM para simulação com mudança das VE 2 e 10 e falha 2



Fonte: Próprio autor

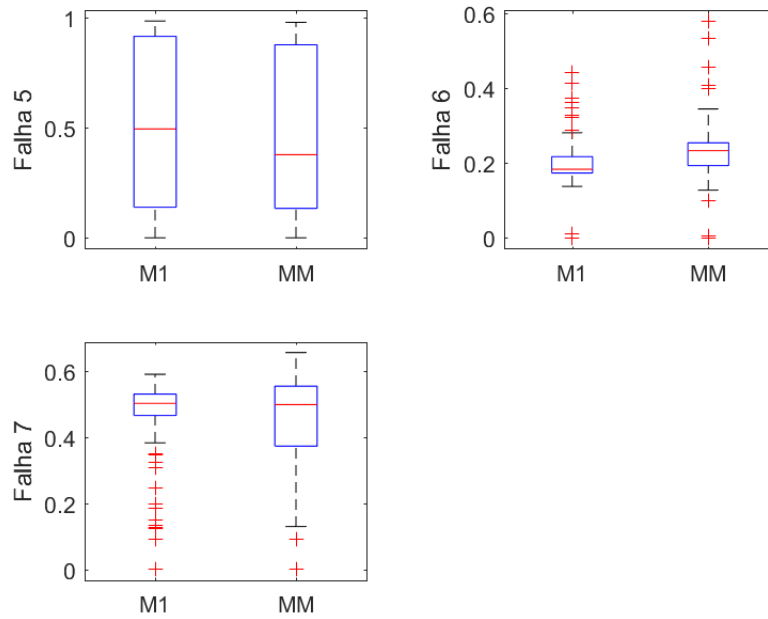
Para avaliar a independência dos métodos em relação à região de operação, cada falha foi simulada a partir de 100 diferentes pontos de operação escolhidos aleatoriamente, e suas taxas de FN foram medidas. As Figuras 81 e 82 apresentam os *box plots* das taxas de FN para essas simulações via M1 e MM. A Figura 83 apresenta a médias das taxas de FN obtidas via M1 e MM para todas as falhas testadas.

