

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE CIÊNCIAS E TECNOLOGIA
UNIDADE ACADÊMICA DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO

**USO DE SENSORES VIRTUAIS (*SOFT SENSORS*) PARA ESTIMATIVA DE IMPUREZAS
EM COLUNAS DE DESTILAÇÃO DE ALTA PUREZA**

Arioston Araújo de Moraes Júnior

Campina Grande – Paraíba

Novembro/2015

ARIOSTON ARAÚJO DE MORAIS JÚNIOR

**USO DE SENSORES VIRTUAIS (*SOFT SENSORS*) PARA ESTIMATIVA DE IMPUREZAS
EM COLUNAS DE DESTILAÇÃO DE ALTA PUREZA**

Tese de doutorado apresentada à Universidade Federal de Campina Grande como parte dos requisitos exigidos para obtenção do título de Doutor em Engenharia Química.

Orientador: Prof. Dr. Romildo Pereira Brito

Campina Grande – Paraíba

Novembro/2015

FICHA CATALOGRAFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

M827u Morais Júnior, Arioston Araújo de.
 Uso de sensores virtuais (*Soft Sensors*) para estimativa de impurezas em colunas de destilação de alta pureza / Arioston Araújo de Morais Júnior. – Campina Grande, 2015.
 166f. : il. color.

 Tese (Doutorado em Engenharia Química) – Universidade Federal de Campina Grande, Centro de Ciências e Tecnologia.
 "Orientação: Prof. Dr. Romildo Pereira Brito".

 1. Destilação. 2. *Soft Sensors* (Sensores Virtuais). 3. Colunas de Destilação. 4. Sistema Azeotrópico. I. Brito, Romildo Pereira. II. Título.


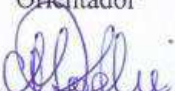

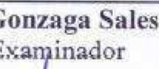
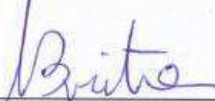

CDU 66.048(043)

**USO DE SENSORES VIRTUAIS (*SOFT SENSORS*) PARA ESTIMATIVA DE IMPUREZAS
EM COLUNAS DE DESTILAÇÃO DE ALTA PUREZA**

ARIOSTON ARAÚJO DE MORAIS JÚNIOR

Tese de Doutorado aprovada em: 06 / 11 / 2015

Banca Examinadora:


Prof. Dr. Romildo Pereira Brito Orientador

Prof. Dra. Cristiane Holanda Sodr� Examinadora

Prof. Dr. Jos� Nilton Silva Examinador

Prof. Dr. Lu� Gonzaga Sales Vasconcelos Examinador

Prof. Dra. N�bia Silva Dantas Brito Examinadora

Prof. Dra. Shirlene Kelly Santos Carmo Examinadora

Campina Grande – Para ba

Novembro/2015

“Há quem diga que todas as noites são de sonhos. Mas há também quem garanta que nem todas, só as de verão. No fundo, isto não tem muita importância. O que interessa mesmo não é a noite em si, são os sonhos. Sonhos que o homem sonha sempre, em todos os lugares, em todas as épocas do ano, dormindo ou acordado”.

William Shakespeare.

AGRADECIMENTOS

Para Dona Neide, minha mãe, testemunha ocular de toda minha dedicação pessoal, que me ensinou que chorar alivia, mas sorrir torna tudo mais bonito. Ao meu pai Arioston, que me norteou quanto às virtudes morais da vida e toda minha família por ser a base sólida de uma pirâmide fortalecida pelo amor.

Ao professor Dr. Romildo pela sua orientação, contribuições e confiança na expectativa do desenvolvimento do trabalho.

À empresa Braskem pelos dados fornecidos e ao Programa de Pós Graduação em Engenharia Química da UFCG (PPGEQ-UFCG) pela possibilidade de desenvolvimento da tese.

A todos os amigo(a)s do Laboratório de Referência em Controle e Automação - Grupo de Processos Químicos da UFCG (LARCA-GSQ-UFCG).

Aos amigos do curso de engenharia química da UFPB, em especial ao Professor Dr. Leopoldo Alcazár Rojas por sua amizade, sabedoria e questionamentos.

À professora Dra. Cristiane Holanda Sodr e e ao professor Dr. William Gonalves Vieira, do Laborat rio de Simula o e Controle da Universidade Federal de Alagoas (LASIC-UFAL).

RESUMO

Neste trabalho é proposta uma metodologia para construção de sensores virtuais implementados em *software*, com objetivo de estimar e prever o comportamento de impurezas na corrente de base de uma coluna de destilação de alta pureza, do processo produtivo do 1,2 Dicloroetano ($C_2H_4Cl_2$). A aquisição dos dados utilizados na construção dos sensores virtuais foi realizada através do modelo matemático do processo, simulado com dados reais de uma planta industrial. O estudo específico engloba a modelagem matemática/termodinâmica e avaliação do comportamento estacionário e dinâmico dessa torre, simulada aqui no *software Aspen Plus e Dynamics*TM. Desse modo, o modelo fornece os dados necessários para inferência das impurezas relacionadas, que são os teores dos compostos tetracloreto de carbono (CCl_4) e clorofórmio ($CHCl_3$), ambos devem ser mantidos, respectivamente, em valores ≤ 3000 e ≤ 400 ppm (partes por milhão). A metodologia também aborda dois algoritmos de seleção de variáveis secundárias, que utilizam técnicas estatísticas multivariadas (algoritmo de todas as regressões possíveis-TRP e da análise de componentes principais-PCA). Verifica-se também nos dados gerados quanto a real ou não necessidade de remoção de erros grosseiros (*outliers*), por isso é também inserida na metodologia uma etapa de pré-processamento de dados. Foram selecionados os dez melhores modelos de inferência para cada uma das saídas. Diante dessa informação, os melhores modelos produzidos não utilizavam as concentrações dos compostos das correntes de alimentação e sim medições de temperaturas ao longo da torre. Uma importante conclusão do ponto de vista de construção de sensores virtuais, porque na maioria dos trabalhos desenvolvidos essas variáveis são cruciais na produção de bons resultados. O treinamento dos sensores virtuais foi efetuado em um ambiente ruidoso, haja vista que foram simulados ruídos inerentes às medições (ruídos brancos Gaussianos). Na etapa final, os sensores virtuais são construídos utilizando uma técnica de modelagem empírica, redes neurais artificiais (RNA), onde foram utilizadas RNA do tipo Perceptron Multicamadas (MLP). Foram também avaliadas diversas variações quanto ao número de neurônios e camadas ocultas das RNA, empregando como critério de parada a técnica de validação cruzada. Os sensores virtuais desenvolvidos apresentaram erros satisfatórios do ponto de vista de engenharia, uma boa análise de regressão e um bom erro médio quadrático. Logo, com essas estimativas espera-se a minimização e a previsão do comportamento transiente dos compostos no referido processo.

ABSTRACT

This work proposes a methodology for software implementation to make soft sensors. The goal is to estimate and predict the behavior of impurities in the bottom current of a high-purity distillation column, for 1,2-Dichloroethane or 1,2-DCE ($C_2H_4Cl_2$) production. The data acquisition used in the construction of soft sensors was performed through a mathematical simulation of the process, with real industrial data taken from an industrial plant. A specific study involves the mathematical modeling, thermodynamics, evaluation of the steady state and the dynamic behavior of this process, simulated here in Aspen Plus and Aspen DynamicsTM software. Thus, the model provides the necessary data to infer the contents of the carbon impurities mentioned above, Tetrachloride (CCl_4) and Chloroform ($CHCl_3$), which are to be fixed approximately below 3000 and 400 ppm (parts per million) respectively. The methodology also covers selection algorithms of secondary variables, using multivariate statistical techniques: All Possible Regressions (TRP) and Principal Component Analysis (PCA). The data generated was checked in order to know whether to include or not a step for removal of outliers, so it was also included in the methodology one preprocessing data step. The ten best inference models were selected for each output concentration. With this information, these models do not use concentrations measurements in the feed streams but measurements of the temperature along the column. This is an important conclusion from the point of view of virtual sensors building, because in most of the literature reported these variables are crucial in getting good results. The training of soft sensors was done in a noisy environment, considering that simulated noise was inherent to measurements (Gaussian noise). In the final step, the soft sensors devices are constructed using an empirical modeling technique of artificial neural networks (ANN), which were generated ANN type Multilayer Perceptron (MLP). Several variations were also evaluated on the number of neurons and hidden layers of networks, employing as a stopping criterion the cross-validation technique. The developed soft sensors presented satisfactory errors from the engineering viewpoint, a good regression and a good mean square error. Finally, with these estimations it is expected to minimize and predict the transient behavior of the compounds in the referred process.

SUMÁRIO

LISTA DE FIGURAS	viii
LISTA DE TABELAS.....	xiii
NOMENCLATURA.....	xiii
CAPÍTULO 1: INTRODUÇÃO, OBJETIVOS, CONTRIBUIÇÕES E ORGANIZAÇÃO DA TESE.....	1
1.1. INTRODUÇÃO	1
1.2. OBJETIVOS.....	3
1.2.1. Objetivo Geral	3
1.2.2. Objetivos Específicos	3
1.3. CONTRIBUIÇÕES	4
1.4. ORGANIZAÇÃO DA TESE	5
CAPÍTULO 2: DESTILAÇÃO: FUNDAMENTOS, MODELAGEM E CONTROLE	7
2.1. INTRODUÇÃO	7
2.2. DESTILAÇÃO NÃO CONVENCIONAL	7
2.2.1. Destilação Azeotrópica	7
2.2.2. Destilação de Alta Pureza	10
2.3. PROCESSO PRODUTIVO E PURIFICAÇÃO DO 1,2-DICLOROETANO (DCE)....	12
2.4. SENSORES VIRTUAIS (SV) EM PROCESSOS DE DESTILAÇÃO.....	15
2.4.1. Metodologia de Construção	15
2.4.2. Dados de Entrada para os Sensores Virtuais	18
2.4.3. Modelos Aplicados na Construção de Sensores Virtuais	19
2.5. MODELAGEM MATEMÁTICA DE ESTÁGIOS EM EQUÍLIBRIO.....	23
2.5.1. Modelagem Termodinâmica	25
2.5.2. Solução Numérica das Equações MESH	26
2.6. CONTROLE DE COMPOSIÇÃO EM DESTILAÇÃO	28
2.6.2. Técnicas de seleção de pares de variáveis ou Sensibilidade da Malha	30
2.6.2. Controle Inferencial de Composição	34
2.7. CONCLUSÕES DO CAPÍTULO	37
CAPÍTULO 3: REDES NEURAIS ARTIFICIAIS PARA CONSTRUÇÃO DE SV	38
3.1. INTRODUÇÃO	38
3.2. NEURÔNIOS ARTIFICIAIS	38
3.3. TIPOS DE FUNÇÕES DE ATIVAÇÃO	40

3.4. ARQUITETURAS DE UMA RNA	41
3.5. PROCESSO DE APRENDIZAGEM DE UMA RNA	45
3.6. ALGORITMOS DE TREINAMENTO DE UMA RNA.....	46
3.7. CRITÉRIO DE PARADA DE UMA RNA.....	49
3.8. CONCLUSÕES DO CAPÍTULO	50
CAPÍTULO 4: SELEÇÃO DE VARIÁVEIS E PRÉ-PROCESSAMENTO DE DADOS	51
4.1. INTRODUÇÃO	51
4.2. TÉCNICAS DE SELEÇÃO DE VARIÁVEIS	51
4.2.1. Modelo de Regressão Linear Múltipla (MLR)	53
4.2.2. Índices de Desempenho	54
4.2.3. Métodos Sequenciais (Procedimentos Stepwise)	55
4.2.3.1. Todas as Regressões Possíveis (TRP, All Possible Regressions).....	56
4.2.3.2. Seleção Progressiva (Forward Selection (FS))	57
4.2.3.3. Regressão por Eliminação (Backward Elimination, (BE))	57
4.2.3.4. Regressão por etapas (Stepwise Regression (SR))	58
4.2.3.5. Algoritmo da Seleção de Entradas	59
4.3. REMOÇÃO DE ERROS GROSSEIROS (OUTLIERS).....	61
4.3.1. Análise de Componentes Principais (PCA)	61
4.3.1.1. Cálculos da Análise de Componentes Principais.....	63
4.3.1.2. Metodologia da PCA para Seleção de Variáveis e Remoção de Outliers.....	65
4.3.2. Análise Residual para remoção de <i>Outliers</i>	66
4.4. CONCLUSÕES DO CAPÍTULO	68
CAPÍTULO 5: COLETA DE DADOS E AVALIAÇÃO DE MODELOS DE INFERÊNCIA	69
5.1. INTRODUÇÃO	69
5.2. SIMULAÇÃO DO PROCESSO	70
5.2.1 Simulação em regime estacionário	70
5.2.1.1. Modelo Termodinâmico Ajustado	73
5.2.1.2. Resultados em Regime estacionário	74
5.2.1.3. Efeito da vazão de levas (F_3).....	76
5.2.1.4. Efeito da Carga Térmica do Refervedor (Q_R).....	77
5.2.1.5. Efeito das Vazões de Alimentação da Torre (F_1 e F_2).	78
5.2.2. Simulação em Regime Dinâmico e Controle	81
5.2.2.1. Estrutura de Controle	81

5.2.2.2. Caracterização das Respostas em Regime transiente.....	83
5.2.2.3. Dinâmica das Concentrações (Perturbações nas Vazões de Alimentação).....	83
5.2.2.4. Dinâmica das Concentrações (Perturbações na Vazão de Destilado, F_3)	85
5.2.2.5. Dinâmica das Concentrações (Perturbações na carga térmica, QR).....	86
5.2.2.6. Dinâmica das Concentrações (Perturbações nas concentrações alimentação) 87	
5.2.2.7. Temperaturas dos Estágios.....	90
5.3. CONJUNTO DE DADOS	94
5.4. DADOS PRÉ-PROCESSADOS	100
5.4.1. Avaliação de <i>Outliers</i> por meio da PCA.....	101
5.4.2. Avaliação de <i>Outliers</i> por meio de Análise Residual.	104
5.5. ETAPA DE SELEÇÃO DE VARIÁVEIS	105
5.5.1. Seleção Utilizando a Técnica de todas as Combinações Possíveis (TRP).	106
5.5.2. Seleção Utilizando a PCA.....	111
5.6. CONCLUSÕES DO CAPÍTULO	113
CAPÍTULO 6: SENSORES VIRTUAIS PARA INFERÊNCIA DAS IMPUREZAS	115
6.1. INTRODUÇÃO.....	115
6.2. SENSORES VIRTUAIS NEURAI (SVN)	116
6.2.1 Modelos de Inferência Avaliados para Treinamento dos SNV.....	116
6.2.2 Topologia de Construção dos SNV.....	117
6.2.3. Ambiente de Treinamento dos SVN.....	118
6.2.4. Tipo de RNA e Especificações da Arquitetura Empregada.....	119
6.2.5. Algoritmos de Treinamento utilizados nas redes MLP.....	121
6.3. RESULTADOS DAS DAS INFERÊNCIAS DAS IMPUREZAS	122
6.3.1. Avaliação de Desempenho e Modelos Finais para Inferência	123
6.3.2. Análise Gráfica dos Sensores Virtual de XCl_4 e XCH_3	128
6.3.3. Inferência do XCl_4 XCH_3 Com Remoção de Erros Grosseiros.....	135
6.4. CONCLUSÕES DO CAPÍTULO	138
CAPÍTULO 7: CONCLUSÕES E SUGESTÕES	140
7.1. CONCLUSÕES DO TRABALHO	140
7.2. SUGESTÕES PARA TRABALHOS FUTUROS.....	142
7.3. PROPOSTA DE CONTROLE POR INFERÊNCIA	142
REFERÊNCIAS BIBLIOGRÁFICAS	144

LISTA DE FIGURAS

Figura 2.1.a) Diagrama Txy e b) Curva xy; Sistema (1) - (2) azeotrópica homogênea.....	9
Figura 2.2. a) Diagrama Txy e b) Curva xy; Sistema (1) - (2) azeotrópica heterogênea	9
Figura 2.3. Processo produtivo do 1,2 Dicloroetano e do Monômero Cloreto de Vinila	12
Figura 2.4. Fluxograma esquemático do processo de purificação do 1,2-DCE.	14
Figura 2.5. Proposta de metodologia para construção de sensores virtuais.	17
Figura 2.6. Sensor virtual com entradas fornecidas por um modelo matemático do processo	18
Figura 2.7. Principais técnicas de identificação empírica utilizadas para construção de sensor virtual.	20
Figura 2.8. a) Representação de um estágio em equilíbrio em regime estacionário. b) Esquema geral de uma cascata em contracorrente com N estágios	23
Figura 2.9. Esquema de Resolução das Equações MESH pelo método Inside-out	27
Figura 2.10. Sistema de controle convencional em processos de destilação.	29
Figura 2.11. Controle realimentado descentralizado multivariável (2 x 2).....	30
Figura 2.12. Diagrama de blocos de um <i>softsensor</i> utilizado para controle inferencial	36
Figura 3.1. Neurônio Artificial.....	39
Figura 3.2. Funções de ativação do tipo: a) Linear; b) Tangente hiperbólica; c) Logarítmica-sigmóide; d) Linear simétrica; e) Linear assimétrica; f) Linear com saturação simétrica; g) Linear com saturação assimétrica; e h) Linear positiva.....	42
Figura 3.3. a) Rede adiante com uma única camada; b) Rede adiante multicamadas..	43
Figura 3.4. a) Rede <i>Perceptron</i> multicamadas; b) Rede recorrente <i>Perceptron</i> multicamadas... ..	44
Figura 3.5. Diagrama de blocos de aprendizagem supervisionada	46
Figura 3.6. Sinais funcionais e de erro numa rede neural.	47
Figura 4.1. Algoritmo em forma de fluxograma da Regressão por Eliminação	58
Figura 4.2. Algoritmo da técnica <i>Stepwise Regression</i>	59
Figura 4.3. Componentes Principais	62
Figura 4.4. Esquema do cálculo da análise de componentes principais (PCA)	65
Figura 4.5. Um ponto que está longe no espaço x.....	66
Figura 4.6. a) Análise residual da saída fornecida pelo modelo versus o erro residual padronizado; e b) Análise residual das entradas X_i versus o erro residual	67
Figura 5.1. Metodologia desenvolvida para construção dos sensores virtuais.....	69
Figura 5.2. Fluxograma da torre de purificação do 1,2-DCE desenvolvido no <i>Aspen Plus</i> TM	71

Figura 5.3. Perfil de composição líquida nos estágios da coluna.....	75
Figura 5.4. Perfil de temperaturas ao longo da coluna.....	75
Figura 5.5. Efeito do aumento de F_3 em: a) $XCCl_4$; e b) $XCHCl_3$	77
Figura 5.6. Efeito do aumento de Q_R em: a) $XCCl_4$; e b) $XCHCl_3$	78
Figura 5.7. Efeito do aumento de F_1 em: a) $XCCl_4$; e b) $XCHCl_3$	79
Figura 5.8. Efeito do aumento de F_2 em: a) $XCCl_4$; e b) $XCHCl_3$	79
Figura 5.9. Efeito da diminuição de F_1 em: a) $XCCl_4$; e b) $XCHCl_3$	80
Figura 5.10. Efeito da diminuição de F_2 em: a) $XCCl_4$; e b) $XCHCl_3$	80
Figura 5.11. Fluxograma de processo em regime dinâmico desenvolvido no <i>Aspen Dynamics</i> TM	81
Figura 5.12. Efeito das perturbações degrau efetuadas em F_1 sobre: a) $XCCl_4$ e b) $XCHCl_3$	84
Figura 5.13. Efeito das perturbações degrau efetuadas em F_2 sobre: a) $XCCl_4$ e b) $XCHCl_3$	85
Figura 5.14. Efeito das perturbações degrau efetuadas em F_3 sobre: a) $XCCl_4$ e b) $XCHCl_3$	86
Figura 5.15. Efeito de perturbações degrau efetuadas em Q_R sobre: a) $XCCl_4$ e b) $XCHCl_3$	87
Figura 5.16. a) PRBS na concentração do CCl_4 da alimentação F_1 e seu efeito sobre $XCCl_4$; b) PRBS na concentração do $CHCl_3$ da alimentação F_1 e seu efeito sobre $XCHCl_3$	89
Figura 5.17. a) PRBS na concentração do CCl_4 da alimentação F_2 e seu efeito sobre $XCCl_4$; b) PRBS na concentração do $CHCl_3$ da alimentação F_2 e seu efeito sobre $XCHCl_3$	89
Figura 5.18. Representação gráfica do método SVD, matrizes de sensibilidade U_1 (F_1) e U_2 (F_3).	93
Figura 5.19. Variação degrau nas vazões efeito sobre as temperaturas: a) T_{71} e b) T_{17} e T_{27}	94
Figura 5.20. Procedimento para coleta dos dados a) Dinâmica de F_1 ; e b) Dinâmica de F_2	97
Figura 5.21. Variações PRBS na variável da carga térmica QR	97
Figura 5.22. a) PRBS na concentração do CCl_4 na alimentação F_1 ; e b) PRBS na concentração do CCl_4 na alimentação F_2	98
Figura 5.23. a) PRBS na concentração do $CHCl_3$ na alimentação F_1 ; e b) PRBS na concentração do $CHCl_3$ na alimentação F_2	98
Figura 5.24. Dinâmica das saídas frente as perturbações e ruídos: a) Temperaturas selecionadas no estudo; b) $XCCl_4$ e c) $XCHCl_3$	99
Figura 5.25. Análise PCA: a) Conjuntos de Dados 1 (Saída $XCCl_4$); b) Conjuntos de Dados 2 (Saída $XCCl_4$); c) Conjuntos de Dados 3 (Saída $XCHCl_3$); e e) Conjuntos de Dados 4 (Saída $XCHCl_3$).....	103

Figura 5.26. Gráfico dos erros padronizados residuais (d_i) versus saídas: a) Dados 1; b) Dados 2; c) Dados 3; e d) Dados 4	105
Figura 5.27. Inferência de XCCl_4 : a) Dez melhores modelos Dados 1; b) Dez piores modelos Dados 1; c) Dez melhores modelos Dados 2; e d) Dez piores modelos Dados 2	108
Figura 5.28. Inferência de XCHCl_3 : a) Dez melhores modelos Dados 3; b) Dez piores modelos Dados 3; c) Dez melhores modelos Dados 4; e d) Dez piores modelos Dados 4.	109
Figura 6.1. Metodologia para construção dos sensores virtuais (em Destaque Treinamento dos SV).	115
Figura 6.2. a) Doze melhores RNA (MQE), Etapa Treinamento das RNA para a saída XCCl_4 b) Doze melhores RNA (MQE), Etapa Validação das RNA para a saída XCCl_4	124
Figura 6.3. a) Doze melhores RNA (MQE), Etapa Treinamento das RNA para a saída XCHCl_3 b) Doze melhores RNA (MQE), Etapa Validação das RNA para a saída XCHCl_3	127
Figura 6.4. Configuração dos SVN na corrente de base da torre de destilação do 1,2-DCE.	128
Figura 6.5. Dados de Treinamento: a) Sensor virtual neural do XCCl_4 e XCCl_4 fornecida pelo simulador b) Sensor virtual neural do XCHCl_3 e XCHCl_3 fornecida pelo simulador.	129
Figura 6.6. a) Sensor virtual neural do XCCl_4 e saída do simulador (Validação Cruzada); e b) Sensor virtual neural do XCHCl_3 e saída do simulador (Validação Cruzada).	130
Figura 6.7. a) Sensor Virtual Neural dinâmico (XCCl_4) versus a saída desejada (Simulador); e b) Sensor Virtual Neural dinâmico (XCHCl_3) versus a saída desejada (Simulador).	132
Figura 6.8. Dispersões dos dados estimados pelos SVN: a) regressão dos dados para XCCl_4 ; e b) regressão dos dados para XCHCl_3	133
Figura 6.9. a) Erro Médio Quadrático do sensor virtual de XCCl_4 ; b) Erro Médio Quadrático do sensor virtual de XCHCl_3	134
Figura 6.10. Sensores virtuais com remoção de <i>outliers</i> : a) MQE do XCCl_4 e XCHCl_3 dados de treinamento; b) MQE do XCCl_4 e XCHCl_3 dados validação.	135
Figura 6.11. Dados de Treinamento: a) Sensor virtual neural do XCCl_4 remoção de outliers b) Sensor virtual neural do XCHCl_3 remoção de outliers.	137
Figura 7.1. Diagrama de realimentação negativa, controle por inferencial do clorofórmio na corrente de base da torre do 1,2-DCE.	143