Figura 81 - Box plots das taxas de FN para as falhas 1, 2, 3 e 4



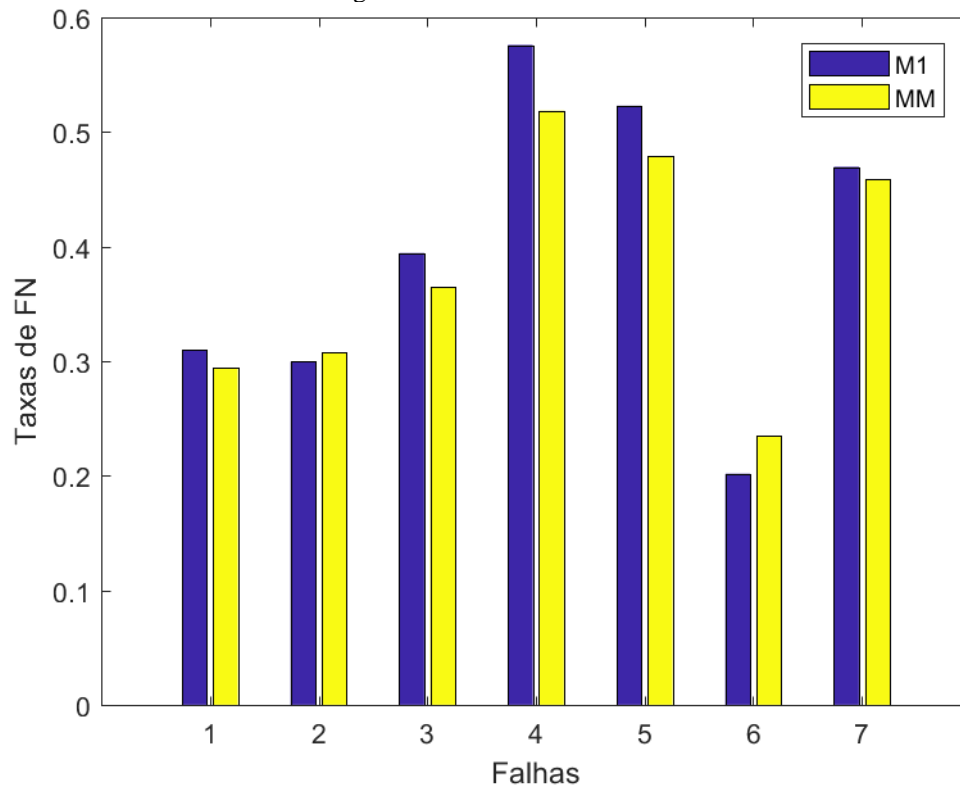
Fonte: Próprio autor

Figura 82 - Box plots das taxas de FN para as falhas 5, 6 e 7



Fonte: Próprio autor

Figura 83 - Média das taxas de FN



Fonte: Próprio autor

Observa-se nas Figuras 81 e 82 que, para as falhas 3 e 7, as dispersões das taxas de FN obtidas via MM foram maiores que as dispersões obtidas via M1. Para as outras falhas, as dispersões de MM e M1 foram similares.

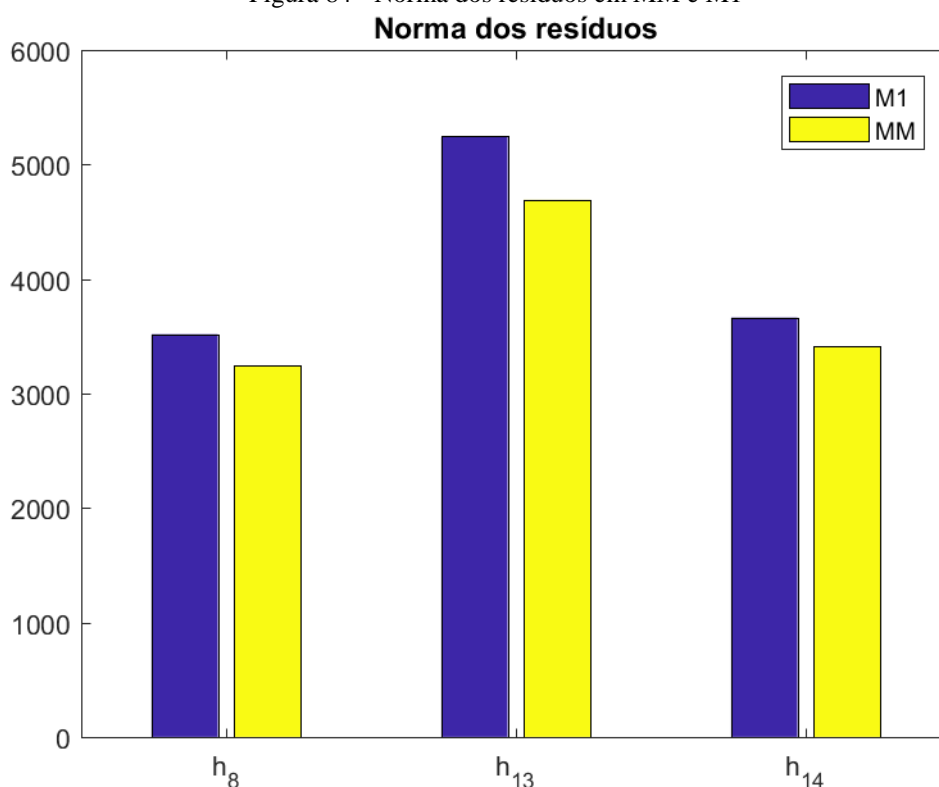
Na Figura 83, observa-se que as médias das taxas de FN foram menores para MM nas falhas 1, 3, 4, 5 e 7, apresentando uma melhora mais acentuada para as falhas 4 e 5. A falha 2 apresentou a média da taxa de FN aproximadamente igual para MM e M1. A falha 6 apresentou uma média ligeiramente maior para MM.

Portanto, vê-se que, para a maioria das falhas, MM apresentou melhorias em relação à média da taxa de FN, porém, a dispersão da taxa de FN foi prejudicada para algumas falhas.

Os resultados da aplicação apresentaram uma melhoria no desempenho de MM, que se mostrou mais sensível a falhas que o método M1. Esse resultado é consequência da característica dos resíduos obtidos via MM, pois quando os resíduos são menores, a magnitude das falhas deverá ser menor para que a mesma seja detectada.

Para realizar uma comparação dos resíduos obtidos via MM e M1, suas normas foram obtidas para as VP do CSTR que fizeram uso de múltiplos modelos lineares, ou seja, h_8 , h_{13} e h_{14} . Esses dados foram obtidos contando com 100 regiões de operação diferentes, variações a cada 300 minutos de três VE em valores aleatórios dentro das suas faixas de operação. As normas obtidas estão apresentadas na Figura 84. Pode ser observado que a norma dos resíduos obtidos via MM realmente foi inferior à obtida via M1, corroborando, então, o que foi discutido.