LISTA DE TABELAS

Tabela 2.1. Principais referências utilizadas ao longo do desenvolvimento do trabalho.....	37
Tabela 3.1. Dados de composição mediante diferentes modelos termodinâmicos.	48
Tabela 4.1. Principais algoritmos de treinamentos, seus respectivos métodos de otimização e função de treinamento no software Matlab®.	56
Tabela 5.1. Componentes presentes em pelo menos em uma das correntes (F_1 e F_2) da Torre...	72
Tabela 5.2. Dados operacionais do Processo utilizados nas Simulações.	82
Tabela 6.3. Componentes considerados presentes nas correntes de Alimentação da Torre.	116
Tabela 5.3. Lista de controladores instalados no processo e representados no simulador.....	82
Tabela 5.4. Lista de válvulas de controle.	82
Tabela 5.5. Correlações para sintonia de controladores, TYREUS-LUYBEN (1997).	83
Tabela 5.6. Rotina de perturbações nas concentrações de alimentação.	88
Tabela 5.7. Rotina de perturbações para o cálculo da SVD.	91
Tabela 5.8. Resultados do Método SVD.	92
Tabela 5.9. Caracterização das perturbações utilizadas nas simulações para coleta dos dados....	96
Tabela 5.10. Variáveis candidatas a serem utilizadas em um modelo de inferência.	101
Tabela 5.11. a) Autovalores e percentual da variância explicada de cada variável candidata com relação a saída $XCCl_4$. b) Autovalores e percentual da variância explicada de cada variável candidata com relação a saída $XCHCl_3$	106
Tabela 5.12. a) Autovetores (Loadings) das PC selecionados com relação a saída $XCCl_4$; b) Autovetores (Loadings) das PC selecionados com relação a saída $XCHCl_3$	111
Tabela 6.1. a) Melhores RNA do tipo MLP obtidas para inferência de composição do $XCCl_4$ através do algoritmo de treinamento Levenberg-Marquardt. b) Melhores RNA do tipo MLP obtidas para inferência de composição do $XCCl_4$ através do algoritmo de treinamento Levenberg-Marquardt.....	116
Tabela 6.2. Topologias das RNA avaliadas para construção dos sensores virtuais.	120
Tabela 6.2. Arquiteturas das RNA avaliadas para construção dos sensores virtuais.	121

NOMENCLATURA

- a - parâmetro da equação do modelo termodinâmico NRTL
- b - parâmetro da equação do modelo termodinâmico NRTL
- C - número de componentes
- B - Matriz Inversa Transposta de $K_{n \times m}$
- E - relação de equilíbrio
- F₁- Vazão do 1,2 DCE proveniente de um processo de oxicloração
- F₂- Vazão do 1,2 DCE proveniente de um processo de cloração direta
- F₃- Vazão de Destilado/Vazão de Hidrocarbonetos Clorados úmidos
- F₄- Vazão de base/Vazão de base do 1,2-DCE
- F_INCOND- Vazão de incondensáveis do vaso de refluxo
- F_ÁGUA- Vazão de água na saída do vaso de refluxo
- G_{ij} - parâmetro de interação binária da equação NRTL
- H - balanço de energia (entalpia (kcal/kg))
- H_{H2O}- Nível de água do vaso de refluxo;
- H_{org}- Nível de orgânicos clorados do vaso de refluxo
- H_{referv} - Nível do refeedor/base da coluna.
- K_{nxm} - Matriz de sensibilidade da SGA
- K - constante de equilíbrio/volatilidade
- K'_{nxm} - Matriz de ganho em malha fechada
- L - vazão de líquido (kg/h)
- M - balanço de massa (kg/h)
- N - número de estágios
- Q - carga térmica (kcal/h)
- Q_r - carga térmica do refeedor (kcal/h)
- P_{VR}- Pressão do vaso de refluxo
- R- Vazão de Refluxo
- R_g - constante universal dos gases perfeitos
- RR - razão de refluxo
- s - somatório das frações molares
- T - temperatura (°C)
- T_b - temperatura de bolha (°C)

U - vazão de retirada lateral de líquido (kg/h)
V - vazão de vapor (kg/h)
W - vazão de retirada lateral de vapor (kg/h)
x - fração do componente na fase líquida
y - fração do componente na fase vapor
z - fração mássica do componente na alimentação

Letras gregas

σ_i - valores singulares da matriz K
 α_{ij} - constante não-randômica para interação binária do modelo termodinâmico NRTL
 ϕ_i - coeficiente de fugacidade do componente i na fase vapor
 γ_i - coeficiente de atividade do componente i na fase líquida
 τ_{ij} - parâmetro de interação binária da equação NRTL
 λ - valor característico (autovalor) da matriz $K_{n \times m}$
 $\Lambda_{n \times m}$ - Matriz de ganho relativo
 Λ_{ij} - Ganho relativo individual

Subscritos

i - componente i (1,2,...,C)
j - estágio j (1,2,...,N)
L - fase líquida
V - fase vapor

Siglas e definições

CCl₄ - Tetracloroeto de Carbono
CHCl₃ - Clorofórmio
C₂H₄Cl₂ - 1,2 Dicloroetano
CN - Número Condicional

ELV - Equilíbrio Líquido-Vapor
EOS - Equação de Estado
HCU - Hidrocarbonetos Clorados Úmidos
LM - Levenberg-Marquardt
NRTL - Non-Random Two Liquid, modelo para o cálculo do coeficiente de atividade
NRTL-HOC - Modelo NRTL com a Equação de Estado de Hayden-O'Connell
OC - Orgânicos Clorados
PCA - Análise de Componentes Principais
PPM - Partes por milhão
PV_i - Variável de processo medida e/ou controlada i;
RGA - Matriz de Ganho Relativo
RNA - Redes Neurais Artificiais
SGA - Matriz de Ganho Estacionário
SSID - Método de Identificação no Subespaço de Dois Estágios
SV - Sensor Virtual
SVD - Decomposição de Valores Singulares
SVN - Sensor Virtual Neural
1,2 - DCE - 1,2 Dicloroetano

CAPÍTULO 1:

INTRODUÇÃO, OBJETIVOS, CONTRIBUIÇÕES E ORGANIZAÇÃO DA TESE

1.1. INTRODUÇÃO

A destilação é um processo de separação amplamente utilizado pelas indústrias químicas e petroquímicas, representando, aproximadamente, um terço do custo de produção e mais de 60% do consumo energético de uma planta, constituindo-se, portanto, em uma das principais fontes de custos de capital e de operação, com possibilidades de serem reduzidos (KISTER, 2006).

Em processos de destilação, bem como em diversos outros processos químicos, as industriais modificam e melhoram suas tecnologias quando são requeridas em aumentar sua produção, condições efetivas de segurança, seus índices de eficiência e a qualidade final do produto. Nesse contexto, a importância do monitoramento das variáveis críticas de processos utilizando dispositivos de medição adequados é evidente. No entanto, uma série de obstáculos para controle e acompanhamento em tempo real de algumas variáveis importantes de certos processos industriais pode tornar este objetivo impraticável. Sejam por razões nas quais o sensor adequado não exista ou por ser demasiadamente caro (FORTUNA et al., 2007).

As medições de variáveis físicas são garantidas de maneira satisfatória com o uso de inúmeros sensores físicos, como exemplo: termômetros e termopares na medição de temperatura, tubo de Bourdon na leitura de pressão, placa de orifício e rotâmetro para medição de vazão, sensores de pressão diferenciais e flutuantes para medição nível, dentre outros. Por outro lado, quando se deseja medir em tempo real variáveis químicas e bioquímicas, não facilmente mensuráveis, tais como teor ou composição de produtos, biomassa e pH, surge o problema de como fazer tais medições. Para resolver essas e outras dificuldades surgiu então o *soft sensor*, também conhecido como sensor virtual (SV) ou modelo inferencial (MI).

O SV é um modelo matemático utilizado para estimar variáveis de interesse em função de variáveis de entrada, ditas secundárias, de fácil medição. Esta alternativa surge de uma dificuldade operacional ou do alto custo na obtenção da variável desejada ou, ainda, utilizado na

previsão de falha de um sensor físico, seja pela sua quebra ou pela sua retirada do processo para manutenção e calibração (MORAIS JR, 2011).

Segundo Lotufo e Garcia (2008), a necessidade de implantação de um SV surge de diversos problemas a serem enfrentados, tais como: inferência de variáveis, detecção de falhas, previsão de desempenho, aproximação de funções não lineares e identificação de padrões de processos os quais são requeridos.

A gama de tarefas a serem cumpridas pelos sensores virtuais é ampla, onde a sua área original de aplicação e ainda mais dominante é a predição de variáveis de processos, que podem ser determinadas a um baixo tempo/período de amostragem ou somente através de análises *off-line*. Nestes casos, o SV pode ser uma boa alternativa em relação às medições tradicionais, desde que as variáveis de entradas (secundárias) possam ser fornecidas/medidas sem problemas e quando a simulação é utilizada para gerar os dados à construção do sensor (geralmente implementada em um *software*) seja capaz de fornecer a variável desejada de forma precisa (KADLEC et al., 2009).

Um dos produtos de maior valor agregado da cadeia produtiva do Policloreto de Vinila (PVC) é o 1,2-DCE, matéria-prima básica para a produção do Monocloreto de Vinila (MVC). O 1,2-DCE para ser convertido no MVC, através do craqueamento térmico, deve estar seco e livre de impurezas. A estimativa dessas impurezas, produtos de base no processo de destilação do 1,2-Dicloroetano, se mostra adequada para construção de uma metodologia de desenvolvimento de sensores virtuais, haja vista que os valores das concentrações dessas variáveis (impurezas) são altamente críticos e obtidos apenas por medições *off-line* (análises de laboratório).

Além disso, o processo de destilação do 1,2-DCE é um processo de alta pureza, caracterizado por ser altamente não linear e sensível às perturbações inerentes ao processo. Ou seja, pequenas alterações nas correntes de alimentação produzem mudanças drásticas nas concentrações dos produtos de base. Logo, a estimativa dessas impurezas por meio de sensores virtuais (*soft sensor*) pode trazer diversos benefícios ao processo, dentre os quais estão a possibilidade da inferência em tempo real das composições e a aplicação de controle inferencial dos compostos.

1.2. OBJETIVOS

1.2.1. Objetivo Geral

Este trabalho tem como objetivo desenvolver uma metodologia de construção de sensores virtuais para o processo de destilação do 1,2-Dicloetano (1,2-DCE). A purificação do 1,2-EDC é feita por meio de um conjunto de três torres de destilação, antes que o mesmo venha convergir através do craqueamento térmico para o Monômero Cloreto de Vinila (MVC). Ao fim dessa etapa exige-se que o composto 1,2-DCE deva estar seco (com um teor água menor que 10 ppm) de forma a evitar a corrosão excessiva dos equipamentos e ter uma alta pureza (tipicamente maior que 99,5 % peso/peso), haja vista que o processo de craqueamento é altamente suscetível à inibição e deposição por traços de impurezas.

As dificuldades para estimativa em tempo real e controle das impurezas na base da primeira torre de destilação, que são as composições do tetracloreto de carbono, CCl_4 , (mantido em torno de 3000 ppm) e do clorofórmio, CHCl_3 , (mantido abaixo de 400 ppm), motivaram e o tornaram o processo como estudo de caso nesse trabalho.

1.2.2. Objetivos Específicos

1. Realizar a modelagem e simulação da torre de destilação do 1,2-DCE, verificando o comportamento nos regimes estacionário e dinâmico das principais variáveis do processo, bem como seus efeitos sobre as composições do CCl_4 e do CHCl_3 na base da torre 1;
2. Propor uma metodologia de construção de dois sensores virtuais dinâmicos, com o objetivo de estimar as concentrações dessas impurezas;
3. Realizar a coleta dos dados no simulador dinâmico (*Aspen Dynamics*TM), verificar a necessidade de remoção de erros grosseiros (*outliers*) e selecionar as variáveis de entrada;
4. Verificar quais os melhores modelos à inferência e configuração das redes neurais artificiais para identificação do processo;

1.3 CONTRIBUIÇÕES

Uma vez que foram definidos os objetivos traçados para este trabalho, as principais contribuições da tese são:

- ❖ Desenvolvimento de uma metodologia inovadora no sentido de construção de sensores virtuais, englobando desde a possibilidade de captação dos dados em diferentes condições de operação do processo até o uso de diferentes técnicas de seleção de variáveis por meio de estatística multivariada;
- ❖ A estimativa de composição, neste trabalho, é considerada não convencional, haja vista o fato que o componente-chave de base, 1,2-DCE, possui ponto de ebulição intermediário em relação aos demais componentes presentes. Nessa condição, as contribuições do trabalho vão desde a seleção rigorosa das variáveis secundárias (regressores) em um processo de alta pureza, até a inferência de compostos com baixíssimas concentrações (em torno de ppm);
- ❖ Comparar diversos modelos selecionados para inferência, avaliar quanto ao uso das variáveis manipuladas e distúrbios junto com as concentrações de alimentação ou uso de temperaturas ao longo da torre;
- ❖ A estimativa de composição que se deseja obter é dupla, haja vista a inferência das concentrações de dois compostos contaminantes na mesma corrente (CCl_4 e CHCl_3). Não foi verificado algo similar na literatura consultada sobre *softsensor* ou em problemas de natureza similar.

Foram publicados os seguintes artigos em anais de congressos e periódicos gerados diretamente por este trabalho:

1. “*Design of a Soft Sensor with Technique Neuro-Fuzzy to Infer the Product Composition a Distillation Process*”. Publicado no periódico: *Lecture Notes in Engineering and Computer Science*, v. II, p. 605-608, 2014;

2. “Análise de Sensibilidade para Fins de Controle em uma Coluna de Destilação azeotrópica de alta pureza”. Publicado no: Congresso Brasileiro de Engenharia Química, Florianópolis, 2014.

3. “*Design of a Soft Sensor with Technique Neuro-Fuzzy to Infer the Product Composition a Distillation Process*”. Publicado no: *World Congress on Engineering and Computer Science* 2014, 2014, San Francisco, USA. Newswood Limited, 2014;

4. “*Seleção Ótima de Entradas no Estudo de Inferência de Composições em Processo de Destilação Industrial de Alta Pureza*”. Publicado no: Congresso Nacional de Matemática Aplicada à Indústria - CNMAI2014, 2014, Caldas Novas, Goiás, Brasil. CNMAI2014, 2014.

5. “*Contribuições à Análise Transiente e Inferência de Composição em uma Coluna de Destilação*”. Publicado no: XXXV *Iberian Latin American Congress on Computational Methods in Engineering*, CILAMCE 2014, 2014, Fortaleza, Ceará, Brasil. CILAMCE 2014, 2014.

6. “*Contribuições à Análise Transiente e à Estimativa de Variáveis Críticas em uma Coluna Industrial de Alta Pureza Utilizando Redes Neurais Artificiais*”. Publicado no: XXXVI *Iberian Latin American Congress on Computational Methods in Engineering*, CILAMCE 2015, Rio de Janeiro-RJ, Brasil, 2015.

1.4. ORGANIZAÇÃO DA TESE

A tese está organizada como descrito a seguir:

O capítulo 02 apresenta a revisão bibliográfica realizada durante o desenvolvimento do trabalho. Essa revisão vai desde a fundamentação teórica sobre colunas azeotrópicas e de alta pureza, estudo da arte sobre sensores virtuais, modelagem matemática de processos de separação de estágios em equilíbrio e técnicas de sensibilidade de controle.

O capítulo 03 descreve uma revisão e a metodologia do modelo caixa preta empregado no trabalho, rede neurais artificiais (RNA), envolvendo suas principais topologias (arquiteturas de redes), funções de ativação, algoritmo de treinamento e critérios paradas utilizados na tese. Esse capítulo é de suma importância, haja vista que trata da melhor maneira de como implementar os sensores virtuais por meio dessa técnica de sistemas inteligentes.

O capítulo 04 aborda quanto à real necessidade ou não de pré-processamento dos dados de entradas do sensor virtual, sendo assim são utilizadas duas técnicas distintas: análise de componentes principais (PCA) e o erro residual padronizado. Além disso, também é abordada a metodologia para seleção de variáveis de entrada dos modelos de identificação propostos, onde são explanados os principais algoritmos utilizados no desenvolvimento do trabalho.

No capítulo 05 são apresentados a primeira parte dos resultados obtidos no trabalho. Esses resultados expressam desde o comportamento estacionário e transiente do processo, quando são inseridas perturbações nas variáveis de entrada, até os obtidos por meio seleção de variáveis e modelos de inferência de composição, sensores virtuais. Para que esse último objetivo fosse alcançado é trazido à tona uma discussão sobre as principais variáveis candidatas a serem os regressores do modelo, também é feita uma análise de sensibilidade por meio da Decomposição de Valores Singulares (SVD) que diz quanto à necessidade ou não do uso de temperaturas para inferência ou controle de composição na torre de alta pureza. Ainda nessa etapa é explanada a maneira mais representativa de como os dados foram coletados, quanto à necessidade de remoção de erros grosseiros (*outliers*).

O capítulo 06 apresenta uma discussão sobre a melhor estrutura de redes neurais, visando a inferência das saídas desejadas do processo. O treinamento dos sensores é efetuado por meio dos melhores modelos, previamente selecionados na etapa de seleção de variáveis.

O capítulo 07 apresenta as conclusões e as sugestões propostas para trabalhos futuros.

CAPÍTULO 2:

DESTILAÇÃO: FUNDAMENTOS, MODELAGEM E CONTROLE

2.1. INTRODUÇÃO

Este capítulo apresenta a revisão bibliográfica realizada ao longo do desenvolvimento do trabalho e suas principais contribuições. A revisão parte, desde uma abordagem sobre misturas azeotrópicas e de alta pureza, bem como da importância do processo produtivo do composto 1,2-DCE, ênfase no objeto de estudo do trabalho, a destilação do 1,2-DCE. O foco principal é apresentar uma revisão geral sobre sensores virtuais, metodologias de construção, dados de entradas fornecidos (ênfase na modelagem matemática) e sobre métodos aplicados em sua construção. Além disso, aqui é efetuada uma revisão sobre o controle de concentração em processos de destilação e dos métodos que auxiliam na escolha da melhor estrutura de controle.

2.2. DESTILAÇÃO NÃO CONVENCIONAL

2.2.1 Destilação Azeotrópica

Um azeótropo é uma mistura líquida em equilíbrio com a fase vapor com composições idênticas, ou seja, a mistura líquida em equilíbrio com o vapor apresenta comportamento de uma substância pura, vaporizando-se e condensando-se com a composição, temperatura e pressão constantes. Muitas misturas de componentes químicos formam misturas azeotrópicas, e esses azeótropos se formam por causa de equilíbrio de fases não ideal resultante da interação molecular (repulsão ou atração) de componentes químicos diferentes.

Caso as moléculas exibam forças de repulsão, o azeótropo de ebulição será mínimo e pode ser homogêneo ou heterogêneo. Caso contrário, o azeótropo de ebulição será máximo (LUYBEN e CHIEN, 2010). O máximo ou mínimo satisfazem as Equações 2.1 e 2.2.

$$\left\{ \begin{array}{l} \left(\frac{\partial T}{\partial x_i} \right)_{P,az} = \left(\frac{\partial T}{\partial y_i} \right)_{P,az} = 0, \text{ com } i = 1, 2, \dots, N-1. \\ x_i = y_i, \text{ com } i = 1, 2, \dots, N-1. \end{array} \right. \quad (2.1)$$

$$\left\{ \begin{array}{l} \left(\frac{\partial T}{\partial x_i} \right)_{P,az} = \left(\frac{\partial T}{\partial y_i} \right)_{P,az} = 0, \text{ com } i = 1, 2, \dots, N-1. \\ x_i = y_i, \text{ com } i = 1, 2, \dots, N-1. \end{array} \right. \quad (2.2)$$

A presença de um azeótropo é uma indicação de que a mistura não apresenta um comportamento ideal, ou seja, de que há desvios em relação à lei de Raoult. O desvio positivo

implica em ser o coeficiente de atividade maior que 1,0, onde as moléculas dos componentes do sistema repelem-se e apresentam uma alta pressão parcial, por este motivo, observa-se a formação de um azeótropo de mínimo ponto de ebulição (temperatura de ebulição menor que a dos componentes puros), esses ocorrem com maior frequência. O desvio negativo significa que o coeficiente de atividade é menor que 1,0, resultando em baixas pressões parciais e na formação de azeótropos de máximo ponto de ebulição.

Se o desvio em relação à Lei de Raoult não é suficiente para provocar a ocorrência de um azeótropo, então os pontos de ebulição dos componentes puros devem estar suficientemente próximos para que haja a possibilidade de ocorrência de uma temperatura de máxima ou de mínima.

Os azeótropos são raros em substâncias cujos pontos de ebulição diferem em mais de 30°C. Os componentes que têm ebulições próximas, com pequenos desvios em relação à idealidade, podem formar azeótropo, enquanto outros compostos, que formam soluções fortemente não ideais, podem não apresentar azeotropismo em virtude da grande diferença entre seus pontos de ebulição. Logo, a existência de azeotropia depende essencialmente de dois fatores: grau de não idealidade da mistura e das diferenças entre os pontos de ebulição dos componentes (SANDLER, 2006; HALDER, 2014).

Independentemente de ser um azeótropo de mínimo ou de máximo de temperatura de ebulição, esse pode ser homogêneo ou heterogêneo. Quando uma mistura azeotrópica formar apenas uma simples fase líquida em equilíbrio com a fase vapor, a mistura é azeotrópica homogênea. Se o comportamento de múltiplas fases líquidas for exibido na condição azeotrópica, o azeótropo é heterogêneo. Nesses últimos, os desvios positivos em relação à lei de Raoult são suficientemente grandes, a imiscibilidade parcial pode ocorrer e a presença de duas fases líquidas pode persistir até a temperatura de ebulição.

Um exemplo de um sistema azeotrópico homogêneo binário é o 1,2-DCE (1)/CCl₄ (2), as Figuras 2.1 a) diagramas T_{xy} e b) diagrama xy apresentam o comportamento desse sistema. Um exemplo de um sistema azeotrópico binário heterogêneo é o CCl₄ (1)/H₂O (2), representado conforme as Figuras 2.2 a) diagramas T_{xy} e b) diagrama xy. Tanto nas Figuras 2.1 quanto nas Figuras 2.2, os dados de T_{xy} e xy foram obtidos a partir de simulações com o *Aspen Plus*TM, utilizando o modelo de composição local NRTL.

Figura 2.1. a) Diagrama Txy e b) Curva xy; Sistema (1) - (2) azeotrópica homogênea.

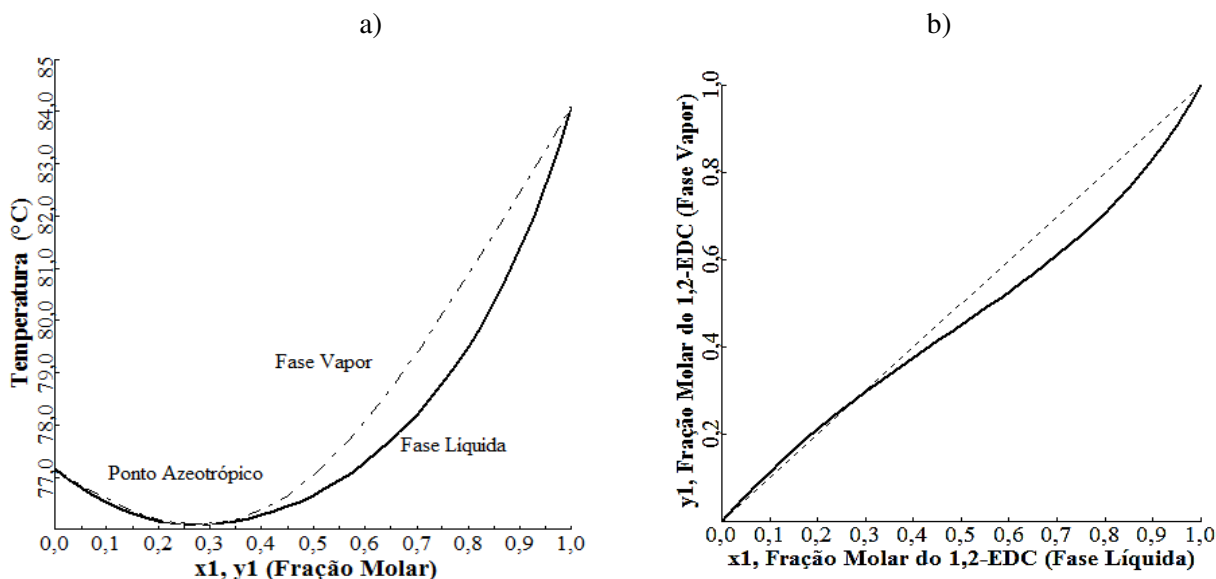
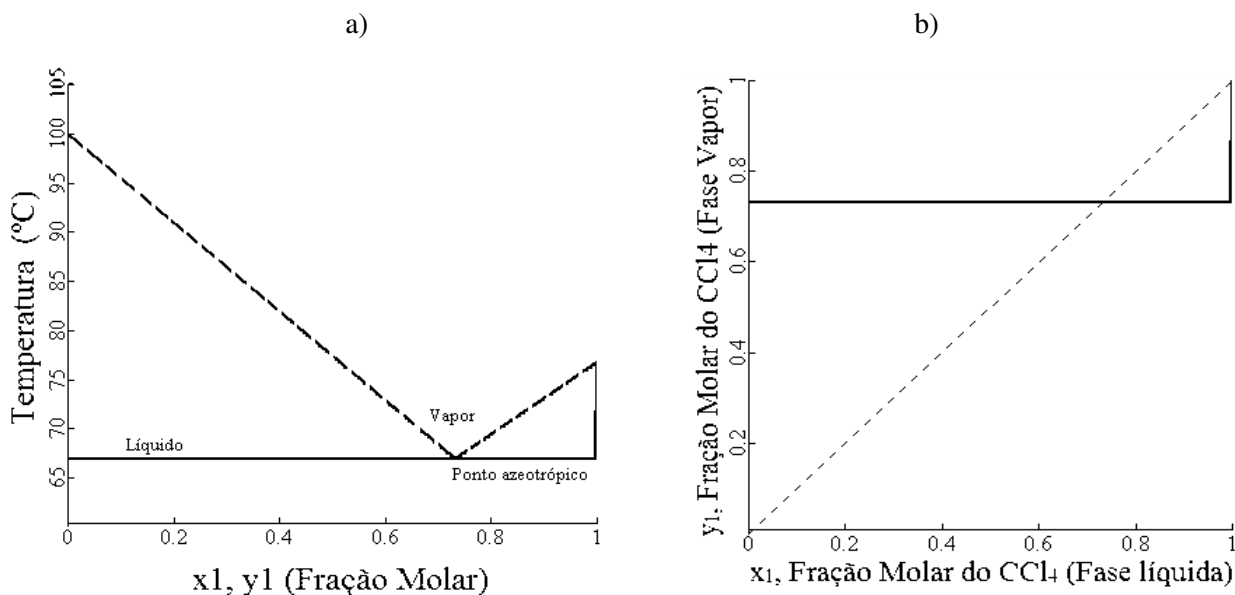


Figura 2.2. a) Diagrama Txy e b) Curva xy; Sistema (1) - (2) azeotrópica heterogênea.



De acordo com os gráficos das Figuras 2.1 e 2.2, que apresentam misturas azeotrópicas, são mostradas as regiões onde a composição da fase vapor é igual à composição global de uma ou mais fases líquidas, pois há um equilíbrio entre uma fase vapor e a(s) fase(s) líquida(s).

Tendo em vista os tipos de azeótropos e a variedade de propriedades físicas, algumas formas de conseguir a separação nesse tipo de mistura podem existir. Todas as técnicas especiais são baseadas na mesma diferença de composição das fases vapor e líquida, semelhante ao que ocorre na destilação convencional. Entretanto, contam com alguns mecanismos adicionais que modificam o comportamento do líquido-vapor dos componentes principais (chaves). Elas foram

aprimoradas e são classificadas de acordo com o efeito sobre a relação entre as composições das fases líquida e vapor, são elas: destilação por balanço de pressão, destilação extrativa, destilação com emprego de sais e destilação reativa (LUYBEN e CHIEN, 2010).

A destilação por balanço de pressão é a estudada nesse trabalho, porque ela causa a formação do azeótropo para alterar as características do ponto de ebulição dos componentes e as condições de separação da mistura. A destilação azeotrópica heterogênea é baseada no mesmo princípio que a destilação azeotrópica homogênea, mas o solvente adicionado é parcialmente miscível com um dos componentes. O solvente favorece a formação de duas ou mais fases líquidas, para em seguida ser recuperado através de um sistema de separação líquido-líquido, seguido de uma destilação convencional.

O comportamento da coluna de destilação em estudo neste trabalho apresenta características de uma coluna de destilação azeotrópica heterogênea, pois há em alguns estágios da torre formação de duas fases líquidas e conseqüente separação em um vaso decantador. Porém, não existe a adição de um solvente para modificar a volatilidade relativa dos compostos. Além disso, a coluna em estudo se apresenta como um processo de alta pureza.

2.2.2. Destilação de Alta Pureza

Colunas de destilação de alta pureza são processos os quais se deseja aumentar minimamente a pureza de um componente chave por separação (LUYBEN, 1992). Esses processos são sensíveis às perturbações, onde pequenas alterações na(s) corrente(s) de alimentação provocam mudanças drásticas nas concentrações dos produtos de base/topo em regime estacionário e dinâmico.

As colunas de destilação de alta pureza podem ser efetivamente controladas, se os distúrbios forem detectados rapidamente e se uma ação corretiva dinâmica for tomada a tempo de manter o processo perto do ponto de operação desejado. Segundo Wenxiang et al. (2010), quando se deseja aplicar técnicas de controle e otimização tradicionais em processos de destilação de alta pureza, as seguintes dificuldades têm de ser enfrentadas:

- ❖ O processo é geralmente não linear e com tempo de resposta longo;
- ❖ Há muitos distúrbios imensuráveis e é difícil manter o processo em estado estacionário;
- ❖ Concentrações da base e do topo são altamente acopladas.

Para superar essas dificuldades nesse tipo de processo alguns trabalhos foram desenvolvidos ao longo dos anos, sendo a maioria relacionado à análise transiente e controle.

Biswas et al. (2009) propuseram uma metodologia sistemática para a implementação de um controlador adaptativo realimentação linearizado, para uma coluna de destilação binária de alta pureza, com parâmetros incertos e de saturação na entrada. Os autores focaram no objetivo principal, que é controlar as composições de base e de topo da coluna com a presença de incertezas estruturadas e não estruturadas.

Segundo Srinivas *et al.* (1995), colunas de destilação de alta pureza representam um problema desafiador à indústria de processo, principalmente pela ausência de um modelo razoavelmente preciso, estes processos são bastante difíceis de controlar. Para uma coluna que separava metanol-etanol de alta pureza, os referidos autores propuseram a identificação e o controle da unidade através de modelos polinomiais auto regressivos não lineares com entradas exógenas (NARX), com isso verificaram as não linearidades inerentes ao processo.

Raríssimos trabalhos, durante o desenvolvimento dessa tese, foram encontrados sobre qual critério deve ser utilizado na seleção de variáveis para compor as entradas de um SV em colunas de alta pureza, se realmente as temperaturas e pressões com baixas variações devem ser incluídas no modelo de identificação utilizado e, principalmente, não foi encontrado nenhum trabalho que tinha como objetivo estimar compostos contaminantes (diluídos) em processos de alta pureza. Além disso, a maioria dos trabalhos encontrados na literatura foram aplicados em processos binários, que possuem menores restrições do que processos multicomponentes.

Exceção é o trabalho de Luo et al. (1995), que enfatizaram o uso de temperaturas em uma coluna de destilação multicomponente do etileno (alta pureza), para controle inferencial utilizaram um modelo Neuro-Fuzzy (FNN), onde a aplicação do método mostrou que o controle FNN foi muito superior ao controle de composição em comparação com uso do PID convencional. E o de Kano et al. (2009), que desenvolveram *soft sensors* altamente robustos, para isso utilizaram o método de identificação no subespaço de dois estágios (SSID), levando em consideração a influência das perturbações não medidas sobre as principais variáveis estimadas de uma fracionadora industrial de etileno (alta pureza). A principal contribuição dos trabalhos mencionados é o fato da possibilidade do uso de variáveis que apresentam pouca variabilidade (temperaturas e pressões), além das mais sensíveis (vazões), como entradas na construção dos sensores virtuais.

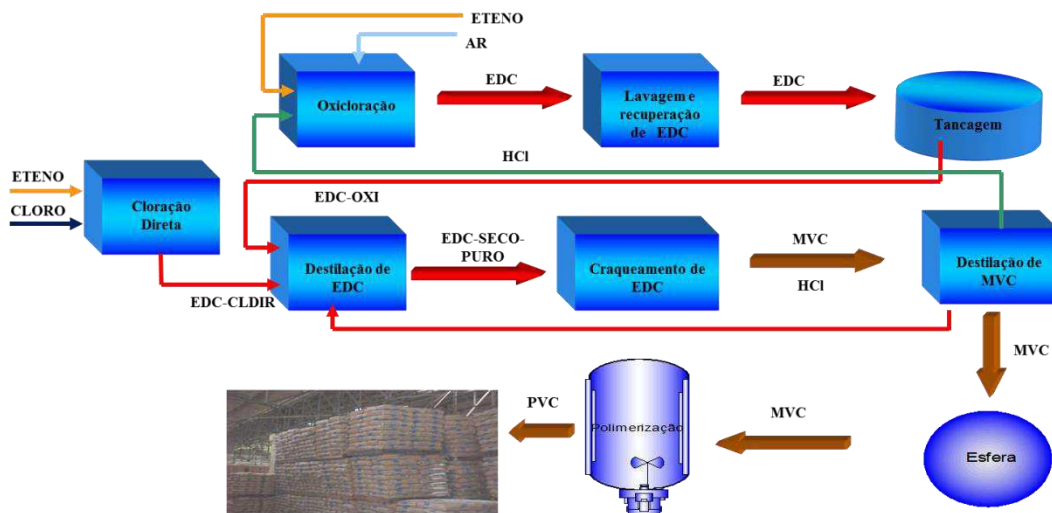
2.3. PROCESSO PRODUTIVO E PURIFICAÇÃO DO 1,2-DCE

O objeto de estudo desse trabalho se concentra na cadeia produtiva da indústria de cloro-soda, que se inicia com a eletrólise da salmoura. Nesta operação, a soda é coproduzida com o cloro, em uma proporção fixa de 1 (uma) tonelada de cloro para 1,12 toneladas de soda cáustica (RODOLFO JR et al., 2006).

Um dos principais produtos gerados a partir do cloro é o 1,2-DCE, matéria-prima básica para a produção da resina PVC. As duas principais rotas produtivas do 1,2 DCE são: a cloração direta e a oxicloração. Na cloração direta (reação líquida), o eteno (C_2H_4) reage com o cloro (Cl_2) para formar o 1,2-DCE e, utiliza-se cloreto férrico ($FeCl_3$) como catalisador, normalmente se processa em temperaturas na faixa de 50 a 70 °C, sob pressões de 4 a 5 atm.

A rota por oxicloração quando comparada com o processo de cloração direta é caracterizada por um maior custo de investimento, maior custo operacional e uma menor pureza do produto. No entanto, a rota produtiva por oxicloração do eteno é reconhecida pela necessidade da recuperação do cloreto de hidrogênio (HCl) gerado na pirólise do 1,2-DCE, produção do MVC em fornalhas, assim proporcionando uma tecnologia alternativa a produção do 1,2 DCE por cloração direta. O Fluxograma de produção do 1,2 Dicloroetano e do Monômero Cloreto de Vinila (MVC) é mostrado na Figura 2.3.

Figura 2.3. Processo produtivo do 1,2 Dicloroetano e do Monômero Cloreto de Vinila.



Fonte: Adaptada do RODOLFO JR et al. (2006) e do Fluxograma de Processo.

Na reação de oxicloração o C_2H_4 reage com o HCl na presença de oxigênio (O_2), proveniente do ar atmosférico, com o catalisador cloreto de cobre ($CuCl_2$) e em temperaturas

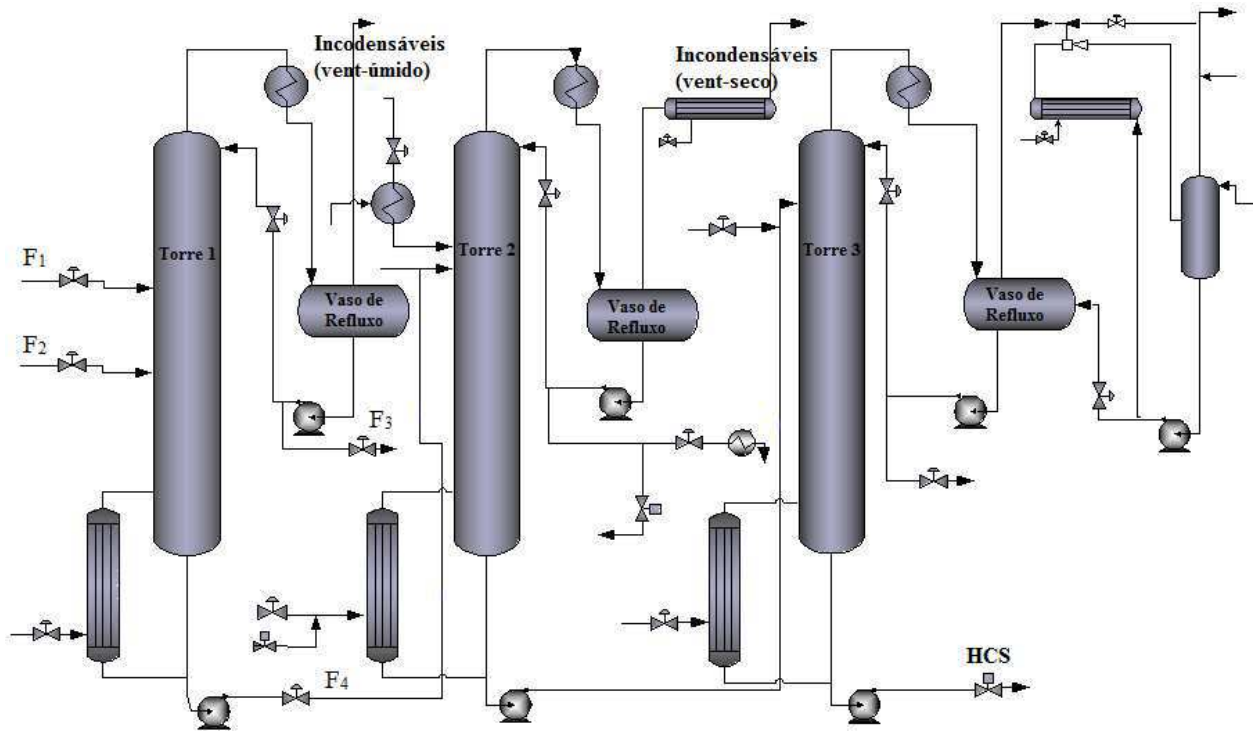
entre 250-350 °C. A maioria dos resíduos clorados são gerados na reação de oxicloração, principalmente pela corrente de HCl que alimenta o reator, haja vista que essa parcela contém uma pequena quantidade de acetileno (C_2H_2) altamente reativa. Outros subprodutos são formados pela reação do próprio C_2H_4 com o HCl, além da reação do etano (C_2H_6) com o Cl_2 .

As impurezas geradas são nocivas ao processo de craqueamento térmico do 1,2-DCE e, portanto, necessitam ser removidas (KIRK-OTHMER, 1997). Em destaque estão a formação das impurezas CCl_4 e $CHCl_3$, cujo teores aceitáveis a níveis de separação de compostos são especificados na etapa de purificação do 1,2-DCE. A conversão do 1,2-DCE para produção do MVC através do craqueamento térmico exige que o 1,2-EDC deva: estar seco (teor de água menor 10 ppm), a fim de evitar a corrosão excessiva dos equipamentos; e que tenha uma alta pureza, tipicamente maior que 99,5 % (peso/peso), haja vista que o processo de craqueamento é altamente suscetível à inibição e deposição por traços de impurezas. A presença de contaminantes na corrente que alimenta as fornalhas de craqueamento pode acarretar diversos efeitos, dentre os quais se destacam (DIMIAN e SORIN, 2008):

- ❖ Inibição parcial da reação, o que proporciona uma redução da conversão e instabilidade do processo à jusante;
- ❖ Acelera a formação de coque sobre as paredes do tubo das fornalhas, causando parada prévia do equipamento para remoção do coque (decoqueamento).

No processo em estudo, a purificação do 1,2-DCE é realizada em um conjunto de três torres de destilação, conforme apresenta o fluxograma da Figura 2.4, com o objetivo de remover água e orgânicos clorados (OC) de alto e baixo ponto de ebulição em relação ao 1,2-EDC. A alimentação da torre 1, chamada de coluna de secagem, é composta de 98 a 99% de 1,2-DCE saturado em H_2O , além de outras substâncias orgânicas cloradas. Apesar de ser menos volátil do que os compostos OC, a H_2O sai pelo topo da coluna de secagem devido à inversão da volatilidade relativa, principalmente, na região acima da primeira alimentação da coluna (F_1). Ainda sobre a torre 1, após o condensador de topo, no vaso de refluxo surgem duas fases líquidas: uma orgânica saturada em H_2O e uma aquosa saturada em orgânicos. A fase orgânica, composta basicamente por substâncias OC saturadas com H_2O dá origem ao refluxo da coluna e ao produto de topo (hidrocarbonetos clorados úmidos, HCU), enquanto a fase aquosa é descartada. Os componentes mais voláteis, em quantidade muito menor e que não condensam, seguem para serem incinerados.

Figura 2.4. Fluxograma esquemático do processo de purificação do 1,2-EDC.



Fonte: Adaptada do Fluxograma de Engenharia (PFD).

A corrente de base da coluna de secagem segue para a torre 2 (coluna de destilação propriamente dita), a qual também é alimentada pela corrente de base de outra coluna de destilação oriunda da recuperação do 1,2-DCE do craqueamento e pela corrente de topo da torre 3 (coluna a vácuo). A torre 3 tem como objetivo produzir uma corrente de topo isenta de componentes pesados para alimentar as fornalhas de craqueamento.

Devido ao seu efeito catalítico no processo de craqueamento, a presença do CCl_4 em concentrações menores ou iguais a 3000 ppm é desejada, haja vista que ele estabiliza a conversão do 1,2-DCE na pirólise e reduz potencialmente a formação de coque. Já a presença do CHCl_3 é indesejada, porque o mesmo é um promotor direto de coqueamento e sua concentração deve ser mantida abaixo de 400 ppm na base da Torre 1. Vale salientar que todo CHCl_3 e CCl_4 que alimentam a Torre 2 saem na corrente de topo, de modo que a definição quanto às quantidades desses componentes deve ser tomada antes da Torre 2, ou seja, na base da torre 1. Logo, um uma estimativa em tempo real das concentrações desses componentes na saída da torre 1 trará diversos benefícios ao processo.

A maneira tradicional de se acompanhar a concentração desses compostos é através do envio regular de amostras para análise de cromatografia em um laboratório qualificado e

certificado. A partir dos resultados, geralmente disponíveis de uma a três vezes ao dia, os operadores atuam no processo para corrigir eventuais desvios. A grande vantagem desse método é a simplicidade, devido à necessidade de se prever nos projetos apenas os pontos de amostragem para a coleta das correntes do processo (TREVISAN e POPPI, 2006).

No entanto, quando se deseja obter ou se aplicar controle em tempo real nas plantas químicas utilizando resultados dessa técnica surgem algumas desvantagens (CAMPOS et al., 2013): existe um atraso entre a coleta da amostra e o resultado do laboratório, que pode chegar a várias horas; este método não permite incluir as informações de qualidade dos produtos em um sistema automático de controle em tempo real; a precisão da medição pode ser afetada no procedimento utilizado pelos operadores durante a amostragem dos produtos na área, tais como limpeza de recipientes, erros sistemáticos de amostragem, etc.; e as análises de laboratório estão sujeitas a outros erros experimentais, considerados antes de se atuar na planta.

Sendo necessária uma frequência maior de resultados sobre as concentrações, de maneira que o operador tenha tempo hábil para corrigir o desvio da variável de processo em questão, poderia ser utilizada instrumentação analítica conectada diretamente com o processo, medindo a propriedade desejada em tempo real, são os chamados analisadores em linha (FORTUNA et al., 2007). No entanto, para atingir uma disponibilidade aceitável de um analisador em linha (> 95% para a maioria das aplicações), é necessário que haja técnico capacitado e disponível para a manutenção do sistema, previsão de materiais de consumo e peças de reposição, requerendo um alto custo de investimento inicial e de manutenção para efetuar essas medições.

Uma nova alternativa para estimar e superar essas dificuldades é o desenvolvimento de sensores virtuais que sejam capazes de fornecer em tempo real os valores de concentrações. Quando se deseja estimar valores de variáveis de processos de difícil medição ele é extremamente interessante e útil à indústria de processo, onde geralmente há um grande número de variáveis físicas medidas continuamente e rapidamente, que podem servir como sinais de entrada para essa tecnologia (BOLF et al., 2008).

2.4. SENSORES VIRTUAIS EM PROCESSOS DE DESTILAÇÃO

2.4.1. Metodologia de Construção

O desenvolvimento acelerado de métodos modernos em aspectos de monitoramento e de controle automático de processos proporcionam diversos benefícios às plantas industriais. Novos desafios estão constantemente focados no desenvolvimento da teoria de controle, métodos de

identificação, atuadores modernos, sensores inteligentes e na possibilidade de se aplicar métodos avançados. Ao mesmo tempo, espera-se que as plantas industriais apresentem maior eficiência e obedeçam as leis que estabelecem limites firmes de qualidade dos produtos e de emissões de poluentes.

Como resultado, um número maior de variáveis de processos necessita ser supervisionado com instrumentos de medições adequados. Entretanto, quando se deseja medir variáveis químicas e bioquímicas não facilmente mensuráveis, tais como, composição de produtos, biomassa e pH, a disponibilidade desses sensores específicos com medições em tempo real é escassa ou até inexistente (ZAMPROGNA et al., 2005; FORTUNA et al., 2005).

Os *soft sensors* surgiram para efetuar essas medições difíceis e desempenhar novos papéis. Ao longo desse trabalho será utilizada a nomenclatura de “sensor virtual (SV)”, no entanto na literatura são utilizados outros acrônimos, tais como: modelo inferencial, sensor por *software*, sensores inferenciais, analisador virtual *on-line* e sensores de observações. Os referidos termos são uma combinação das palavras "*software*", porque os modelos são programas de computador, e "*sensors*", porque os modelos fornecem informações semelhantes aos seus homólogos, *hardwares*.