Figura 84 - Norma dos resíduos em MM e M1



Fonte: Próprio autor

5 CONCLUSÕES E TRABALHOS FUTUROS

5.1 CONCLUSÕES

A monitoração de processos industriais, quando feita de forma eficiente, traz diversas vantagens para as indústrias, tornando-as economicamente mais eficientes. O desenvolvimento de sensores inteligentes e o aumento da capacidade de armazenamento de dados proporcionou o desenvolvimento de várias técnicas baseadas em dados, que vêm sendo estudadas com o objetivo de alcançar resultados satisfatórios visando a detecção e diagnósticos de falhas em processos industriais.

O trabalho aqui apresentado contribui para o problema de detecção de falhas em múltiplas regiões de operação através da investigação da possibilidade da utilização de múltiplos modelos lineares aplicados à análise externa. Objetiva-se, com isso, tratar a presença de não linearidades no processo sem utilizar modelos não lineares, uma vez que estes apresentam a necessidade de definições complexas de funções e parâmetros que influenciam consideravelmente na qualidade do modelo final.

A pesquisa desenvolvida para este trabalho propõe uma metodologia para detecção da necessidade do uso de múltiplos modelos lineares e para a construção dos mesmos. Através da aplicação do RLS em dados com diferentes regiões de operação, foi possível identificar a existência de variações nos coeficientes dos modelos que relacionam as variáveis externas e principais. Esses coeficientes foram agrupados via *k-means* variando o número de *clusters* para gerar diferentes números de modelos, o que proporcionou – através da comparação dos resíduos gerados – selecionar a quantidade de modelos lineares que reduzissem tais resíduos.

Classificadores kNN foram treinados para associar as variáveis externas já rotuladas durante a fase de agrupamento aos grupos de modelos, e demonstraram uma acurácia significativa. Então, o modelo PCA foi calculado através das variáveis de resíduos obtidas com a aplicação de modelos selecionados via classificadores.

Para a aplicação dos múltiplos modelos na monitoração do processo, foi proposto um método que consiste na obtenção dos resíduos e na monitoração através do cálculo da estatística combinada.

O desempenho da metodologia proposta foi avaliado através de sua aplicação na supervisão do processo industrial simulado de um CSTR e apresentou taxas de FP próximas ao nível de significância definido. As taxas de FN obtidas através da aplicação da monitoração do processo via MM foram comparadas às taxas obtidas através da aplicação da metodologia de análise externa via M1. Foi verificado que, para a maioria das falhas, MM apresentou melhorias em relação à média da taxa de FN. Porém, as dispersões das taxas de FN em relação às regiões de operação que precederam a falha foram prejudicadas em alguns casos.

Portanto, foi verificado que os resultados da aplicação do método proposto, quando comparados aos resultados obtidos via M1, apresentaram uma melhoria na sensibilidade de MM para a maioria das falhas testadas como consequência da menor magnitude dos resíduos obtidos.

5.2 TRABALHOS FUTUROS

Dentre as diversas possibilidades de estender e aperfeiçoar a metodologia proposta, pontuam-se os seguintes trabalhos futuros:

1. Analisar uma adequação da metodologia que permita o uso de VE que apresentem comportamento contínuo e comportamentos de *setpoints* (que variam entre regiões de valores constantes).
2. Analisar a possibilidade de melhorias nos resultados com a aplicação de outros métodos de *clusterização*.
3. Analisar uma adequação da metodologia que permita sua aplicação na análise externa dinâmica.
4. Investigar o fato da dispersão das taxas de FN para algumas falhas terem sido maiores com a utilização de MM – quando comparadas às taxas para M1 – e buscar possíveis soluções para melhorar esse aspecto dos resultados.
5. Aplicação da abordagem proposta por este trabalho à supervisão de dados de operações reais.