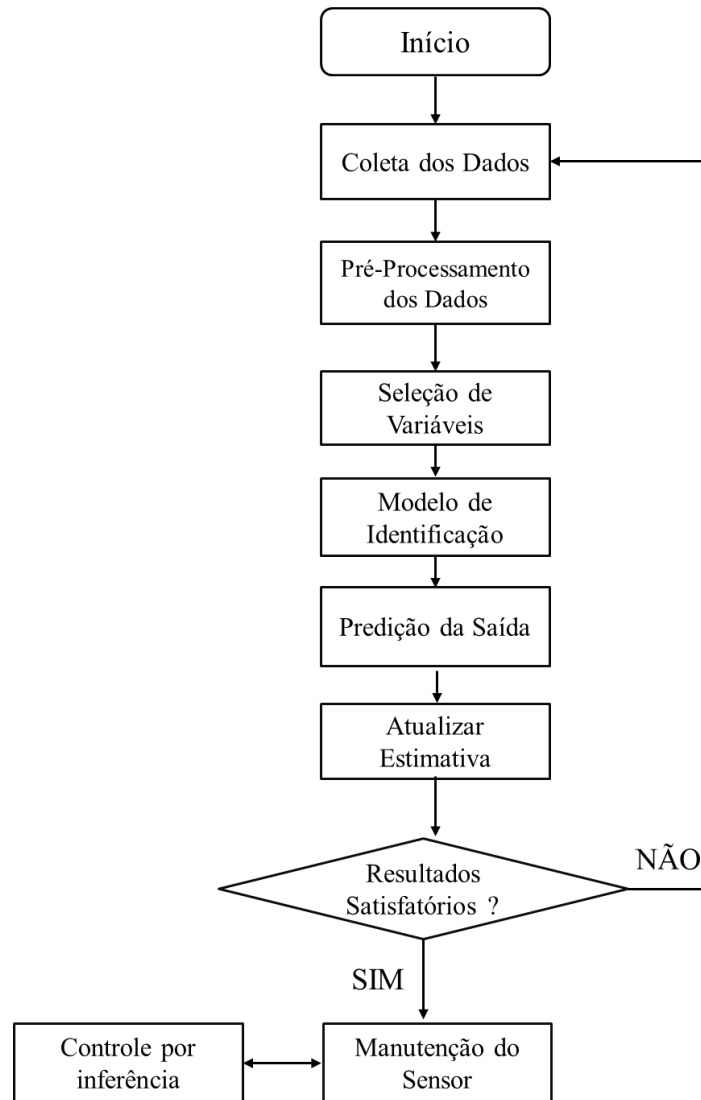
Conceitualmente um SV é um modelo matemático, implementado em *software*, utilizado para estimar variáveis de interesse de difícil medição em função de variáveis de entrada, ditas secundárias, de fácil medição. Esta alternativa surge de uma dificuldade operacional ou do alto custo na obtenção da variável desejada (MORAIS JR, 2011). Algumas das propriedades mais interessantes dos SV são descritas abaixo:

- ❖ São uma alternativa de baixo custo se comparados com sensores de *hardware* caros como, por exemplo, os analisadores em linha;
- ❖ Trabalham em paralelo com sensores de *hardwares* como, por exemplo, sensores de medições físicas, fornecendo informações úteis para tarefas de detecção de falhas;
- ❖ Ao serem implementados em um *hardware* existente como, por exemplo, micro-controladores, podem ser adaptados em casos de mudanças de parâmetros do sistema;
- ❖ Permitem a estimativa de dados em tempo real, superando os atrasos introduzidos pelos sensores de *hardware* com respostas lentas, por exemplo, cromatógrafos gasosos.

Portanto, qualquer grandeza que possa ser estimada por meio de algoritmos matemáticos poderá ser determinada por meio de um sensor virtual, como é o caso da composição dos

produtos de um processo de destilação, e seus resultados poderão ser utilizados em aplicações de controle, acompanhamento em tempo real ou diagnóstico da planta industrial. A Figura 2.5 descreve as etapas típicas e questões de prática comum no desenvolvimento de sensores virtuais.

Figura 2.5. Proposta de metodologia para construção de sensores virtuais.



Fonte: Adaptada de Fortuna et al. (2007), Lotufo e Garcia (2008) e Kadlec et al. (2009).

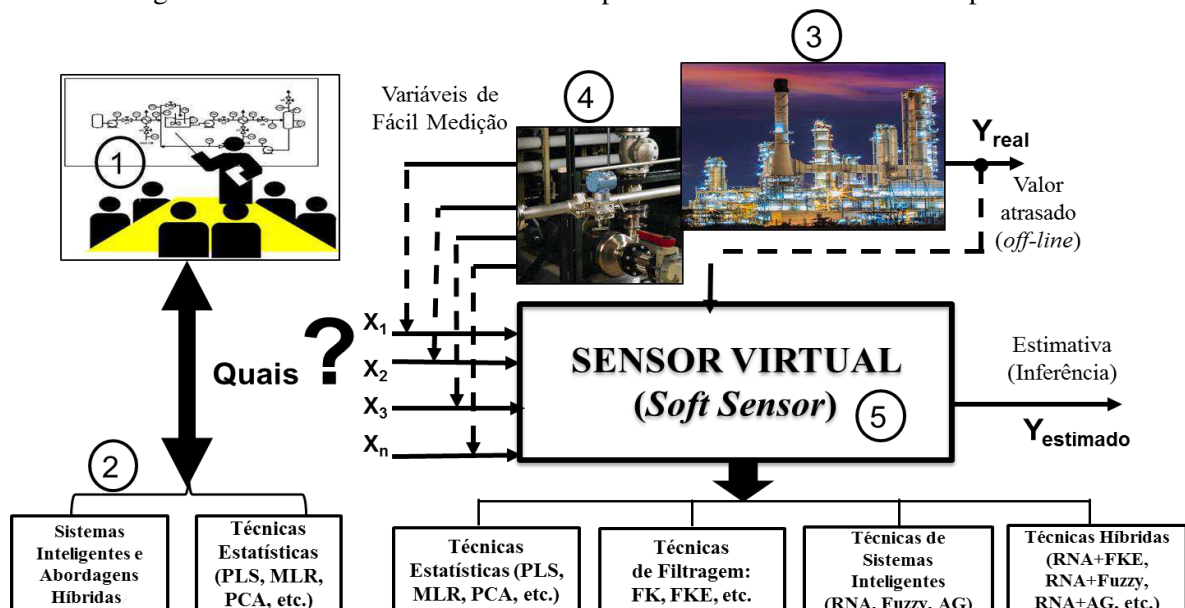
O fluxograma da Figura 2.5 representa uma adaptação de algumas das metodologias propostas para construção de SV. A metodologia em etapas apresentada traz uma abordagem geral, podendo ser aplicada em processos contínuos e batelada, bem como para qualquer um desses objetivos de: predições *on-line*, monitoramento de falhas de processos, detecção de falhas e reconstrução de sensores físicos. A primeira etapa do problema em questão trata, especificamente, da maneira como os dados devem ser coletados, por exemplo, através de sensores físicos instalados na planta e salvo em um banco de dados de um supervisor

industrial. Logo em seguida, etapa 2, verifica-se quanto à necessidade de tratamento dos dados armazenados. A etapa 3 é considerada por muitos autores que propuseram SV em colunas como, por exemplo, Ming-Da et al. (2009) e Zamproga *et al.* (2005), a fase mais importante do *soft sensor*, haja vista que o desempenho final do sensor é proporcional às escolhas das variáveis de entrada do modelo de identificação (etapa 4). O modelo de identificação deve ser capaz de fornecer a saída predita de maneira satisfatória ou a níveis aceitáveis. Geralmente, utiliza-se um critério de erros para avaliação de desempenho. Etapas adicionais podem ser inseridas como, por exemplo, uma maneira de proporcionar a manutenção ou aferição do SV (através análises de laboratório) ou até a implementação de controle inferencial.

2.4.2. Dados de Entrada para os Sensores Virtuais

A interatividade entre muitas das variáveis de processos químicos, por exemplo, em colunas de destilação, permite que o estado dessas variáveis reflita no estado de variáveis complexas como composição. Dessa forma, um conjunto de dados obtidos por análises laboratoriais, provenientes de sensores físicos instalados na própria planta ou através de um modelo matemático validado com dados reais, podem fornecer as entradas, tais como: temperaturas, pressões, níveis, vazões, e outras; com o objetivo de estimar as concentrações do produto (base ou topo) de uma torre em uma frequência adequada. Quando a simulação do processo é utilizada para geração de dados do SV, cujas entradas para este modelo são outras variáveis de controle medidas, saídas da planta e distúrbios mensuráveis, uma configuração representativa em é mostrada de acordo com a Figura 2.6.

Figura 2.6. SV com entradas fornecidas por um modelo matemático do processo.



A coleta dos dados das variáveis secundárias, etapa 1 da Figura 2.6, é a fase mais importante à obtenção da inferência de composição, pois se os dados forem ruins, os sensores virtuais também serão ruins (CAMPOS et al., 2013). Muitas vezes, os dados fornecidos necessitam da remoção de estados estacionários ou de um pré-processamento (remoção de erros grosseiros). No entanto, a influência de certas variáveis pode ser apenas parcialmente independente de outras, ou pode variar com a região operacional do processo. Então, para se obter o grupo de variáveis independentes (entradas) que irão compor o modelo, eliminam-se influências ambíguas. Nesse trabalho serão utilizadas técnicas de estatística multivariada para seleção de variáveis secundárias, por meio de modelos de regressão linear múltipla (MLR) e da análise de componentes principais (PCA).

Os seguintes métodos estatísticos para seleção ótima de variáveis de entrada podem ser utilizados: todas as regressões possíveis (*All Possible Regressions*), adição sequencial (do Inglês *Forward Selection*, FS), seleção por Eliminação (do Inglês *Backward Elimination*, BE), regressão por etapas (do Inglês *Stepwise Regression*, SR) e a regressão por etapas nas predições do modelo (do Inglês *Stepwise Regression on Model Predictions*, SRMP). A descrição de cada uma dessas técnicas e seus respectivos algoritmos são apresentados no Capítulo 4.

Dessa forma, deseja-se que o SV combine o conhecimento teórico do processo de destilação, por meio do modelo matemático, com o conhecimento prático do seu real funcionamento através das medições efetuadas. Se essas entradas que agem no sistema são conhecidas, satisfazendo as condições teóricas e se, além disso, o modelo é uma aproximação adequada do sistema real, então o sensor virtual estima o completo estado do sistema. Assim, se o modelo desempenhar as chamadas propriedades de observabilidade, o SV fornece estimativas em tempo real das variáveis chaves do processo (SIEGEL, 2001).

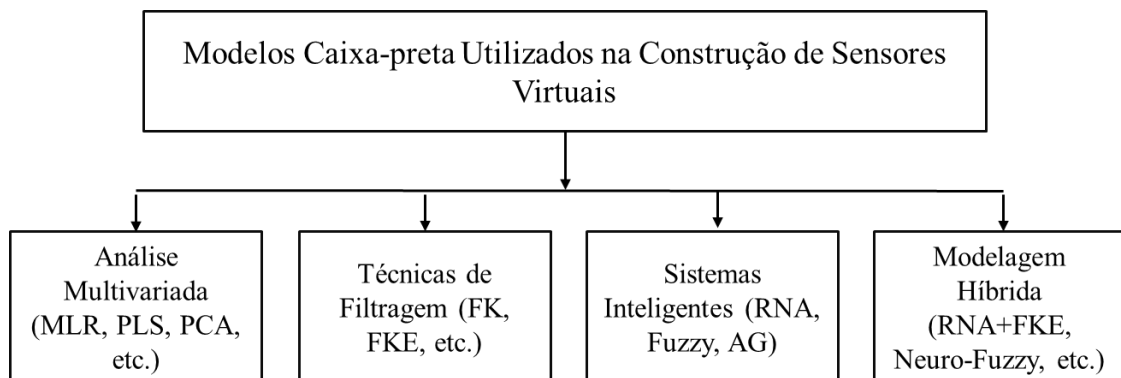
2.4.3. Modelos Aplicados na Construção de Sensores Virtuais

Muitas das técnicas, para a construção de SV, estão amplamente disponíveis e a seleção de métodos adequados é dependente do comportamento dinâmico do processo específico e da adequação prática ao modelo de aplicação (PIANG, 2005). Esses modelos podem ser formulados pelas teorias das equações dos princípios de conservações, chamados de modelagem caixa branca (*white-box*) ou modelos de princípios fundamentais (FPM), pelas abordagens empíricas através de identificação caixa preta (modelos *black-box*), ou ainda, utilizando os modelos por analogia caixa cinza (*grey-box*). Mas, apesar dos modelos FPM possuírem, geralmente, uma

faixa de validação mais ampla do que os demais modelos, pode ser difícil construir modelos dinâmicos de processos complexos, como é o caso coluna de destilação do 1,2-DCE em estudo. Outro problema é que com modelo FPM podem ser gerados modelos extremamente complexos com pouco valor prático, haja vista que muitas vezes não se conhecem com precisão todos os parâmetros desses modelos.

É possível observar na literatura que diversas técnicas caixa preta vem sendo empregadas na construção dos SV. Na Figura 2.7 são apresentadas as principais técnicas de identificação de sistemas utilizadas para tal objetivo. Isso se deve ao fato que os modelos orientados por dados são baseados nos históricos das próprias unidades de industriais, descrevendo as reais condições de operação do processo e sem a necessidade conhecimento fenomenológico do mesmo. A modelagem caixa preta quando é utilizada para construção de SV se mostra mais realista ao relacionar e descrever as condições reais do processo, principalmente, quando comparada com os sensores virtuais orientados a modelos FPM (KADLEC et al., 2009).

Figura 2.7. Técnicas de identificação empírica utilizadas para construção de analisador virtual.



Fonte: Adaptada de MORAIS JR (2011).

O modelo empírico mais simples é o de regressão linear múltipla (MLR), justamente por ser uma regressão linear pelo método de mínimos quadrados apresenta algumas limitações. Na construção de sensores virtuais é comum o uso de técnicas estáticas, como, por exemplo, o mínimo quadrado parcial (PLS) e a análise de componentes principais (PCA), para pré-processamento de dados e no treinamento de outros modelos inferenciais caixa preta.

O filtro de Kalman (FK) também vem sendo utilizado no algoritmo de construção de SV. Ele é um estimador otimizado de estados para sistemas lineares, sendo empregado em casos que se conhece o modelo do sistema e quando certas propriedades estocásticas estão disponíveis, sua

extensão é o filtro de Kalman estendido (FKE). Algumas vezes, o FK e o FKE são utilizados com outras técnicas de identificação para construção de sensores virtuais, por exemplo, com uma abordagem híbrida junto com a rede neural artificial.

Quando são utilizados modelos do tipo espaços de estados para construção de SV, se dá o nome de “observador de estados”. Estes modelos são mais aplicados a casos lineares, porque em casos não lineares se tornam muito complexos. Os mais utilizados são os de identificação do tipo entrada-saída como, por exemplo, o modelo auto regressivo com média móvel e entradas exógenas (ARMAX). Entretanto, com o crescente emprego das máquinas de vetores de suporte (SVM) pela comunidade de aprendizado científica, existem algumas aplicações recentes dessa técnica para SV. O sucesso do uso dessa técnica foi destaque em alguns trabalhos como, por exemplo, em YAN et al. (2004). No entanto, alguns problemas com as SVM foram relatados, especialmente com grandes conjuntos dados, haja vista que a complexidade computacional do processo de treinamento da SVM torna-se proporcionalmente restrita à quantidade de dados.

O fato de que seres humanos são aptos a conduzir tarefas complexas sob significativa incerteza tem estimulado pesquisas por padrões alternativos de modelagem, como as técnicas de sistemas inteligentes (SI) ou inteligência artificial (AI). Os SI são metodologias que empregam técnicas motivadas por sistemas biológicos e inteligência humana no desenvolvimento de modelos de sistemas dinâmicos. Dessa forma, elas vêm sendo muito empregadas em modelos de construção de SV, dentre as quais se destacam as redes neurais artificiais (RNA), a lógica *fuzzy*, o algoritmo genético A(G) e modelos híbridos.

As RNA consistem em um conjunto de neurônios artificiais interligados por conexões sinápticas. Ela está inserida dentro de uma área conhecida como sistemas inteligentes (conexionistas) ou inteligência computacional. Para substituir modelos que são demasiadamente complicados de serem resolvidos em tempo real, a RNA demonstra ter diversas habilidades (JANG et al. 1997; HIMMELBLAU, 2008). A aplicação de RNA na modelagem dinâmica e controle de processos químicos teve início no final de 1980 e entre os pioneiros destacam-se Bhat et al. (1990). Eles propuseram o uso das RNA para estimativa de pH em um reator CSTR. Os resultados mostraram que a rede tem capacidade de prever a variável nesse reator. Por possuir características não lineares, a RNA apresentou uma maior confiabilidade quando comparada com um modelo do tipo ARMAX. Por isso, o desenvolvimento de SV, para inferir a qualidade do produto (composição) em processos de destilação, a partir de outras variáveis do processo

facilmente medidas através de arquiteturas de RNA, é um tema que está sendo de grande aplicação.

Zanata (2005) apresentou um SV empregando RNA à estimativa da composição em uma coluna de destilação de mistura ideal multicomponente. Para aquisição dos dados de entrada do SV desenvolveu-se uma simulação dinâmica da torre no *software* Simulink[®]. A principal contribuição do trabalho relaciona-se ao estudo sobre a influência do treinamento parcial no desempenho da rede e o estudo realizado sobre os principais erros que podem ocorrer nos sensores virtuais desenvolvidos com RNA, raramente tratado em outras publicações. Kano et al. (2009) utilizaram o método de identificação de subespaços de duas fases (SSID), com isso desenvolveram SV precisos que levaram em conta a influência das perturbações não medidas sobre as principais variáveis estimadas. Os autores citam que a principal vantagem do SSID é que se podem estimar perturbações não medidas sem as simplificações que a técnica convencional de filtro Kalman faz. Os autores afirmaram que o SSID é robusto aos ruídos brancos Gaussianos inerentes ao processo, já que o SV foi aplicado para uma coluna de etileno industrial de alta pureza. A principal contribuição do trabalho relaciona-se ao fato dos autores utilizarem medições secundárias no estudo, tais como, medições de quatro temperaturas ao longo da torre e a composição de um dos compostos na alimentação, cujo objetivo foi estimar a concentração do etano no produto de etileno. As observações relatadas pelos autores são importantes na tomada de decisão de quais variáveis candidatas devem ser incluídas em um estudo de seleção variáveis em colunas de alta pureza, haja vista que normalmente temperaturas e pressões apresentam pequenas variações nesses processos. Além disso, os autores fazem uso da composição do composto a ser inferido na corrente de alimentação, o que pode se tornar dispensável porque essa é também um distúrbio não medido.

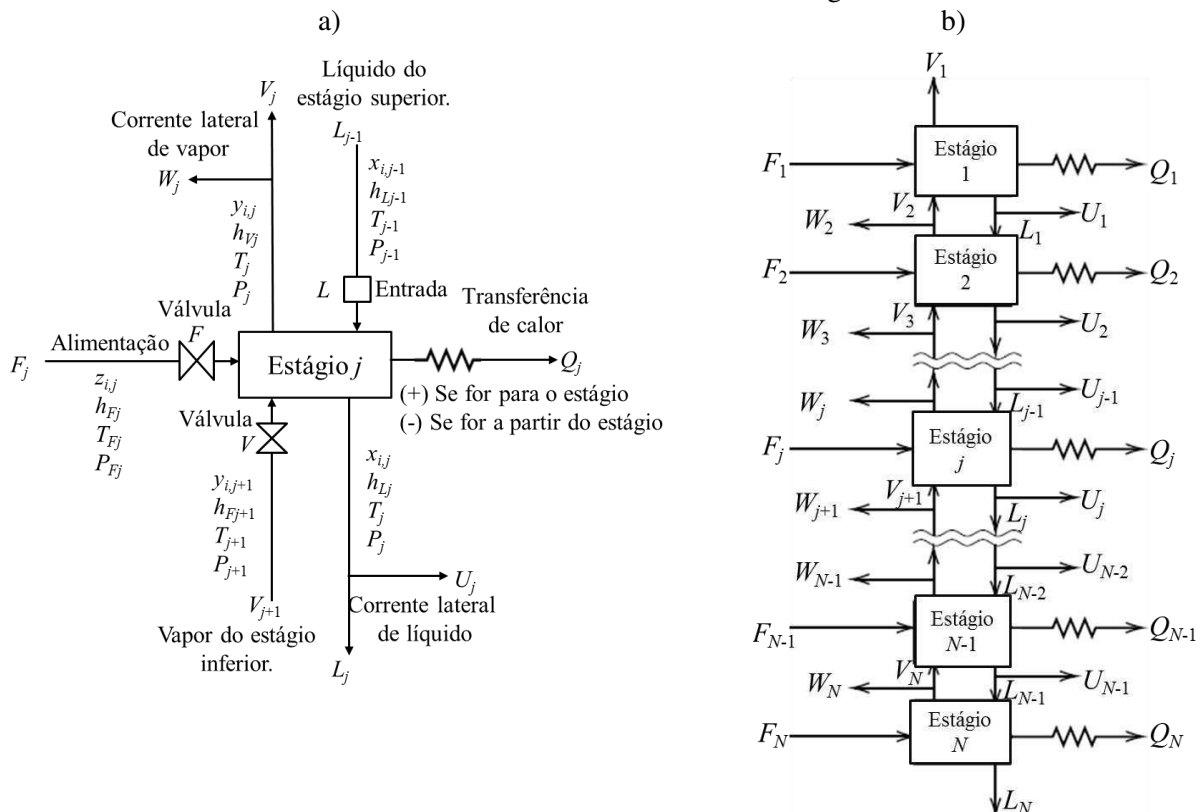
Conforme apresentado anteriormente, alguns trabalhos de sensores virtuais que utilizam RNA ,para estimar a qualidade do produto, podem ser encontrados na literatura. Entretanto, não foi encontrado até aqui temas que tratam da estimativa das impurezas em colunas multicomponentes de alta pureza. As diversas características positivas quando se utilizam redes neurais em processos não lineares, por exemplo, a adaptação por experiência, a capacidade de aprendizado, a habilidade de generalização, organização dos dados, tolerância a falhas e armazenamento distribuído, motivaram o emprego dessa técnica nesse trabalho.

2.5. MODELAGEM MATEMÁTICA DE ESTÁGIOS EM EQUÍLIBRIO

Nesta seção serão apresentadas as principais equações e o método de resolução do modelo matemático e termodinâmico em processos de separação por estágios. Essas equações são utilizadas na modelagem e simulações estacionárias e dinâmicas da coluna de destilação do 1,2-DCE, de modo que sejam ajustadas para o objetivo em questão, metodologia de construção dos SV, onde seus resultados tendem a ser confiáveis mesmo quando feitas extrapolações (FORTUNA et al., 2007).

A modelagem e simulação de colunas de destilação não convencionais é mais indicada por meio de *softwares* de simulação orientados a módulos, como o *Aspen Plus* e *Dynamics*TM. Esses simuladores utilizam sub-rotinas para representar uma unidade inteira de processo, simulam vários tipos de configurações, uma parte do processo ou a planta inteira, e requerem que o programador forneça valores numéricos dos parâmetros e chamadas de sub-rotinas aos módulos apropriados (SMITH e CORRÍPIO, 2008). Considere uma coluna de destilação com N estágios, numerados de cima para baixo, o esquema representativo de um prato j é apresentado na Figura 2.8 a) e uma cascata de separação em contracorrente representada na Figura 2.8 b).

Figura 2.8. a) Representação de um estágio em equilíbrio em regime estacionário. b) Esquema geral de uma cascata em contracorrente com N estágios.



Fonte: Adaptada de SEADER et al. (2011).

Nas Figuras 2.8 a) e b) a alimentação entrando no estágio j pode ser um líquido, um vapor ou uma mistura de líquido e vapor. O sinal da carga térmica Q será negativo se o calor estiver sendo retirado do prato e positivo caso contrário. O cálculo preciso de uma coluna de destilação multicomponente requer a determinação rigorosa da temperatura, das vazões, composições das correntes e da taxa de transferência de calor em cada estágio. Esse objetivo é alcançado através da resolução do balanço material, de energia e das relações de equilíbrio em cada estágio.

O desenvolvimento do modelo rigoroso matemático para uma coluna de destilação é baseado no conceito de estágio de equilíbrio e as expressões representativas formam um conjunto de equações básicas: balanços de massa (M), as relações de equilíbrio (E), o somatório das frações molares (s) e balanços de energia (H), normalmente designadas de equações MESH. De acordo com a Figura 2.8, onde os balanços são realizados, se obtêm as equações MESH (Equações de 2.3 até 2.8):

Equações M (Equações 2.3 e 2.4): balanço de Massa e para o componente i no prato j (C equações para cada estágio):

Para $i = 1, 2, \dots, C$; e $j = 1, 2, \dots, N$.

$$\frac{dM_j}{dt} = F_j + L_{j-1} + V_{j+1} - (L_j + U_j) - (V_j + W_j). \quad (2.3)$$

$$\frac{d(M_j x_{i,j})}{dt} = F_j z_{i,j} + L_{j-1} x_{i,j-1} + V_{j+1} y_{i,j+1} - x_{i,j} (L_j + U_j) - y_{i,j} (V_j + W_j). \quad (2.4)$$

Equação E (Equação 2.5): relações de Equilíbrio de fases (C equações para cada estágio):

$$E_{i,j} = y_{i,j} - K_{i,j} \cdot x_{i,j} = 0. \quad (2.5)$$

Equações S (Equações 2.6 e 2.7): Somatório das frações molares, uma equação para cada estágio:

$$(S_y)_j = \sum_{i=1}^C y_{i,j} - 1 = 0. \quad (2.6)$$

$$(S_x)_j = \sum_{i=1}^C x_{i,j} - 1 = 0. \quad (2.7)$$

Equação H (Equação 2.8): balanço de Energia no prato j , uma equação para cada estágio:

$$d \frac{(M_j h_j)}{dt} = F_j \cdot h_{F,j} + L_{j-1} \cdot h_{L,j-1} + V_{j+1} \cdot h_{V,j+1} - (L_j + U_j) \cdot h_{L,j} - (V_j + W_j) \cdot h_{V,j} - Q_j. \quad (2.8)$$

São funções: $K_{i,j} = K_{i,j}\{T_j, P_j, x_j, y_j\}$, $h_{V_j} = h_{V_j}\{T_j, P_j, y_j\}$ e $h_{L_j} = h_{L_j}\{T_j, P_j, x_j\}$. Se estas relações não são contabilizadas como equações e estas três propriedades não são contabilizadas como variáveis, cada estágio de equilíbrio é definido unicamente por $2C + 3$ equações MESH.

A determinação do número de graus de liberdade do sistema em questão é realizada considerando-se o esquema de cascata em contracorrente com N estágios, Figura 2.8 b), cujo número de equações (N_E) é igual a $N(2C+3)$ e o número de variáveis (N_V) é dado por $[N(3C+10) + 1]$. Se N e todos os F_j , $z_{i,j}$, T_{F_j} , P_{F_j} , P_j , U_j , W_j e Q_j são especificados, o modelo é então representado por $N(2C+3)$ equações algébricas simultâneas, no caso de regime estacionário, e por $N(2C+3)$ variáveis desconhecidas (saídas), compreendendo $x_{i,j}$, $y_{i,j}$, L_j , V_j e T_j as variáveis a serem calculadas.

No caso de regime estacionário, as taxas de acúmulo das Equações 2.3., 2.4 e 2.8 são nulas, ou seja, igualadas a zero. Como resultado as equações M, E e H formam um conjunto de equações não lineares que devem ser resolvidas por técnicas iterativas, onde o número de graus de liberdade do sistema ($N_L = N_V - N_E$) é, portanto, igual a 2 (dois), de modo que têm-se que especificar duas variáveis ou arranjar duas equações para tornar o sistema determinado.

As equações MESH aplicam-se a cada um dos N estágios em equilíbrio de uma coluna de destilação, representada na Figura 2.8 b), ou seja, para $j = 1, 2, \dots, N$. O condensador (total ou parcial) é o andar 1, o refeedor (total ou parcial) é o andar N , o líquido L_N é a corrente de resíduo (para condensador total), $V_1 = 0$ é a sangria e U_1 é o destilado.

2.5.1. Modelagem Termodinâmica

A modelagem termodinâmica (complexa ou aproximada) é de suma importância nos resultados de uma simulação em colunas de destilação, uma vez que toda operação de separação do processo em estudo está embasada no equilíbrio das fases envolvidas.

Alguns trabalhos trazem discussões acerca de qual modelo deve ser utilizado para o cálculo do coeficiente de atividade da fase líquida (γ), (Wilson, NRTL, UNIFAC ou UNIQUAC). POLING et al. (2001) defendem a aplicação da equação UNIQUAC. Smith et al. (2007) e Halder (2014) indicam o uso da equação de NRTL para o equilíbrio líquido-líquido.

Independentemente da equação a ser utilizada, deve-se ter muito cuidado ao determinar os parâmetros a partir dos dados experimentais.

O complexo modelo termodinâmico pode incluir qualquer tipo de modelo inclusive os baseados nas equações de estado e os baseados na energia livre de Gibbs para predição do coeficiente de atividade da fase líquida (citados no parágrafo anterior). Estes modelos são utilizados para gerar os parâmetros do modelo termodinâmico aproximado. Em geral, os modelos termodinâmicos rigorosos são representados conforme as Equações 2.9, 2.10 e 2.11:

$$K_{i,j} = K_{i,j} \{P_j, T_j, x_j, y_j\} \quad (2.9)$$

$$h_{vj} = h_{vj} \{P_j, T_j, y_j\} \quad (2.10)$$

$$h_{Lj} = h_{Lj} \{P_j, T_j, x_j\} \quad (2.11)$$

Neste trabalho o ajuste do modelo termodinâmico, cálculo do coeficiente de atividade da fase líquida e da fugacidade da fase vapor, será feito através da comparação dos diversos modelos termodinâmicos disponíveis no simulador. O intuito é o de verificar as respostas e aproximar a solução do modelo quanto à concentração do DCE, CCl_4 e do CHCl_3 (corrente de topo e de base da coluna do 1,2-DCE em estudo).

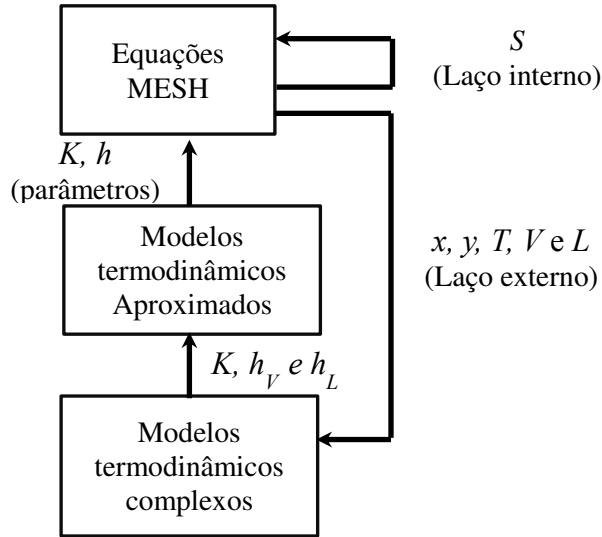
2.5.2. Solução Numérica das Equações MESH

Na resolução estacionária do modelo de equilíbrio os termos de acúmulo das Equações 3.1, 3.2 e 3.6 são igualados a zero, resultando em um conjunto de equações algébricas não lineares. Assim, os simuladores necessitam de uma técnica iterativa para encontrar o valor das variáveis dependentes com o objetivo de satisfazer o valor desejado para a variável especificada. Logo, na resolução das equações MESH existem alguns métodos que já conhecidos e utilizados na literatura, como os métodos: *Bubble-Point* (BP), *Sum-Rates* (SR), *Newton-Raphson* e o *inside-out*.

O simulador *Aspen Plus*TM, aqui empregado, utiliza a técnica *inside-out* proposta Boston e Sullivan (1974) para promover essa convergência. Isso se deve ao fato que os métodos BP e SR requeem um maior esforço computacional, principalmente nos cálculos dos valores das constantes de equilíbrio ($K_{i,j}$) e entalpias das fases líquida e vapor, particularmente quando são utilizados modelos rigorosos para o cálculo das propriedades termodinâmicas.

O algoritmo de *inside-out* se baseia em resolver dois problemas da modelagem: o primeiro (1º) é a determinação dos parâmetros termodinâmicos utilizando um “laço externo” e o segundo (2º) é a resolução das equações MESH em um “laço interno”, conforme a Figura 2.9.

Figura 2.9. Esquema de resolução das equações MESH pelo método *Inside-out*.



Fonte: Adaptada de SEADER et al. (2011).

De acordo com a Figura 2.9 as equações MESH são resolvidas pelo laço interno utilizando os parâmetros do modelo termodinâmico aproximado, resolvendo os valores de K (razão de equilíbrio entre fases) e h que é a entalpia. Assim, as equações MESH são sempre resolvidas no *loop* interno com o modelo aproximado. Os parâmetros das equações empíricas do modelo simples são atualizados no *loop* externo a partir do modelo rigoroso. A presença dos *loops* internos e externos é o que diferencia o método de *inside-out* dos demais e confere a esta classe de métodos a denominação de métodos *inside-out* (SEADER et al., 2011).

Ao serem considerados os termos de acúmulos de massa e energia em cada estágio nas equações M, E e H, temos então o modelo dinâmico de estágio de equilíbrio da coluna de destilação multicomponente. Diversos algoritmos têm sido desenvolvidos para solução dessas equações transientes. Boston et al. (1974), sugerem uma variação do algoritmo *inside-out* fazendo uma distinção entre as variáveis do laço externo (equações diferenciais) e as variáveis do laço interno (equações algébricas). Alternativamente é possível resolver o sistema de equações diferenciais ordinárias de forma simultânea, e separadamente resolver os subconjuntos de equações algébricas (GANI et al., 1986).

A simulação desenvolvida neste trabalho é realizada com auxílio do simulador *Aspen*

*Dynamics*TM. Durante a simulação o *holdup* da fase vapor é desprezado, devido ao fato da coluna operar em baixa pressão. É utilizada a equação de Francis (Equação 2.12) para obtenção da relação entre o *holdup* líquido do prato e a vazão de líquido que deixa este prato.

$$Q_L = K_w L_w h_c^{1.5}. \quad (2.12)$$

Onde: Q_L é o fluxo volumétrico líquido do estágio; K_w é a constante do vertedouro; L_w é o comprimento total do vertedouro; e h_c é a altura de transbordamento de líquido no vertedouro.

A altura do transbordamento (h_c), Equação 2.13, é a diferença entre a altura do líquido no prato (h_t) e a altura do vertedouro (h_w). A altura do líquido, Equação 2.14, é a razão do *holdup* volumétrico líquido (M_v) pela área ativa do prato (A).

$$h_c = h_t - h_w. \quad (2.13)$$

$$h_t = \frac{M_v}{A}. \quad (2.14)$$

Um dos métodos de integração utilizado pelo simulador *Aspen* é o de “Euler implícito com passo variável”, que combina a vantagem da velocidade convergência com a robustez de integradores de passos variáveis. O método numérico de resolução de sistemas de equações algébrico-diferenciais (EDA), utilizado em Euler implícito, consiste na solução da Equação 2.15.

$$y(t+h) = y(t) + h \frac{dy(t+h)}{A}. \quad (2.15)$$

Onde: h é incremento do passo de integração; $y(t)$ é o valor da variável de estado no tempo t ; e $y(t+h)$ e $\frac{d}{dt} y(t+h)$ são, respectivamente, o valor e a derivada da variável de estado no tempo $(t+h)$.

As vantagens dos métodos integrativos à resolução do conjunto EAD, utilizado pelo simulador *Aspen Dynamics*TM, são: variam o tamanho dos passos durante a simulação; mantém a exatidão com a estratégia do controle de erro; e consomem menos tempo computacional, fornecendo resultados mais exatos às condições altamente transientes em comparação com os métodos de passo fixo (SMITH e CORRÍPIO, 2008).

2.6. CONTROLE DE COMPOSIÇÃO EM DESTILAÇÃO

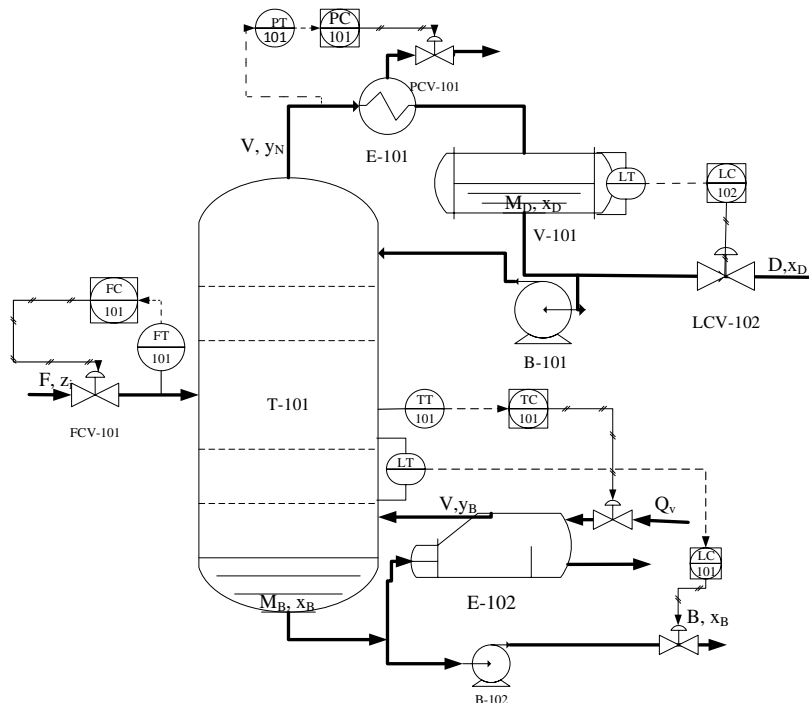
Em processos de destilação, desde processos ideais simples binários até sistemas complexos não ideais multicomponentes, existem dois graus de liberdade a serem fixados. Logo,

existem apenas duas variáveis adicionais que podem (e devem) ser controladas em uma coluna de destilação. Segundo Luyben et al. (1992), as duas variáveis que são escolhidas para serem controladas dependem de vários fatores, de acordo com situações comuns:

- ❖ Controle de composição do componente leve-chave no topo ou base da coluna e da composição do componente pesado-chave no topo ou base da coluna;
- ❖ Controle de temperatura na seção de retificação e da temperatura na seção de esgotamento da coluna;
- ❖ Controle de nível do vaso refluxo e de alguma temperatura ao longo da coluna;
- ❖ Controle de nível de base do refeedor e de alguma temperatura na coluna;
- ❖ Controle da vazão de refluxo e de uma temperatura próxima ao topo da coluna.

Existe um número grande de possibilidades para a estrutura do sistema de controle de colunas de destilação. Estruturas convencionais apenas explicitam as variáveis manipuladas (MV) utilizadas no controle da qualidade da separação (variável de processo-PV). A Figura 2.10 apresenta um esquema de um sistema típico de controle em uma coluna de destilação.

Figura 2.10. Sistema de controle convencional em processos de destilação.



Estes exemplos ilustram que unicamente duas variáveis devem ser controladas: pelo menos uma composição ou uma temperatura em algum lugar na coluna (LUYBEN, 1992). Para fins de controle, é necessário então, selecionar o par ideal de PV-MV o que é feito geralmente

por meio de uma análise de sensibilidade em regime estacionário na malha. A análise de sensibilidade é importante para selecionar essas variáveis e também obter dados do processo antes que ocorra à mudança para o regime transiente. Além disso, pode-se ainda obter informações sobre os estágios da coluna onde apresentam irregularidades no processo, como uma variação de temperatura ou composição não esperada.

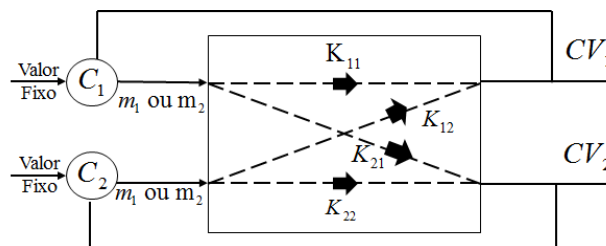
2.6.1. Técnicas de seleção de pares de variáveis ou Sensibilidade da Malha

Algumas colunas de destilação usam a estrutura de controle “ideal”, ou seja, os dois graus de liberdade que estão concebidos no processo são normalmente especificados para o produto de fundo e para o produto de destilado. Logo, a estrutura “ideal” de controle iria medir as composições dos dois produtos (base e topo) e manipular duas variáveis de entrada (por exemplo, vazão de refluxo e entrada de calor refeedor) para manter os níveis desejados do componente de interesse nas duas correntes de produto (LUYBEN, 2013; SEBORG et al., 2011).

Algumas configurações de controle são propostas para processos multivariáveis utilizando medições de temperatura ou composição. O controle simultâneo de duas composições ou temperaturas é denominado de controle dual (*dual composition control*). Segundo Kalid (2009), esta estrutura de controle, quando funciona, é a mais indicada. Entretanto, em muitos casos, não é possível controlar as duas malhas, então apenas uma variável é controlada, controle singular (*single composition control*). Além disso, em processos de alta pureza pode ocorrer transientes mínimos de temperaturas e pressões, o que pode inviabilizar a implementação de controle *single ou dual composition*.

A utilização de controladores Proporcional-Integral-Derivativo (PID), C_1 e C_2 , resulta em uma estrutura descentralizada de múltiplas malhas, a Figura 2.11 representa um sistema 2x2, m_j representa as variáveis manipuladas de um processo, CV_i as possíveis variáveis controladas e K_{ij} os ganhos estacionários do mesmo.

Figura 2.11. Controle realimentado descentralizado multivariável (2 x 2).



Fonte: Adaptada de SMITH e CORRIPIO (2008).

Em processos multicomponentes, as temperaturas dos estágios fornecem informações precisas sobre as concentrações dos compostos de interesse. Além disso, os sensores de temperatura são baratos, apresentam boa confiabilidade e um rápido tempo de resposta. Entretanto, como no caso de processos de alta pureza em estudo, o uso de temperaturas para inferência ou controle de composição deve ser avaliado/investigado por algum critério base. Piang (2005) empregou um dos critérios, que serão citadas a seguir, para investigar quais temperaturas deveriam ser utilizar como entradas para a construção de um SV em uma coluna.

Existem alguns critérios ou métodos matemáticos que são utilizados na análise de sensibilidade de malha em processos de destilação, são eles: método de pratos sucessivos, simetria de sensibilidade, matriz de sensibilidade, resposta à variação na composição de alimentação, decomposição de valores singulares (SVD) da matriz de sensibilidade, além de outros, como a matriz de ganho relativo (RGA), índice de resiliência de Morari (MRI), número condicional (CN) e índice de Niederlinski (NI).

Segundo Luyben (2013), as duas técnicas mais utilizadas são a análise da SVD e da RGA. Essas duas são ferramentas para análise de sensibilidade e controlabilidade em malha de controle, ambas necessitam do conhecimento da matriz de ganho estacionário (SGA) do processo, dada pela Equação 2.15.

$$K_{n \times m} = \begin{bmatrix} \left. \frac{\partial PV_1}{\partial MV_1} \right|_{M,j \neq 1} & \left. \frac{\partial PV_1}{\partial MV_j} \right|_{M,j \neq j} & \left. \frac{\partial PV_1}{\partial MV_m} \right|_{M,j \neq m} \\ \left. \frac{\partial PV_2}{\partial MV_1} \right|_{M,j \neq 1} & \left. \frac{\partial PV_2}{\partial MV_j} \right|_{M,j \neq j} & \left. \frac{\partial PV_2}{\partial MV_m} \right|_{M,j \neq m} \\ \left. \frac{\partial PV_n}{\partial MV_1} \right|_{M,j \neq 1} & \left. \frac{\partial PV_n}{\partial MV_j} \right|_{M,j \neq j} & \left. \frac{\partial PV_n}{\partial MV_m} \right|_{M,j \neq m} \end{bmatrix}. \quad (2.15)$$

Em um estudo quantitativo do efeito das variáveis manipuladas sobre as variáveis controladas de um sistema distribuído (as propriedades do mesmo são função da posição espacial) é necessário conhecer a SGA em malha aberta, que por definição $K_{n \times m}$ é a matriz de sensibilidade do sistema para as n variáveis medidas ou controladas e m variáveis manipuladas; PV_i = Variável de processo medida e/ou controlada i ; e MV_j = Variável manipuladas. As colunas da matriz K estabelecem como cada variável controlada ou medida sofre a influência de cada variável manipulada. A interação entre as PV_i e as MV_j pode ser avaliada a partir dos valores singulares da matriz sensibilidade, fornecendo o grau de acoplamento entre as variáveis do processo.

O método da SVD é muito útil em análises e projetos de sistemas de controle de processos industriais que tendem às anomalias durante a operação (SMITH e CORRÍPIO, 2008). Assim, as análises feitas em processos industriais reais com a SVD se mostraram extremamente satisfatórias e ajudam tanto na implementação do controle, quanto na escolha de algumas variáveis no estudo de inferência. Proposta para aplicações industriais por Moore (1980), a SVD é utilizada na análise de sensibilidade de sistemas multivariáveis e baseia-se no seguinte teorema da álgebra linear: a matriz $K_{n \times m}$ pode ser decomposta no produto de três outras matrizes, conforme Equação 2.16.

$$K_{n \times m} = U S V^T. \quad (2.16)$$

Sendo: $K_{n \times m}$ a matriz de sensibilidade SGA das PV_i pelas m MV_j ;

U uma matriz unitária, $n \times n$, ou ortonormal e suas colunas são os vetores singulares à esquerda da matriz K , que são os vetores característicos da matriz $K \cdot K^T$;

V uma matriz unitária $m \times m$, onde suas colunas são os vetores singulares à direita da matriz K , que são os vetores característicos da matriz $K^T \cdot K$;

S uma matriz diagonal $n \times m$, caracterizada pelos valores singulares (σ_i) de K .

Os valores singulares (σ_i), dados pela matriz S , são organizados de forma decrescente, isto é, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m \geq 0$, fornecendo ganho ideal dissociado do processo em malha aberta. A razão do maior valor singular pelo o menor $\sigma_{\max}/\sigma_{\min}$ é o chamado Numero Condicional (NC) da matriz de ganho e é uma medida da dificuldade do problema do controle multivariável. Os valores singulares são dados pela Equação 2.17:

$$\sigma = +\sqrt{\lambda(K^T K)} \text{ com } i = 1, 2, \dots, n. \quad (2.17)$$

Sendo: λ o valor característico (autovalor) de K .

O método da SVD quando é aplicado a um processo em regime estacionário e a análise é realizada através da decomposição da SGA, as conclusões sobre a sensibilidade do processo são coerentes, haja vista que os valores singulares possuem um significado físico muito forte. Essa importância física permite uma maior clareza sobre os problemas envolvendo a medição/inferência/controle à natureza física desses problemas. Além disso, maneiras de atuar no controle desses processos podem estar disponíveis.