REFERÊNCIAS

- ALLEN, D. M. The relationship between variable selection and data augmentation and a method for prediction. **Technometrics**, **16(1)**, p. 125-127, 1974.
- BONETT, D. G.; WRIGHT, T. A. Sample size requirements for estimating Pearson, Kendall and Spearman correlations. **Psychometrika**, n. 65(1), p. 23-28, 2000.
- CAO, S.; RHINEHART, R. R. An efficient method for on-line identification of steady state. **Journal of Process Control**, **5(6)**, p. 363-374, 1995.
- CAO, S.; RHINEHART, R. R. Critical values for a steady-state identifier. **Journal of Process Control**, **7(2)**, p. 149-152, 1997.
- CHALAM, V. V. **Adaptive control systems: Techniques and applications**. [S.l.]: Marcel Dekker, Inc., 1987.
- CHAMBERS, J. M. **Graphical methods for data analysis**. [S.l.]: CRC Press, 2018.
- CHIANG, L. H.; RUSSELL, E. L.; BRAATZ, R. D. **Fault detection and diagnosis in industrial systems**. [S.l.]: Springer Science & Business Media, 2001.
- COHEN, J.; COHEN, P.; WEST, S.G.; AIKEN, L.S. **Applied multiple regression/correlation analysis for the behavioral sciences**. [S.l.]: Routledge, 2013.
- DERKSEN, S.; KESELMAN, H. J. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. **British Journal of Mathematical and Statistical Psychology**, **45(2)**, p. 265-282, 1992.
- FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. A. D. Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r), 2009.
- FINCH, F. E. Automated fault diagnosis of chemical process plants using model-based reasoning. (**Doctoral dissertation, Massachusetts Institute of Technology**), 1989.
- GARCIA-ALVAREZ, D.; FUENTE, M. J.; SAINZ, G. I. Fault detection and isolation in transient states using principal component analysis. **Journal of Process Control** **22.3**, p. 551-563, 2012.
- GE, Z.; SONG, Z. Online monitoring of nonlinear multiple mode processes based on adaptive local model approach. **Control Engineering Practice**, **16**, p. 1427-1437, 2008.
- GE, Z.; YANG, C.; SONG, Z.; WANG, H. Robust online monitoring for multimode processes based on nonlinear external analysis. **Industrial & Engineering Chemistry Research** **47(14)**, p. 4775-4783., 2008.

GE, Z.; SONG, Z.; KANO, M. External analysis-based regression model for robust soft sensing of multimode chemical processes. **AIChE Journal** **60.1**, p. 136-147., 2014.

GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection. **Journal of machine learning research**, **3(Mar)**, p. 1157-1182, 2003.

HAMERLY, G.; ELKAN, C. Learning the k in k-means. **Em Advances in neural information processing systems** , p. 281-288, 2004.

HWANG, D. H.; HAN, C. Real-time monitoring for a process with multiple operating modes. **Control Engineering Practice**, v. 7, n. 7, p. 891-902, 1999.

JAIN, A. K. Data clustering: 50 years beyond K-mean. **Pattern recognition letters**, **31(8)**, p. 651-666, 2010.

JOE QIN, S. Statistical process monitoring: basics and beyond. **Journal of chemometrics** **17.8-9**, p. 480-502, 2003.

KANO, M.; HASEBE, S.; HASHIMOTO, I.; OHNO, H. Evolution of multivariate statistical processcontrol: application of independent component analysis and external analysis. **Comput Chem Eng.**, p. 1157–1166, 2004.

KAYNAK, O.; HOYER, H. Predictive control of a robotic arm. **In Robot Control 1988 (Syroco'88)**, p. 213-218, 1989.

KRUGER, U.; XIE, L. **Advances in statistical monitoring of complex multivariate processes: with applications in industrial process control**. [S.l.]: John Wiley & Sons, 2012.

LANDAU, I.D.; LOZANO, R.; M'SAAD, M.; KARIMI, A. **Adaptive control: algorithms, analysis and applications**. [S.l.]: Springer Science & Business Media, 2011.

LEE, J. E. A. Recent advances and trends in predictive manufacturing systems in big data environment. **Manufacturing Letters** **1.1**, p. 38-41, 2013.

LJUNG, L. **System Identification - Theory for the User**. 2. ed. [S.l.]: Prentice-Hall, Engle- wood Cliffs, NJ, 1987.

LJUNG, L. System identification. **Em Signal analysis and prediction**, Birkhäuser, Boston, MA, 1998. 163-173.

MA, H.; HU, Y.; SHI, H. A novel local neighborhood standardization strategy and its application in fault detection of multimode processes. **Chemometrics and Intelligent Laboratory Systems** **118**, p. 287-300, 2012.

MAESTRI, M.; FARALL, A.; GROISMAN, P.; CASSANELLO, M.; HOROWITZ, G. A robust clustering method for detection of abnormal situations in a process with multiple steady-state operation modes. **Computers & chemical engineering**, **34(2)**, p. 223-231, 2010.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C. Adaptive filtering: An integrated autoregressive/moving average filter for time series forecasting. **Journal of the Operational Research Society**, **28(2)**, p. 425-437, 1977.