Entretanto, o uso da SGA, considera apenas o regime estacionário das variáveis das diversas variáveis de processos, que podem assumir unidades diferentes. Já a análise de sensibilidade efetuada pela Matriz de Ganho Relativo (RGA), proposta por Bristol (1966), fornece o grau de interação entre as malhas de controle de processos e qual o par mais efetivo das variáveis controladas e manipuladas. A matriz de ganho em malha fechada $\mathbf{K}'_{n \times m}$ é calculada pela Equação 2.18.

$$\mathbf{K}'_{n \times m} = \begin{bmatrix} \left. \frac{\partial PV_1}{\partial MV_1} \right|_{C,i \neq 1} & \left. \frac{\partial PV_1}{\partial MV_j} \right|_{C,i \neq 1} & \left. \frac{\partial PV_1}{\partial MV_m} \right|_{C,i \neq 1} \\ \left. \frac{\partial PV_2}{\partial MV_1} \right|_{C,i \neq 2} & \left. \frac{\partial PV_2}{\partial MV_j} \right|_{C,i \neq 2} & \left. \frac{\partial PV_2}{\partial MV_m} \right|_{C,i \neq 2} \\ \left. \frac{\partial PV_n}{\partial MV_1} \right|_{M,i \neq n} & \left. \frac{\partial PV_n}{\partial MV_j} \right|_{C,i \neq n} & \left. \frac{\partial PV_n}{\partial MV_m} \right|_{C,i \neq n} \end{bmatrix}. \quad (2.18)$$

Ao contrário da SVD, a RGA não é sensível às unidades dos ganhos de processo. Isso acontece por que, para cada um dos ganhos de malha aberta da matriz $\mathbf{K}_{n \times m}$ há um ganho em malha fechada. Assim, para um sistema $n \times m$ a matriz de ganho relativo ($\Lambda_{n \times m}$) é obtida dividindo cada elemento da matriz $\mathbf{K}_{n \times m}$ por $\mathbf{K}'_{n \times m}$, ou seja, o respectivo ganho em malha aberta pelo respectivo ganho em malha fechada, onde em um sistema genérico $n \times m$ o ganho relativo é fornecido pela Equação 2.19.

$$\lambda_{ij} = \frac{\left. \frac{\partial PV_i}{\partial MV_j} \right|_{MV_k, k \neq j}}{\left. \frac{\partial PV_i}{\partial MV_j} \right|_{PV_k, k \neq i}}. \quad (2.19)$$

As análises de interação entre as malhas de controle de um sistema usando informações da RGA são realizadas a partir da magnitude dos valores de λ_{ij} conforme apresentado abaixo:

- ❖ $\lambda_{ij}=0$ significa que a variável de entrada j não atua sobre a saída i . Neste caso, para sistemas 2×2 , a melhor escolha para o controle da variável de saída seria a outra variável de entrada (não a variável de entrada j);
- ❖ $\lambda_{ij}=1$ significa que o sistema é desacoplado, isto é, somente a entrada j atua sobre a saída i , o que possibilita que o sistema multivariável seja tratado como um sistema de múltiplas malhas individuais monovariáveis (SISO; *single input – single output*), resultando no chamado controle descentralizado;
- ❖ $0 < \lambda_{ij} < 1$ significa que há interações entre as malhas de controle;

- ❖ $\lambda_{ij} < 0$ significa a impossibilidade de controlar o sistema, pois ao fechar a outra malha de controle muda o ganho do par ij inviabilizando a operação do controlador da variável de saída i ;
- ❖ $\lambda_{ij} > 1$ significa que há interação e menos opções de pareamento (nenhuma no caso 2×2), pois outros λ terão valores negativos, para que a soma resulte 1.

Para sistemas $n \times m$, n e m superior a duas unidades, Bristol (1966) propôs encontrar os ganhos relativos com base em uma equação matricial. O 1º passo consiste em encontrar a matriz $\mathbf{B} = (\mathbf{K}_{n \times m}^{-1})^T$. Em seguida, multiplicar cada elemento da matriz resultante \mathbf{B} por cada termo da matriz original, $\mathbf{K}_{n \times m}$. Isso resulta na matriz de ganho relativo, Λ_{ij} , para sistemas $n \times m$, calculada conforme a Equação 2.20.

$$\Lambda_{ij} = (\mathbf{K}_{n \times m}^{-1})^T \cdot \mathbf{K}_{n \times m}. \quad (2.20)$$

2.6.2 Controle Inferencial de Composição

Normalmente, em processos de destilação a composição que se deseja controlar não pode ser medida em tempo real e, como, consequência, não é possível utilizar qualquer configuração de controle realimentação que necessite da medição direta dessa variável. No entanto, se os distúrbios (d) que geram o problema de controle possam ser medidos e se o modelo adequado do processo estiver disponível, é possível empregar controle antecipatório. Entretanto, se as medições dos distúrbios e da variável de processo não pode ser efetuada em tempo real, nenhuma das estratégias de controle tradicionais poderá ser utilizadas para manter a composição em seu valor desejado (*setpoint*). Este é o tipo de problema de controle onde o controle inferencial é a única solução (STEPHANOPOULOS, 1984; SEBORG et al., 2011).

Assim, o controle inferencial pode ser utilizado para o controle de composição em colunas de destilação. Nessa estratégia, a composição do produto é a variável normalmente inferida, porque faltam sensores rápidos, economicamente viáveis e confiáveis. A temperatura é a variável secundária mais comumente utilizada para inferir a composição não diretamente medida. No entanto, em processos de alta pureza as temperaturas apresentam pouca variabilidade, o que de fato pode comprometer a efetividade da inferência de composição. Nesse sentido, trabalhos como os de Brosilow e Joseph (1978) e Kano et al. (2000) apresentam boas práticas de controle inferencial, principalmente por utilizarem medições de temperaturas em vários estágios.

Os estudos sobre controle inferencial começaram em 1970, com a concepção de algoritmo, bem como a exploração desse tipo de controle, foi abordada por diversos autores, dentre os quais destacam-se pelo pioneirismo: Brosilow e Joseph (1978) e Doyle (1998).

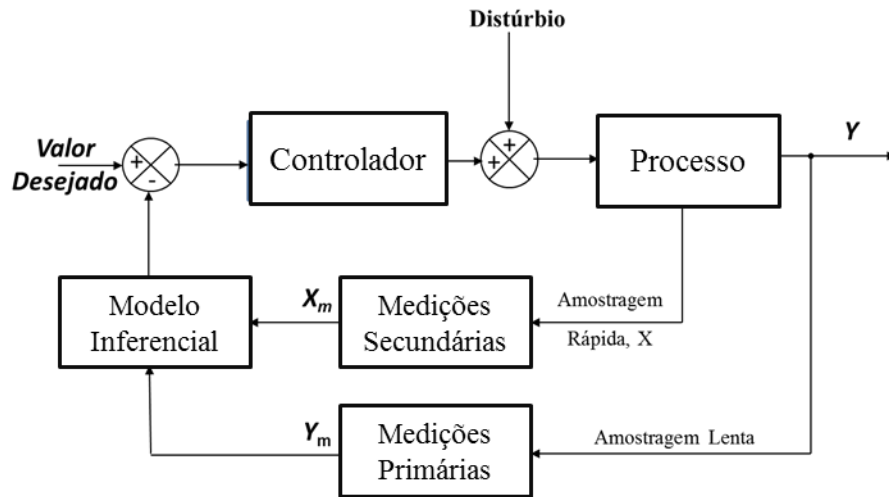
Brosilow e Joseph (2002) abordaram a questão de como selecionar medições secundárias para usar em um sistema de controle inferencial *feedback* (por exemplo, como localizar os estágios em uma coluna de destilação para instalar sensores de temperaturas para se obter um bom controle da composição do produto). Segundo os autores, os critérios utilizados para a seleção de medições secundárias incluem: boa capacidade preditiva da saída a partir das medições secundárias, custo de instalação e manutenção e capacidade do sistema de controle superar as falhas dos sensores.

A discussão nessa seção será limitada, visando à escolha das medições com base no primeiro critério, haja vista que o segundo e o terceiro critérios são viáveis quando são utilizados um pequeno número de medições. O último critério favorece estruturas de controle mais simples, que não dependem de muitas medições. À medida que a complexidade do sistema de controle aumenta, torna-se mais difícil lidar com falhas de sensores temporários.

Se forem utilizadas medições secundárias, então pode-se subdividir o primeiro critério em duas partes: (1) capacidade de estimar a variável primária corretamente (a exatidão ou o desempenho do sensor virtual) e (2) a capacidade de lidar com erros de modelagem, a sensibilidade ou robustez do sensor virtual (BROSILOW e JOSEPH, 2002).

Como a composição pode ser inferida a partir das medições secundárias (temperaturas, vazões, etc.) o algoritmo de controle inferencial realimentação pode ser aplicado. A Figura 2.12 apresenta a estrutura geral de controlador inferencial, onde o modelo inferencial (sensor virtual) substitui o elemento de medição da variável de processo. Nessa mesma figura, X_m são as medições das variáveis secundárias, que está disponível em uma base quase contínua de dados (amostragem rápida), enquanto que Y é a medida primária da variável de interesse, que é obtida de forma intermitente e com menor frequência (por exemplo, análise *off-line*/amostra de laboratório). Note-se que X e Y podem ser utilizadas para o controle, através de um modelo não linear usado pelo sensor virtual (*Soft Sensor*).

Figura 2.12. Diagrama de blocos de um *softsensor* utilizado para controle inferencial.



Fonte: Adaptado de SEBORG et al. (2011).

A estratégia de controle implementada na Figura 2.12 é denominada de controle preditivo inferencial baseada em modelo (MPIC), onde são utilizadas várias medições das variáveis secundárias, bem como estimadores não lineares e controladores. Brosilow e Joseph (1978), propuseram que a composição de destilado de uma coluna de destilação fosse inferida através da medição de três temperaturas do processo, previamente selecionadas, onde a vazão de refluxo era utilizada como variável manipulada. O modelo de inferência era um algoritmo linear de mínimo quadrado parcial (*Partial Least Squares*, PLS). No entanto, várias questões devem ser consideradas para uma generalização de SV em colunas de destilação, dentre as quais se destacam: seleção de medidas secundárias, estrutura dinâmica do sistema de controle e métodos de construção de estimadores (PIANG, 2005). Principalmente quando se trata de colunas de alta pureza, haja vista que as mesmas tipicamente têm perfis de temperatura praticamente invariáveis nas correntes de produto (topo e base), modificando-se bruscamente no meio da coluna, indicando que a separação primária dos componentes ocorre.

A baixa sensibilidade nesses processos limita a relação termodinâmica temperatura-concentração, exceto nos estágios onde ocorre a separação. Em tais casos, o problema pode ser reduzido pelo uso de várias outras medições do processo ou em casos de medianas variações, os sensores de temperatura, são colocados em toda a seção de sensibilidade da torre, onde pode ser calculada a média para fornecer um sinal de realimentação líquida que tem menor ganho e melhor linearidade (LUYBEN e CHIEN, 2010).

2.7. CONCLUSÕES DO CAPÍTULO

Neste capítulo foram abordados aspectos sobre destilação azeotrópica, alta pureza e do processo produtivo do 1,2-DCE, ênfase na etapa de purificação desse processo. Também foi tratado aqui o tema: sensores virtuais (SV) em processos de destilação. A abordagem trouxe características importantes sobre a metodologia de construção, dados de entrada e modelos de identificação aplicados na construção de sensores virtuais, tanto em processos genéricos quanto, especificamente, para colunas de destilação. Além disso, foi feita uma revisão sobre a controlabilidade multivariável e a possibilidade da aplicação de controle inferencial em torres de alta pureza. A Tabela 2.1 apresenta as referências bibliográficas de maior relevância, em ordem alfabética, e suas principais contribuições ao longo do desenvolvimento do trabalho.

Tabela 2.1. Principais referências utilizadas ao longo do desenvolvimento do trabalho.

TEMA	AUTOR	PRINCIPAIS CONTRIBUIÇÕES
Destilação azeotrópica e de alta pureza.	Luyben e Chien (2010) e Seader et al. (2011).	Fundamentação, considerações e modelagem termodinâmica.
Processo de purificação o 1,2-DCE.	Rodolfo Jr et al. (2006) e Silva et al. (2002).	Características e considerações.
Modelagem de estágios de equilíbrio e termodinâmica.	Boston e Sullivan (1974); Seader et al. (2011); Halder (2014) e Sandler (2006).	Equações MESH, cálculos γ - ϕ e solução numérica.
Modelagem nos <i>softwares Aspen Plus and Dynamics™</i> e controle de processos.	Luyben (2013); Schefflan (2011); Seborg et al. (2011) e Smith e Corrípio (2008).	Modelo estacionário e dinâmico, conversão de regimes, gráficos, exportação de dados, sinais e sintonia de controladores.
Sensores virtuais: metodologia de construção.	Fortuna et al. (2007); Kadlec et al. (2009); Kano et al. (2000) e Lotufo e Garcia (2008).	Fundamentos, construção e estudo de casos.
Seleção de variáveis e remoção <i>outliers</i> .	Fachin (2005); Fortuna et al. (2007); Jolliffe (1972) e 2002); e Montgomery e Runger (2014).	Métodos sequenciais e índices de desempenho.
Redes neurais artificiais.	Haykin (2001); Braga et al. (2000); Silva et al. (2010); e Taylor et al. (2006).	Conceitos, funções de ativação, arquiteturas, aprendizagem, treinamento e critérios de parada.

CAPÍTULO 3:

REDES NEURAIAS ARTIFICIAIS PARA CONSTRUÇÃO DE SV

3.1. INTRODUÇÃO

Uma rede neural artificial (RNA) consiste em um conjunto de neurônios artificiais interligados por conexões sinápticas. Ela está inserida dentro de uma área conhecida como sistemas inteligentes (conexionistas) ou inteligência computacional. A área de sistemas inteligentes é composta por várias outras ferramentas, que são: sistemas *fuzzy*, computação evolutiva, inteligência coletiva (*swarm intelligence*), sistemas imunológicos artificiais, agentes inteligentes e uma área que tem surgido em diversas aplicações, os sistemas híbridos inteligentes (JANG et al., 1997).

A principal característica da RNA é a sua adaptabilidade, resultado de sua arquitetura (paralela ou com capacidade de classificação) e dos algoritmos de aprendizagem utilizados, que são inspirados biologicamente e incorporam aspectos de organização e funcionalidade dos neurônios das células do cérebro humano (SILVA et al., 2010). A aplicabilidade das RNA tem crescido com sucesso nos mais diversos tipos de problemas, dentre os quais, bons resultados foram apresentados na construção de sensores virtuais de variáveis químicas e bioquímicas, como, por exemplo, no trabalho de Duwaish et al. (2002) para inferência da concentração do oxigênio. As diversas aplicações das RNA podem em parte ser atribuídas às suas propriedades, tais como: não linearidade, mapeamento de entrada e saída, adaptabilidade e generalização.

3.2. NEURÔNIOS ARTIFICIAIS

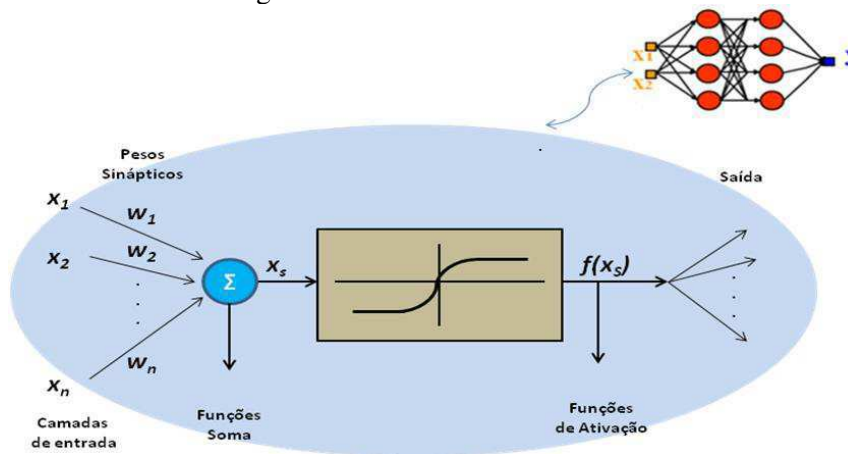
Um neurônio é uma unidade de processamento de informação que é fundamental para a operação de uma rede neural (HAYKIN, 2001). As RNA consistem em uma estrutura formada pela interligação entre esses elementos processadores que se assemelham ao comportamento, forma e até mesmo às funções dos neurônios biológicos dos seres humanos.

A inspiração e analogia da RNA, a rede neural biológica (sistema nervoso), é um sistema que consiste basicamente de três estágios: estímulo/recepção “receptores”, o centro do sistema nervoso “cérebro” e percepção/tomada de decisões apropriadas “atuadores”. A presença dos receptores indica a função de converter os estímulos do corpo humano ou do ambiente externo em impulsos elétricos, transmitindo informações para a rede neural (cérebro). Os atuadores tem a

função de converter os impulsos elétricos, gerados pela rede neural, em repostas discerníveis como saídas do sistema.

Trabalhos como o de McCulloch e Pitts (1943) foram pioneiros na modelagem matemática inspirada no neurônio biológico. O neurônio artificial é o elemento básico da RNA, ele pode ser chamado de nó ou nodo ou elemento processador, de tal forma que os padrões de conexões entre os neurônios artificiais determinarão como será feito o processamento das informações. As forças de conexões entre os neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido. A função de ativação processa os estímulos obtidos através das entradas. A Figura 3.1 apresenta a estrutura de um neurônio artificial.

Figura 3.1. Neurônio Artificial.



Fonte: Adaptada de SAITO (2010).

Sabendo que $f(x_s)$ é a função de ativação e que x_s é igual a combinação entre as entradas e os pesos sinápticos, tem-se a Equação 3.1:

$$x_s = \sum_{i=1}^n x_n w_n. \quad (3.1)$$

Sendo: $\{x_1, x_2, \dots, x_n\}$ as entradas, $\{w_1, w_2, \dots, w_m\}$ os pesos sinápticos e $\{y_1, y_2, \dots, y_n\}$ as saídas. O principal objetivo da função de ativação é limitar a saída do neurônio dentro de um intervalo de valores razoáveis a serem assumidos pela sua própria imagem funcional. A função de combinação linear, Equação 3.1, tem como objetivo agregar todos os sinais de entrada que foram ponderados pelos respectivos pesos sinápticos, a fim de produzir um valor de potencial de ativação. As saídas consistem no valor final produzido pelo neurônio em relação a um determinado conjunto de sinais de entrada, podendo ser também utilizadas por outros neurônios que estão sequencialmente interligados.

3.3. TIPOS DE FUNÇÕES DE ATIVAÇÃO

As funções de ativação podem ser divididas em dois grandes grupos principais: funções parcialmente diferenciáveis e funções totalmente diferenciáveis, considerando-se para tanto todo o domínio de definição das mesmas (SILVA et al., 2010). A seguir serão apresentadas os tipos de funções de ativação com maior empregabilidade.

i) Funções parcialmente diferenciáveis: são aquelas que possuem pontos cujas derivadas de primeira ordem são inexistentes. As duas principais funções do grupo são: a função degrau e a função rampa simétrica.

a) Função degrau (Equação 3.2). Produz um resultado que assume valores unitários positivos quando o potencial de ativação (entradas) do neurônio for maior ou igual a zero, caso contrário assume valores nulos.

$$f(x_s) = \begin{cases} 1, & x_s \geq 0 \\ 0, & x_s < 0 \end{cases} \quad (3.2)$$

b) Função rampa simétrica (Equação 3.3). Propõe que os valores retornados sejam iguais aos próprios valores dos potenciais de ativação quando estes estão definidos dentro do intervalo $[-a, a]$, restringindo-se aos valores limites em caso contrário.

$$f(x_s) = \begin{cases} a, & x_s > a \\ x_s, & -a \leq x_s \leq a \\ -a, & x_s < -a \end{cases} \quad (3.3)$$

ii) Funções totalmente diferenciáveis: são aquelas cujas derivadas de primeira ordem existem e são conhecidas em todos os pontos de seu domínio de definição. As principais funções de ativação pertencentes a esse grupo são mostradas a seguir:

a) Função linear (Equação 3.4). Produz resultados de saída semelhantes aos valores de sua entrada x_s .

$$f(x_s) = ax_s. \quad (3.4)$$

Sendo a um número real que define a saída linear para os valores de x_s .

b) Função logística (Equação 3.5). A saída produzida assume valores reais entre zero e um.

$$f(x_s) = \frac{1}{1 + e^{-\alpha 2x_s}}. \quad (3.5)$$

Sendo α uma constante real associada ao nível de inclinação da função logística frente ao seu ponto de inflexão.

- c) Função Tangente Hiperbólica (Equação 3.6). A saída produzida por essa função assume valores reais entre -1 e 1.

$$f(x_s) = \frac{1 - e^{-\alpha x_s}}{1 + e^{-\alpha x_s}}. \quad (3.6)$$

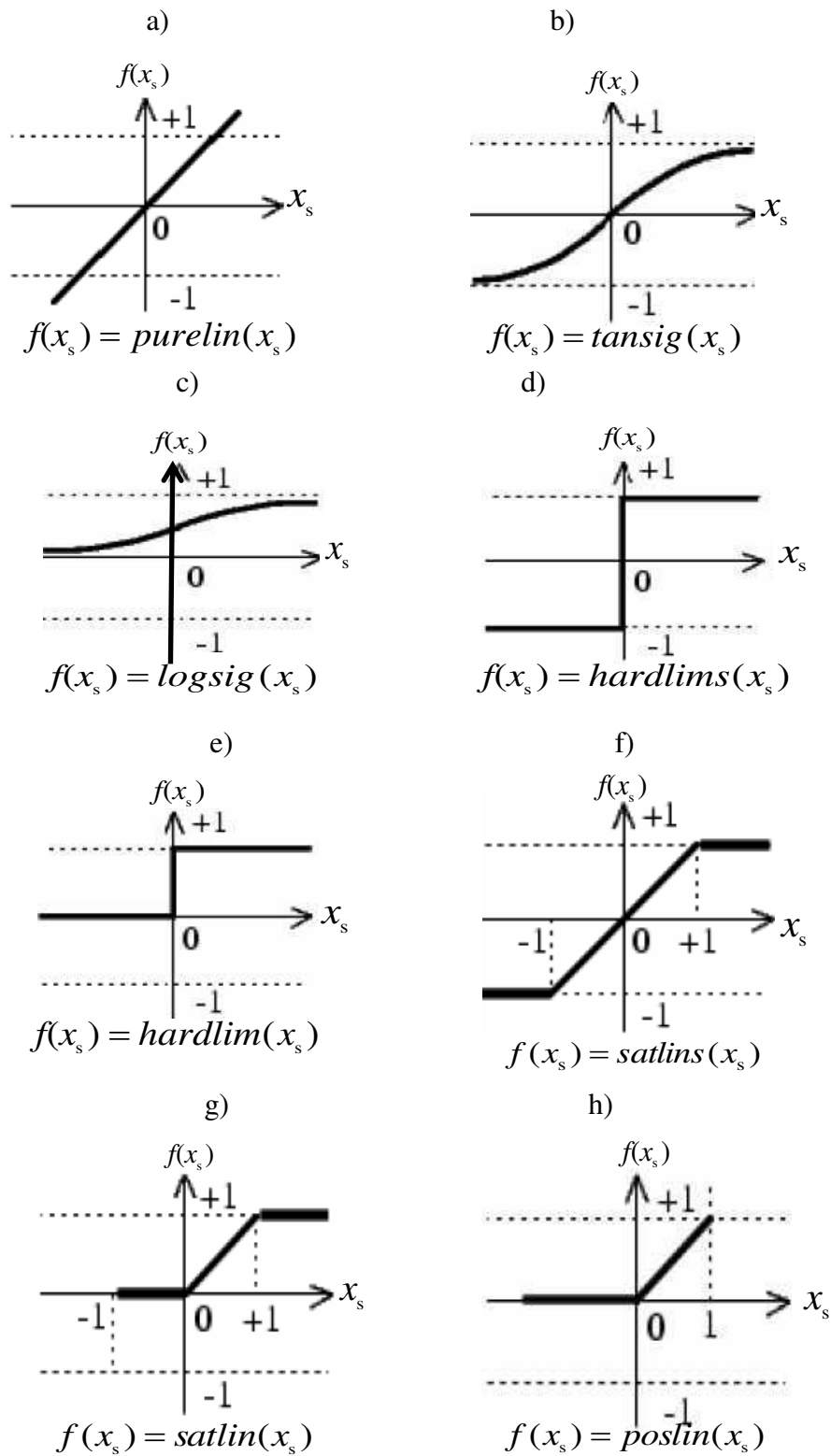
A principal propriedade da função de ativação é a não linearidade, o que permite um mapeamento não linear. O uso ideal de uma função de ativação deve satisfazer as exigências da melhor aproximação do problema proposto e da minimização do esforço computacional (HAYKIN, 2001). Os gráficos das principais funções de ativação, disponíveis no *software* Matlab®, são apresentados na Figuras 3.2.