MEJÍA, R.I.G.; COELHO, A.A.R.; FARENZENA, M.; TRIERWEILER, J.O.; NUNES, G.C. Auditoria de Malhas de Controle em Processos Industriais. **Tutoriais do XVII Congresso Brasileiro de Automática CBA. Juiz de Fora, MG, Brasil: Book Editora, ISBN** , p. 978-85, 2008.

NIKAM, S. S. A comparative study of classification techniques in data mining algorithms. **Oriental Journal of Computer Science and Technology**, **8(1)**, p. 13-19, 2015.

NOSE-FILHO, K.; LOTUFO, A. D. P.; MINUSSI, C. R. Preprocessing data for short-term load forecasting with a general regression neural network and a moving average filter. **In PowerTech, 2011 IEEE Trondheim. IEEE**, p. 1-7, 2011.

OAKLAND, J. S. **Statistical process control**. [S.l.]: Routledge, 2007.

PATAN, K. **Artificial neural networks for the modelling and fault diagnosis of technical processes**. [S.l.]: Springer, 2008.

PETERSON, L. E. K-nearest neighbor. *Scholarpedia*, 4(2), p. 1883, 2009.

RAY, S.; TURI, R. H. Determination of number of clusters in k-means clustering and application in colour image segmentation. **Proceedings of the 4th international conference on advances in pattern recognition and digital techniques**, Dezembro 1999. 137-143.

SAEED, M. O. B.; ZERGUINE, A.; ZUMMO, S. A. Blind distributed estimation algorithms for adaptive networks. **EURASIP Journal on Advances in Signal Processing**, **2014(1)**, p. 136, 2014.

SATOYAMA, Y.; FUJIWARA, K.; KANO, M. Variable Elimination-Based Contribution for Accurate Fault Identification. **IFAC-PapersOnLine**, **49(7)**, p. 383-388, 2016.

SHLENS, J. A tutorial on principal component analysis. **arXiv preprint arXiv:1404.1100**, 2014.

SONG, Y.; HUANG, J.; ZHOU, D.; ZHA, H.; GILES, C.L. **Iknn**: Informative k-nearest neighbor pattern classification. Em European Conference on Principles of Data Mining and Knowledge Discovery. Berlin, Heidelberg: Springer. 2007. p. 248-264.

TOU, J. T.; GONZALEZ, R. C. **Pattern Recognition Principles**, , Massachusetts: Addison-Wesley, 1974.

VAHIDI, A.; STEFANOPOULOU, A.; PENG, H. Recursive least squares with forgetting for online estimation of vehicle mass and road grade: theory and experiments. **Vehicle System Dynamics**, **43(1)**, p. 31-55, 2005.

WESTERHUIS, J. A.; GURDEN, S. P.; SMILDE, A. K. Generalized contribution plots in multivariate statistical process monitoring. **Chemometrics and intelligent laboratory systems** **51.1**, p. 95-114, 2000.

WOLD, S.; ESBENSEN, K.; GELADI, P. Principal component analysis. **Chemometrics and intelligent laboratory systems**, n. 2(1-3), p. 37-52, 1987.

YONG, G.; XIN, W.; ZHENLEI, W. Fault detection for a class of industrial processes based on recursive multiple models. **Neurocomputing**, n. 169, p. 430-438, 2015.

YUE, H. H.; QIN, S. J. Reconstruction-based fault identification using a combined index. **Industrial & engineering chemistry research**, n. 40(20), p. 4403-4414, 2001.

ZAR, J. H. Significance testing of the Spearman rank correlation coefficient. **Journal of the American Statistical Association**, n. 67(339), p. 578-580, 1972.

ZHANG, H. Statistical Process Monitoring and Modeling Using PCA and PLS. **Dissertação(Mestrado) — University of Alberta**, 2000.

ZHAO, S. J.; ZHANG, J.; XU, Y. M. Monitoring of processes with multiple operating modes through multiple principle component analysis models. **Industrial & engineering chemistry research**, **43(22)**, p. 7025–7035, 2004.

ZHENG, Y.; QIN, S. J.; CHAI, T. Fault Detection of Multimode Processes Using Concurrent Projection to Latent Structures. **IFAC-PapersOnLine** , n. 49.7, p. 705-710, 2016.

ZULLO, L. Validation and verification of continuous plants operating modes using multivariate statistical methods. **Computers & chemical engineering**, n. 20, p. S683-S688., 1996.