3.4. ARQUITETURAS DE UMA RNA

A arquitetura de uma RNA é o modo pelo qual os neurônios estão estruturados/dispostos ou arranjados (CAMPOS e SAITO, 2004). De acordo com as principais arquiteturas das redes neurais artificiais, uma RNA pode ser dividida em três camadas, que são nomeadas da seguinte forma:

- a) Camada de entrada: responsável por receber os dados (informações), sinais, características ou medições. Essas entradas são geralmente, normalizadas com relação às faixas de variações dinâmicas produzidas por alguma das funções de ativações anteriormente definidas;
- b) Camada oculta (escondida): responsável por extrair as características que estão associadas ao processo ou sistema, como, por exemplo, a inferência de composição em um processo de destilação;
- c) Camada de saída: responsável pela produção e apresentação dos resultados finais da rede, os quais são oriundos dos processamentos efetuados pelos neurônios das camadas anteriores.

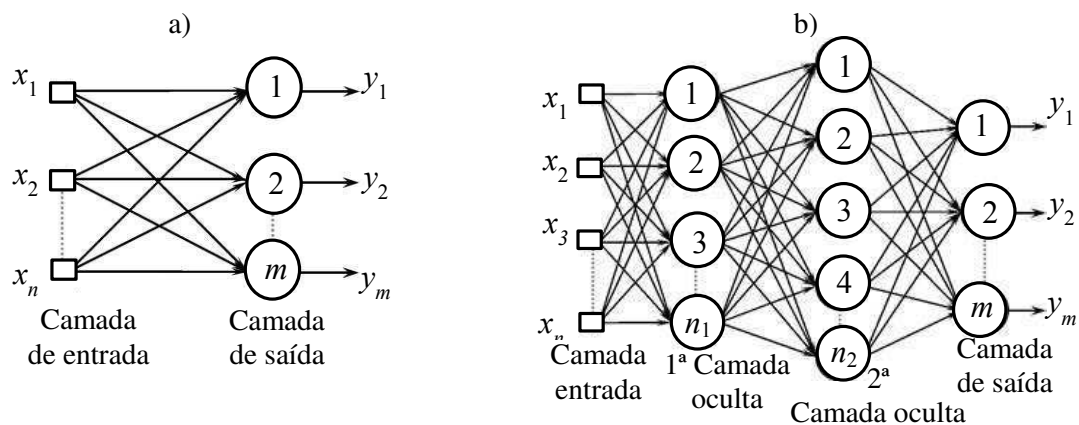
Figura 3.2. Funções de ativação do tipo: a) Linear; b) Tangente hiperbólica; c) Logarítmica-sigmóide; d) Linear simétrica; e) Linear assimétrica; f) Linear com saturação simétrica; g) Linear com saturação assimétrica; e h) Linear positiva.



Fonte: Adaptadas do *Neural Network Toolbox* do Matlab® R2010b e R2015a.

A forma estrutural dos neurônios apresentam três tipos de arquiteturas distintas, que são: redes alimentadas adiante (*feedforward*) de uma única e com múltiplas camadas, e redes recorrentes de *Elman*. As redes alimentadas adiante possuem apenas uma camada de entrada, contendo os nós fonte e uma camada de saída contendo os nós computacionais. A camada de entrada não é considerada na contagem do número de camadas, pois são neurônios são considerados especiais e seu papel é distribuir as entradas da rede (sem modificá-las) aos neurônios da camada seguinte. As redes *feedforward* multicamadas se distinguem das anteriores pelo fato de possuírem uma ou mais camadas de neurônios ocultos. São utilizadas em diversas aplicações, dentre as quais se destacam: aproximações de funções, classificação de padrões, identificação, otimização e controle de processos. As Figuras 3.3 a) e b) apresentam, respectivamente, a arquitetura de RNA redes alimentadas adiante com uma e múltiplas camadas.

Figura 3.3 a) Rede adiante com uma única camada; b) Rede adiante multicamadas.



Fonte: Adaptadas de SILVA et al. (2010).

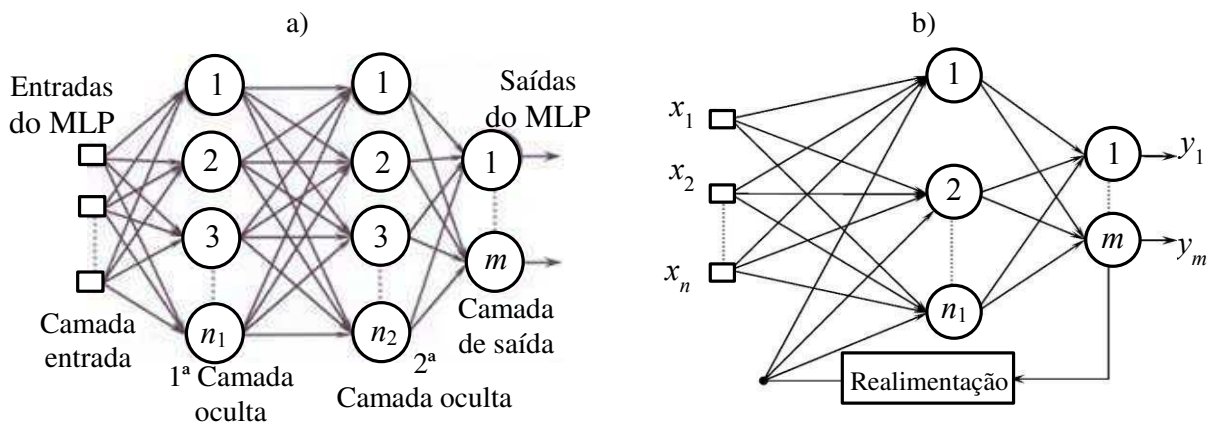
De acordo com as Figuras 3.3 a) e b), a quantidade de sinais de saídas coincide com o número de neurônios da camada de saída. A Figura 3.3 b) apresenta um exemplo de uma RNA *feedforward* multicamadas, com uma camada de entrada, duas camadas neurais ocultas (n_1 e n_2) uma camada neural de saída, composta por m neurônios que representam os valores das saídas aplicadas. Na estrutura alimentada a quantidade de neurônios que fazem parte da primeira camada escondida é normalmente diferente do número de sinais que compõem a camada de entrada da rede (HAYKIN, 2001). Sendo assim, o número de camadas escondidas e seus respectivos neurônios constituintes, dependem do tipo e da complexidade do problema a ser mapeado pela rede, quantidade e qualidade dos dados disponíveis sobre o problema. Dentre os principais tipos de redes com arquitetura de rede alimentada adiante se encontra o *Perceptron* multicamadas (MLP, do inglês *Multilayer Perceptron*) e as redes de base radial (RBF, do inglês *Radial Basis Function*), cujos algoritmos de aprendizado utilizados em seus processos de

treinamento são baseados, respectivamente, na regra delta generalizada e na regra delta competitiva.

As RNA do tipo MLP são as mais difundidas na literatura, principalmente para o objetivo de inferência de variáveis, sua arquitetura pode ser vista na Figura 3.4 a). Durante o processo de aprendizagem da MLP é de costume dividir os dados disponíveis em dois conjuntos: um conjunto de treinamento, com um total de 60 a 90% dos dados e um conjunto de teste, que contém um total de 10 a 40% dos dados. O conjunto de treinamento permite observar a aprendizagem da rede e o conjunto de testes permite avaliar a capacidade de generalização da mesma (HIMMELBLAU, 2008).

As redes recorrentes ou realimentadas apresentam “ciclos” nas suas conexões, ou seja, a saída do neurônio de uma camada i constitui a entrada de neurônios de uma camada $i - j$, com $j \geq 0$. Em resumo, são RNA cujas saídas dos neurônios são realimentadas como sinais de entrada para outros neurônios. Essas características qualificam esse tipo de rede para processamento dinâmico de informações, ou seja, elas podem ser utilizadas tanto em sistemas variantes no tempo, como previsões de séries temporais, identificação, otimização e controle de processos (HAYKIN, 2001; FORTUNA et al, 2005). Dentre as principais arquiteturas de RNA que possuem realimentação, destacam-se: rede recorrente *Perceptron* multicamadas (RMLP) com realimentação entre neurônios situados em camadas distintas (utilizam algoritmos de treinamento baseados na minimização e na regra delta generalizada); rede de Hopfield que emprega um princípio chamado de armazenamento de informação como atratores dinamicamente estáveis (HAYKIN, 2001; SILVA et al., 2010). O esquema de uma arquitetura de RNA do tipo RMLP é apresentado conforme a Figura 3.4 b).

Figura 3.4 a) Rede *Perceptron* multicamadas; b) Rede recorrente *Perceptron* multicamadas.



Fonte: Adaptadas de SILVA et al. (2010).

É importante ressaltar que a rede recorrente, Figura 3.4 b), recupera as informações através do processo dinâmico de atualização dos estados dos neurônios, sendo o neurônio a ser atualizado escolhido aleatoriamente. O atraso na conexão e a alimentação dos valores internos em um passo seguinte fazem com que a rede recorrente seja capaz de armazenar informações do passo anterior e a utilize em um momento futuro. Com isso, ela desenvolve habilidades e é capaz de inferir e reconhecer padrões temporais e espaciais (HIMMELBLAU, 2008).

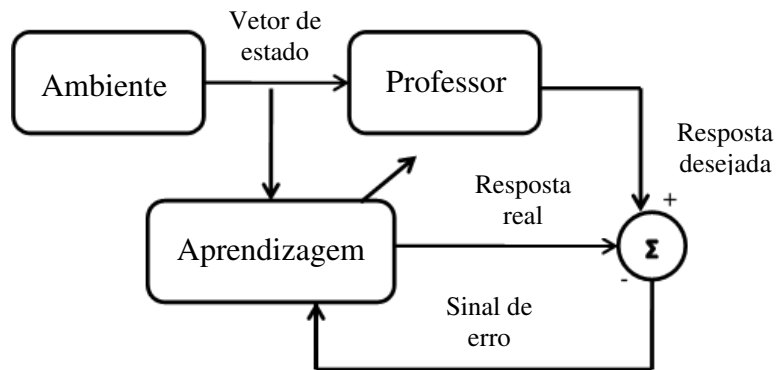
3.5. PROCESSO DE APRENDIZAGEM DE UMA RNA

A propriedade de importância primordial para uma rede neural é a sua habilidade de aprender a partir do ambiente e com isso, conseguir melhorar seu desempenho através do processo de aprendizagem. Essa etapa é considerada a maior capacidade atribuída a uma RNA, haja vista que ao aprender a partir da apresentação de amostras (relacionamento entre as entradas e as saídas) ela é capaz de generalizar soluções e produzir uma saída próxima à desejada (HAYKIN, 2001).

O processo de treinamento de uma RNA utiliza passos ordenados para ajuste dos pesos sinápticos e limiares de seus neurônios, tendo-se como objetivo final a generalização de soluções a serem produzidas pelas suas saídas, onde as respostas são representativas dos sistemas físicos que estão sendo mapeados. O conjunto de passos necessários para o treinamento da RNA é denominado de “processo de aprendizagem” (BRAGA et al., 2000). Normalmente, o conjunto total das amostras é dividido em dois subconjuntos: treinamento e teste. O primeiro subconjunto é utilizado essencialmente no processo de aprendizado da RNA, enquanto o segundo é usado para verificar se os aspectos referentes à generalização de soluções são aceitáveis, o que possibilita a validação da topologia assumida (SILVA et al., 2010).

Os métodos desenvolvidos para treinamento de RNA podem ser agrupados em dois paradigmas principais: aprendizado supervisionado e aprendizado não supervisionado. O aprendizado supervisionado, como o próprio nome sugere, é realizado sob a supervisão de um (professor) externo que tem a função de monitorar a resposta da rede, para cada entrada, conhecendo previamente a saída esperada. O aprendizado supervisionado tem atraído a atenção de muitos pesquisadores, uma das motivações talvez seja o fato que ele pode ser visto como um problema de otimização e também, por usar ferramentas que já mostraram sua eficácia, tanto em programação linear quanto em não linear. A Figura 3.5 apresenta o diagrama de blocos de uma RNA treinada por aprendizado supervisionado.

Figura 3.5. Diagrama de blocos de aprendizagem supervisionada.



Fonte: Adaptada de HAYKIN (2001).

O aprendizado não supervisionado é baseado na auto-organização, dispensando a utilização, tanto do aprendizado “crítico” quanto do “professor”. Não existe o conhecimento das saídas desejadas para as entradas, ou seja, a própria RNA deve se auto organizar em relação às particularidades existentes entre os elementos que compõem o conjunto total de amostras e ser capaz de identificar os *clusters* (subconjuntos) que contenham maior similaridades.

Um modelo que tem boa generalização é o que responde corretamente aos exemplos contidos na base de aprendizado, mas também a outros exemplos, diferentes dos usados durante o processo de aprendizagem, estando contidos em uma base de teste. Enfim, a habilidade de generalizar é a principal capacidade na busca das tarefas que envolvem o aprendizado. Uma RNA, porém, pode se especializar demasiadamente em relação aos exemplos contidos na base de aprendizado. Este tipo de comportamento gera um problema, conhecido como sobre aprendizado ou sobre ajuste (*over-training/over-fitting*). Outra situação que pode ocorrer é a RNA não conseguir generalizar, gerando o problema chamado de sub-ajuste (*under-fitting*).

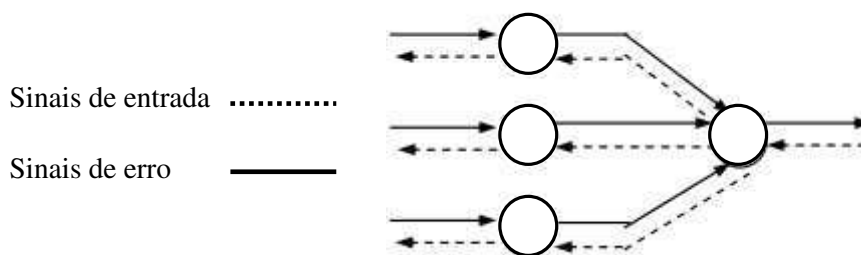
3.6. ALGORITMOS DE TREINAMENTO DE UMA RNA

Pode-se dizer que o treinamento é o ponto onde se define o sucesso ou o fracasso da RNA, haja vista que nesse procedimento, submete-se a RNA ao aprendizado, onde alguns fatores relevantes devem ser levados em conta, dentre eles: o algoritmo de treinamento e o número de épocas (ciclos ou iterações). Essencialmente o aprender das RNA consiste em colocar valores de conexões sinápticas. Em alguns casos, estes valores são colocados representando certo conhecimento, como no caso usado em sistemas especialistas; em outros, é utilizado um algoritmo para encontrá-los, chamado algoritmo de aprendizagem.

Neste trabalho são propostas RNA do tipo MLP para construção dos sensores virtuais. Estão disponíveis na literatura diversos algoritmos para treinamento de uma RNA, onde os mesmos podem ser classificados em duas categorias distintas: métodos de treinamento determinísticos e os de treinamentos estocásticos.

Este trabalho foca apenas nos algoritmos de treinamento determinísticos, sendo assim vale a pena realizar uma abordagem sobre o algoritmo de retropropagação do erro. Ele é um algoritmo utilizado em *perceptrons* de múltiplas camadas e utiliza funções de transferência diferenciáveis e não lineares em um treinamento supervisionado. O método consiste basicamente, de dois passos através das diferentes camadas da RNA: um passo para frente (propagação) e um passo para trás (retropropagação). Na propagação, submete-se a RNA a uma entrada e seu efeito se propaga camada por camada através da mesma. Então, um conjunto de saídas é produzido como a resposta real da RNA, de modo que nesta etapa os pesos sinápticos sejam todos fixos. Na retropropagação os pesos sinápticos são ajustados de acordo com uma regra de correção de erro, a saída é comparada com o valor desejado e calcula-se o gradiente do erro. Este sinal de erro é então propagado para trás da RNA, contra a direção das conexões sinápticas, vindo daí o nome “retropropagação do erro” (HIMMELBLAU, 2008). A Figura 3.6 representa os sinais de entrada, também conhecidos como sinais funcionais e de erro em um *perceptron* de múltiplas camadas.

Figura 3.6. Sinais funcionais e de erro numa rede neural.



Fonte: Adaptada de HAYKIN (2001).

Como na retropropagação a correção a cada passo é proporcional à derivada, os pesos mudam mais para as unidades que têm uma ativação de valor intermediário, isto contribui para a convergência do algoritmo. Convém notar que a retropropagação pode ser usada (teoricamente) para neurônios onde as entradas são combinadas usando multiplicação, para RNA com realimentação ou recorrentes.

Neste trabalho será utilizada a RNA do tipo MLP, onde o objetivo é reduzir o erro de aprendizagem por correção de erro. Sabe-se que o treinamento das MLP é um problema de otimização não linear de uma função custo, que mede o erro quadrático médio calculado pela saída da RNA frente a uma saída desejada. A Tabela 3.1 apresenta alguns dos principais algoritmos de treinamento utilizados para treinamento dos SV, que se distinguem pelo emprego da estratégia de otimização utilizada, de forma a realizar e ajustar os pesos e a *bias* da RNA.

Tabela 3.1. Principais algoritmos de treinamentos, seus respectivos métodos de otimização e função de treinamento no software Matlab®.

MÉTODO	ALGORITMO DE TREINAMENTO	SIGLA	FUNÇÃO NO MATLAB®
Gradiente Descendente (1ª ordem)	Gradiente descendente básico	GD	<i>traingd</i>
	Gradiente descendente com taxa de aprendizado adaptativa	GDA	<i>traingda</i>
	Gradiente descendente com momento	GDM	<i>traingdm</i>
	Gradiente descendente com momento e com taxa de aprendizado adaptativa	GDX	<i>traingdx</i>
	Resilient backpropagation	RP	<i>trainrp</i>
Gradiente Conjugado (2ª ordem)	Gradiente conjugado escalonado	SCG	<i>trainscg</i>
	Gradiente conjugado de Powell-Beale	CGB	<i>traingb</i>
	Gradiente conjugado de Polak-Ribière	CGP	<i>traincgp</i>
	Gradiente conjugado de Fletcher-Reeves	CGF	<i>traincgf</i>
Quase-Newton (2ª ordem)	<i>Broyden-Fletcher-Goldfarb-Shanno</i> (BFGS)	BFG	<i>trainbfg</i>
	<i>Levenberg-Marquardt</i>	LM	<i>trainlm</i>
	Regularização de <i>Bayesian</i>	BR	<i>trainbr</i>
Gradiente Conjugado e Quase-Newton	Método da secante de único passo (2ª ordem)	OSS	<i>trainoss</i>

Fonte: WERLE (2012).

Nesse trabalho, foram utilizados quatro tipos de algoritmo de treinamento: Levenberg-Marquardt (LVM), o *Resilient Backpropagation* (RP), Regularização Bayesiana (BR) e o *One Step Secant Method* (OSS). Esses algoritmos de treinamento produziram bons resultados à estimativa de composição nos trabalhos de Plumb et al. (2005) e Werle (2012).

De fato, nenhum algoritmo pode ser considerado melhor que os demais, o melhor será aquele que se adapta à situação específica (ZANATA, 2005). Alguns desses métodos requerem poucas modificações no algoritmo de retropropagação do erro. Com isso, eles conseguem acelerar a convergência média na fase de treinamento e não exigem a escolha dos parâmetros críticos da RNA, tais como a taxa de aprendizado e o coeficiente de momento.

3.7. CRITÉRIO DE PARADA

Uma das dificuldades ao se utilizar as RNA é estabelecer o ponto de parada do algoritmo de treinamento, pois o erro de treinamento inicia com um alto valor que decresce rapidamente e continua diminuindo lentamente, tendendo a atingir um mínimo local na superfície de erro. Logo, para encontrar um valor mínimo global (ponto ideal de parada), deve-se obter a melhor generalização da rede (BRAGA et al., 2000; HAYKIN, 2001).

Existem vários métodos para determinação do momento em que o treinamento de uma RNA deve ser encerrado. A determinação destes critérios é fundamental para um bom treinamento e conseqüentemente, uma boa generalização das RNA utilizadas no desenvolvimento dos sensores virtuais. Os critérios de parada mais utilizados são (TAYLOR et al., 2006; HAYKIN, 2001):

- ❖ Número de ciclos: define o número de ciclos de treinamento, ou seja, o número de vezes em que o conjunto de treinamento é apresentado à rede. Se houver números excessivos de ciclos, a RNA pode perder o poder de generalização (*overfitting*). Entretanto, com um pequeno número de ciclos a rede pode não alcançar seu melhor desempenho (*underfitting*). Sugerem-se, valores entre 500 e 3000 épocas de treinamento;
- ❖ Erro: consiste em encerrar o treinamento após o erro médio quadrático ficar abaixo de uma tolerância especificada. Um erro médio quadrático muito pequeno não implica necessariamente em uma boa generalização, haja vista que este valor depende do problema em questão. Sugere-se estabelecer um valor de 0,01 no primeiro treinamento e depois ajustá-lo em função do resultado;
- ❖ Combinação dos métodos anteriores: pode-se estipular como critério de parada uma combinação dos métodos anteriormente mencionados. Neste caso, o treinamento é encerrado quando qualquer um dos critérios acima seja satisfeito;
- ❖ Validação (*Best Model*): no critério por validação, o treinamento é interrompido a cada C ciclos e é realizada uma estimativa do erro da rede com base no conjunto de dados de teste. A partir do momento em que o erro medido no conjunto de teste apresentar crescimento, o treinamento é encerrado. O que se deseja com esta técnica é descobrir o momento exato em que a rede começa a perder generalização.

Nesse trabalho adotou-se a técnica de parada antecipada, com base na validação cruzada. Essa é uma técnica estatística que valida o modelo obtido durante o treinamento da rede, que

utiliza um conjunto de dados diferentes dos usados para estimar os parâmetros do treinamento. O método consiste em acompanhar a evolução do aprendizado nas curvas correspondentes aos subconjuntos de dados de treinamento e de validação. Deste modo, o treinamento é interrompido, quando a curva de validação decresce a um erro mínimo e antes de começar a crescer, conforme o treinamento continua (CAMPOS e SAITO, 2004).

3.8. CONCLUSÕES DO CAPÍTULO

Nesse capítulo foram apresentados os principais aspectos das redes neurais artificiais a serem utilizadas e avaliadas para construção dos sensores virtuais. Esses aspectos vão desde a arquitetura das RNA, funções de ativação, processo de aprendizagem, algoritmo de treinamento e definição dos critérios de parada durante o treinamento e simulação das RNA. A definição dessas características é de suma importância para o desenvolvimento do trabalho, haja vista que sem esse conhecimento acarretaria em diversos problemas, tanto no aspecto do conhecimento prático da identificação do processo, quanto na possibilidade de melhorias a serem alcançadas, acurácia e confiabilidade dos sensores virtuais desenvolvidos.

CAPÍTULO 4:

SELEÇÃO DE VARIÁVEIS E PRÉ-PROCESSAMENTO DE DADOS

4.1. INTRODUÇÃO

Um dos grandes desafios à construção de modelos não lineares de identificação, como é o caso das RNA, é a correlação entre as variáveis de entrada e saída. Na prática, a retirada de algumas variáveis formando modelos menores é uma abordagem popular, entretanto, isso diminui o desempenho de previsão do modelo. Por outro lado, incluir todas as variáveis aumentará sua complexidade, como consequência sobre ajuste pode ocorrer (PIANG, 2005).

De fato, a presença de variáveis adicionais representa uma fonte de perturbação ou ruído, conduzindo a um sistema com deficiência em sua construção (ZAMPROGNA et al., 2001). As características requeridas para os dados a serem utilizados na geração da inferência confundem-se com a natureza e as características do algoritmo e também, das variáveis integrantes do mesmo, haja vista que a coleta dos dados é uma das etapas mais importantes da obtenção das variáveis, pois se os dados forem ruins, o sensor virtual também será ruim.

Ao se escolher os dados coletados utilizados para inferência deve-se preocupar com as seguintes características: origem, periodicidade, precisão, abrangência, coerência e tendência desses dados (CAMPOS et al., 2013). A abrangência dos dados é uma das principais características, haja vista que eles devem ser capazes de representar o processo a ele associado em todo o espaço operacional em que a unidade de processo normalmente trabalha, garantindo assim que ela seja representativa para toda essa faixa operacional.

Nesse capítulo é abordada a metodologia para a seleção de variáveis e de pré-processamento dos dados a serem utilizados. Os tópicos seguintes apresentam uma discussão desses métodos e propostas de implementação de seus algoritmos computacionais.

4.2 TÉCNICAS DE SELEÇÃO DE VARIÁVEIS

Devido à natureza dos processos químicos e sistemas de engenharia de processo, os estados de muitas das variáveis de entrada refletem nos estados de variáveis complexas (JERBIĆ et al., 2007). Assim, as observações em regime estacionário e dinâmico trazem informações importantes e possibilitam a coleta dos dados do sensor virtual. No entanto, a seleção das

variáveis de entrada a serem efetivamente utilizadas, em geral, não pode ser feita baseando-se apenas em aspectos teóricos, haja vista que a influência de certas variáveis pode ser apenas parcialmente independente de outras ou pode variar com a região operacional do processo (KANO et al., 2009).

Para se obter o grupo de variáveis independentes ideal na construção dos SV, existem algumas técnicas de seleção de variáveis, dentre as quais destacam-se: os métodos de busca exaustiva (todas as possibilidades), os métodos evolutivos e outras técnicas estatísticas. Nos métodos de busca exaustiva é possível identificar o melhor subconjunto de variáveis, quando existe um grande número de variáveis candidatas. Os métodos evolutivos consistem em propor modelos divididos em etapas, onde a cada etapa introduz-se ou remove-se uma nova variável. As etapas terminam quando não ocorre melhoria significativa no modelo ou quando todas as variáveis já foram inclusas/removidas (CHONG et al., 2005; MONTGOMERY et al., 2001).

Os métodos evolutivos ou procedimentos sequenciais “*stepwise*” mais utilizados para seleção de variáveis são: todas as regressões possíveis (do Inglês *All Possible Regressions*, TRP), adição sequencial (do Inglês *Forward Selection*, FS), seleção por eliminação (do Inglês *Backward Elimination*, BE) e a regressão por etapas (do Inglês *Stepwise Regression*, SR). É importante citar que em todas as abordagens anteriores aplicam-se em princípio, o comportamento de dependência linear. Eles são baseados em estatística multivariada, onde o modelo de referência pode ser, por exemplo, um modelo de regressão linear múltipla (MLR) (CAMPOS et al., 2013; MING-DA MA et al., 2009).

Existem ainda, outras técnicas que utilizam a análise de sensibilidade para reduzir a dimensão do espaço de entradas através de técnicas de projeções. Dentre essas abordagens a análise de componentes principais (PCA), talvez seja a mais utilizada. A ideia central da PCA é reduzir a dimensionalidade de um conjunto de dados nos quais existe um número grande de variáveis inter-relacionadas, mantendo o máximo possível da variação presente desse conjunto de dados (FORTUNA et al., 2007; JOLLIFE, 2002).

Nesse trabalho, foi construído um algoritmo de seleção de variáveis de modo que qualquer um dos procedimentos sequenciais podem ser implementados. Entretanto, o algoritmo escolhido foi o da TRP, pois esse procedimento é o mais adequado quando o número de variáveis secundárias (regressores) não é muito grande e por não proporcionar distorções pelas dependências entre os regressores. Além da técnica de seleção de variáveis TRP, também foi

desenvolvido um algoritmo de seleção que utiliza a PCA, visando projetar o conjunto de dados originais em um número reduzido de características e assim, conseguir-se extrair as informações relativas à variabilidade e dependência dos dados.

4.2.1. Modelo de Regressão Linear Múltipla (MLR)

Um modelo linear múltiplo com constante diferente de zero, com N variáveis independentes e i pontos experimentais ($i = 1, 2, \dots, n; i > N$), é representado pela Equação 4.1:

$$\hat{y}_i = y_i - \varepsilon_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_N x_{iN}. \quad (4.1)$$

Sendo: ε_i o resíduo ou erro do ponto experimental i .

Os dados experimentais y_i são representados pela Equação 4.2.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_N x_{iN} + \varepsilon_i. \quad (4.2)$$

A Equação 4.2 é reescrita para cada ponto experimental, dando origem a Equação (4.3):

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_N x_{1N} + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_N x_{2N} + \varepsilon_2 \\ y_3 &= \beta_0 + \beta_1 x_{31} + \beta_2 x_{32} + \dots + \beta_N x_{3N} + \varepsilon_3 \quad \cdot \\ &\vdots \\ y_m &= \beta_0 + \beta_1 x_{m1} + \beta_2 x_{m2} + \dots + \beta_N x_{mN} + \varepsilon_m \end{aligned} \quad (4.3)$$

Agora, a Equação 4.3 é reescrita em notação matricial, tem-se a Equação 4.4 ao:

$$Y = X\beta + \varepsilon$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1N} \\ 1 & x_{21} & x_{22} & \cdots & x_{2N} \\ 1 & x_{31} & x_{32} & \cdots & x_{3N} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{m1} & x_{m2} & \cdots & x_{mN} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_m \end{bmatrix}. \quad (4.4)$$

A Equação 4.4 pode ser rearranjada, onde as equações resultantes a serem resolvidas são equações normais (Equações 4.5 e 4.6). A estimativa de mínimos quadrados dos coeficientes β é feita utilizando a Equação 4.6.

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{y}). \quad (4.5)$$

$$\mathbf{X}^T \mathbf{X} \hat{\beta} = \mathbf{X}^T \mathbf{y}. \quad (4.6)$$

Existem $p = N+1$ equações normais para $p = N+1$ incógnitas (os valores de $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_N$). A matriz $\mathbf{X}^T \mathbf{X}$ é sempre não singular, de modo que sua inversão é descrita facilmente.

É fácil verificar que a forma matricial das equações normais é idêntica à forma escalar, conforme descreve a Equação 4.7.

$$\begin{bmatrix} m & \sum_{i=1}^m x_{i1} & \sum_{i=1}^m x_{i2} & \cdots & \sum_{i=1}^m x_{iN} \\ \sum_{i=1}^m x_{i1} & \sum_{i=1}^m x_{i1}^2 & \sum_{i=1}^m x_{i1}x_{i2} & \cdots & \sum_{i=1}^m x_{i1}x_{iN} \\ \sum_{i=1}^m x_{i2} & \sum_{i=1}^m x_{i2}x_{i1} & \sum_{i=1}^m x_{i2}^2 & \cdots & \sum_{i=1}^m x_{i2}x_{iN} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sum_{i=1}^m x_{iN} & \sum_{i=1}^m x_{iN}x_{i1} & \sum_{i=1}^m x_{iN}x_{i2} & \cdots & \sum_{i=1}^m x_{iN}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_{i1}y_i \\ \sum_{i=1}^m x_{i2}y_i \\ \vdots \\ \sum_{i=1}^m x_{iN}y_i \end{bmatrix}. \quad (4.7)$$

Se a multiplicação matricial indicada pela Equação 4.7 for feita, resultará na forma escalar das equações normais, Equações 4.5 e 4.6. Dessa maneira, é fácil verificar que $\mathbf{X}^T \mathbf{X}$ é uma matriz simétrica ($p \times p$) e $\mathbf{X}^T \mathbf{y}$ é um vetor coluna ($p \times 1$), onde os elementos da diagonal de $\mathbf{X}^T \mathbf{X}$ são somas dos quadrados dos elementos nas colunas de \mathbf{X} e os elementos fora da diagonal são as somas dos produtos cruzados dos elementos nas colunas de \mathbf{X} .

A seleção de variáveis para compor o modelo do sensor virtual através dos procedimentos *stepwise* é realizada através da comparação entre os diferentes modelos gerados mediante as combinações das variáveis disponíveis. Essa comparação deve ser realizada de forma sistemática e, para isso, podem ser utilizados diversos índices de desempenho capazes de quantificar o ajuste do modelo aos dados experimentais.

4.2.2. Índices de Desempenho

A determinação do modelo que mais se ajusta aos dados é feita através da utilização de índices de ajuste. Existem diversos de índices, a maioria utiliza a soma quadrática do erro (SQE, Equação 4.8) como medida do desvio dos dados originais (y_i) e os fornecidos pelo modelo (\hat{y}_i).

$$SQE = \sum_{i=1}^N (y_i - \hat{y}_i)^2. \quad (4.8)$$

O primeiro índice que é função da SQ_E é a média quadrática do erro (MQE), para p variáveis contidas no modelo e n número de amostras testadas (Equação 4.9).

$$MQE = \frac{SQE}{(n - p)}. \quad (4.9)$$

O índice mais difundido para medir a capacidade de ajuste de um modelo é o coeficiente de correlação, normalmente referenciado como R^2 . Ele estabelece uma relação entre o erro originado pelo modelo e a soma quadrática total corrigida (SQ_T) (Equação 4.10).

$$SQ_T = \sum_{i=1}^N (y_j - \bar{y}_j)^2. \quad (4.10)$$

Complementando em relação a 1 (um) a relação entre SQ_E e SQ_T obtém-se a expressão final para o cálculo do R^2 (Equação 4.11):

$$R^2 = 1 - \frac{SQE}{SQ_T}. \quad (4.11)$$

O índice R^2 não é confiável na determinação do melhor conjunto de variáveis, pois ele apresenta crescimento assintótico para a unidade, à medida que novas variáveis são incluídas ao modelo (MONTGOMERY e RUNGER, 2014). Para contornar esse problema, utiliza-se o R^2_{ajustado} (Equação 4.12) que nada mais é que o coeficiente R^2 acrescido de um termo de penalização decorrente da complexidade do modelo.

$$R^2_{\text{ajustado}} = 1 - \left(\frac{n-1}{n-p} \right) (1 - R^2). \quad (4.12)$$

Sendo: n o número de amostras testadas e p o número de variáveis contidas no modelo.

Outros índices são citados na literatura, no entanto a maioria deles baseia-se ou deriva-se dos citados nesta seção, principalmente na minimização da MQE.

4.2.3. Métodos Sequenciais (Procedimentos *Stepwise*)

Os métodos sequenciais surgiram devido à impossibilidade de avaliar os modelos gerados pelo método de busca aleatória. Eles identificam pequenos grupos de modelos de regressão que são "bons", de acordo com um critério especificado (KUTNER et al., 2004). As próximas seções descrevem os principais métodos empregados e seus algoritmos de implementação, cujo objetivo é selecionar as principais variáveis secundárias para os modelos de inferência.

4.2.3.1 Todas as Regressões Possíveis (TRP, *All Possible Regressions*)

Esse tipo de abordagem requer que sejam ajustadas todas as equações do modelo MLR envolvendo uma variável candidata, todas as equações do MLR com duas variáveis candidatas e assim por diante. Considere um modelo MLR e suas suposições (Equação 4.13). O método de TRP possibilita a análise do ajuste de todos os submodelos compostos pelos possíveis subconjuntos das p variáveis e identifica os melhores desses subconjuntos, através de um ou mais índices de desempenho citados na seção anterior.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p. \quad (4.13)$$

Suponha que exista um número de variáveis candidatas a compor um modelo definido por K elementos. Como o número de modelos para serem avaliados cresce na potência 2^K , deverá existir um termo constante β_0 entre todos os modelos e suas suposições. Supondo que se deseja avaliar um modelo que possua cinco variáveis candidatas $N=4$. Neste caso, seria necessário avaliar 16 modelos. A Tabela 4.1 apresenta os possíveis modelos existentes para o caso de 4 variáveis candidatas.

Tabela 4.1. Modelos possíveis para 4 (quatro) variáveis candidatas.

NÚMERO DE VARIÁVEIS [K]	ESTRUTURAS DO MODELO
0	β_0
1	$[X_1], [X_2], [X_3], [X_4]$
2	$[X_1 X_2], [X_1 X_3], [X_1 X_4]$
	$[X_2 X_3], [X_2 X_4], [X_3 X_4]$
3	$[X_1 X_2 X_3], [X_1 X_2 X_4]$
	$[X_1 X_3 X_4], [X_2 X_3 X_4]$
4	$[X_1 X_2 X_3 X_4]$

Esse procedimento requer cuidado, visto que um esforço computacional considerável pode ser requerido quando o número de regressores candidatos é grande. Por exemplo, em um processo de destilação com 10 conjuntos de variáveis candidatas seria necessário a avaliação de 1024 modelos, sendo que o número de modelos dobraria a cada nova variável introduzida. Entretanto, esse método é o mais recomendado quando o número de variáveis candidatas não ultrapassa 30 (trinta). Isso não constitui uma limitação na busca do melhor modelo, dados os grandes recursos computacionais altamente disponíveis (MONTGOMERY e RUNGER, 2014; KUTNER et al., 2004).

4.2.3.2. Seleção Progressiva ou por Adição (*Forward Selection*, FS)

Segundo FINKLER (2003), o método FS é uma alternativa computacional eficiente para seleção de subconjuntos. O procedimento inicia com um pressuposto de um modelo no qual não existe nenhuma variável. A construção do modelo se dá em etapas e a cada etapa, uma variável é introduzida. A melhoria do modelo pode ser avaliada através de um índice de desempenho ou através de um teste de hipótese estatístico F_{teste} . A descrição das etapas desse procedimento (algoritmo descritivo) é apresentada a seguir (KUTNER et al., 2004):

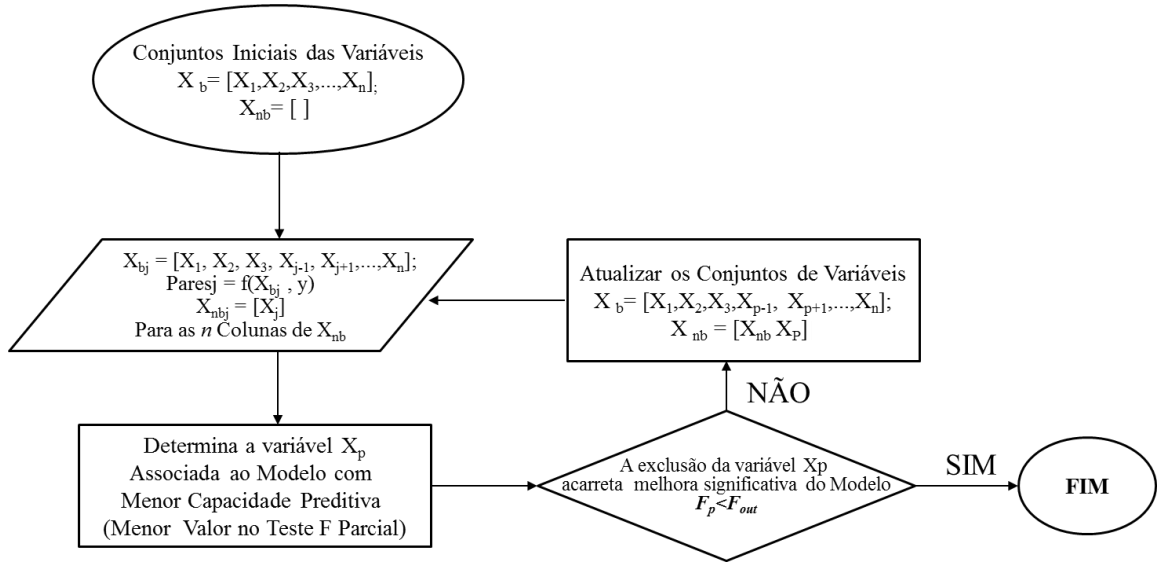
- ❖ 1ª Etapa: são criados modelos com cada uma das variáveis candidatas e a variável que apresentar melhor desempenho, é adicionada ao conjunto de variáveis que irão compor o modelo prévio.
- ❖ 2ª Etapa: são criados modelos os quais contêm combinações da variável previamente incluída, por exemplo, u_i , com cada uma das variáveis não selecionadas anteriormente. A variável que gerar o melhor modelo em combinação com aquela introduzida na etapa anterior é pré-selecionada para compor o modelo.
- ❖ 3ª Etapa: a variável pré-selecionada é submetida a um teste F_{teste} ou outros critérios de desempenho, com o intuito de verificar se a sua adição acarreta melhora significativa ao modelo;
- ❖ 4ª Etapa: se a variável pré-selecionada adicionada ao modelo resultou em sua melhoria, ela é adicionada automaticamente ao conjunto que irá fazer parte do modelo final. O procedimento é repetido para as demais variáveis.
- ❖ 5ª Etapa: repetir a 4ª etapa até que a adição de uma nova variável não resulte em melhoria significativa ao modelo ou quando todas as variáveis tiverem sido inclusas.

4.2.3.3. Regressão por Eliminação (*Backward Elimination*, BE)

Esse algoritmo começa com K variáveis/regressores candidata(o)s no modelo. Em seguida, a variável com menor F_{teste} ($F_p < F_{\text{saída}}$, a estatística F é insignificante) ou pior desempenho é removida. Em seguida, o modelo ($K-1$) regressores é ajustado e o próximo regressor para potencial eliminação é encontrado. O algoritmo termina quando nenhum regressor a mais é eliminado. A ideia desse algoritmo é obter o melhor subconjunto de variáveis através da sucessiva remoção das variáveis do modelo e a variável que foi excluída do modelo pré-selecionado é removida do modelo.

O procedimento é repetido até que a remoção de uma variável piore significativamente o modelo ou até que todas as variáveis tenham sido eliminadas. As etapas da BE são descritas no fluxograma da Figura 4.1.

Figura 4.1. Algoritmo em forma de fluxograma da Regressão por Eliminação.



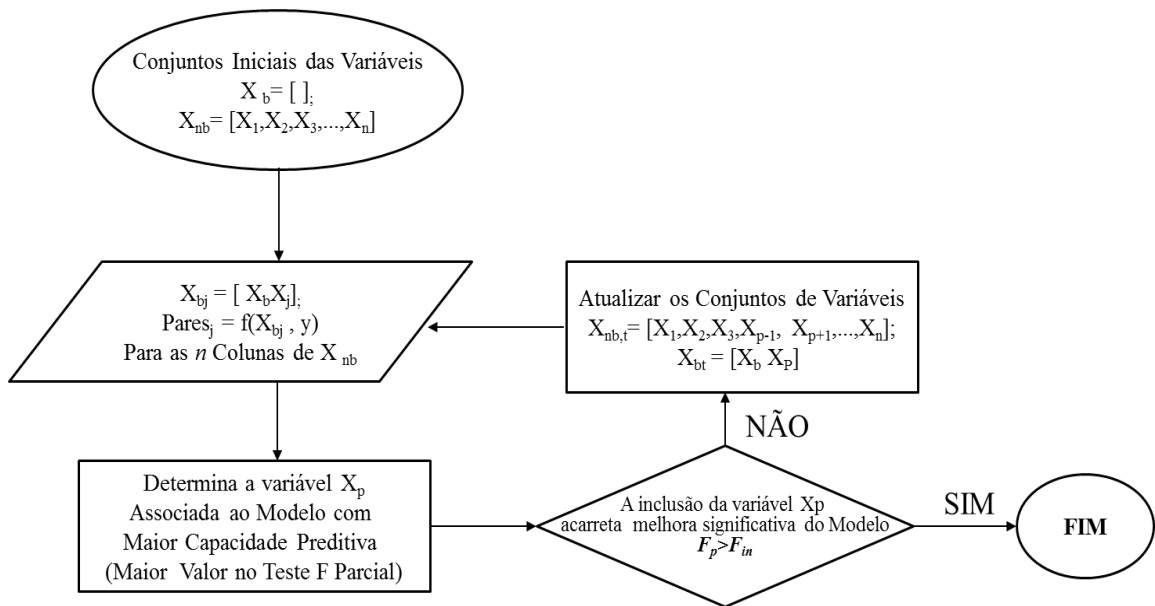
Fonte: Adaptada de FINKLER (2005) e MORAIS JR (2011).

4.2.3.4. Regressão por etapas (*Stepwise Regression, SR*)

A regressão por etapas talvez seja o procedimento mais difundido para seleção de variáveis em processos químicos (MING-DA MA et al., 2009). A técnica consiste em uma combinação dos métodos *Forward Selection* e do *Backward Elimination*, visto que os dois processos descritos anteriormente sugerem uma série de combinações possíveis. A cada adição de variável, é realizada uma etapa de verificação para eliminação de variáveis redundantes, ou seja, uma variável participou das primeiras etapas de seleção pode ser eliminada nas últimas etapas.

O procedimento inicia-se como no procedimento FS, construindo modelos com cada uma das variáveis candidatas e selecionando-se aquela que resultar no melhor modelo. A diferença é que a partir da segunda etapa, a cada inclusão de variável o procedimento *BE* é realizado, visando verificar se a variável adicionada não é redundante a outra previamente incluída. O algoritmo dessa metodologia é representado pela Figura 4.2.

Figura 4.2. Algoritmo da técnica *Stepwise Regression*.



Fonte: Adaptada de FINKLER (2003).

4.2.3.5. Algoritmo Proposto para Seleção de Variáveis

Neste trabalho foi desenvolvido um algoritmo de seleção de variáveis com base em análise de regressão, de modo que qualquer um dos procedimentos sequenciais possa ser implementado. Neste caso, o interesse é filtrar as variáveis candidatas para obter, por exemplo, um modelo de regressão que contenha o “melhor” subconjunto de variáveis para compor as entradas finais do sensor virtual. Obviamente, a experiência prévia do processo e as considerações teóricas do modelo matemático auxiliaram na seleção das variáveis candidatas. Vamos supor um modelo de múltiplas entradas e única saída (MISO), onde existam K variáveis candidatas, x_1, x_2, \dots, x_k , e uma única variável a ser inferida y . Todos os modelos devem incluir o termo de interseção β_0 , de modo que o modelo com todas as variáveis incluídas tenham $K+1$ termos. Além disso, segundo Montgomery e Runger (2014), a forma funcional de cada variável candidata, por exemplo, $x_1=1/x$, $x_2=\ln(x)$, dentre outras, é considerada correta.

O algoritmo proposto é considerado inovador no sentido de que, ao se ajustar o modelo de MLR, é possível utilizar qualquer uma das técnicas citadas na seção anterior e os resultados podem ser avaliados com quaisquer critérios de desempenho da seção 4.2.2. Esse algoritmo foi implementado no *software* Matlab® e suas principais etapas são descritas no algoritmo narrativo expresso abaixo:

1. Entre com o número de elementos da amostra N e com o número de variáveis p ;
2. Defina o tipo de técnica de seleção de variáveis (TRP, FS, BE e SR), conforme algoritmos apresentados na seção 4.2.3. Neste trabalho utilizou-se a metodologia da TRP. A definição do procedimento a ser utilizado é importante porque várias matrizes \mathbf{X} e \mathbf{Y} serão formadas, dependendo das combinações das variáveis. Por exemplo, caso se tenham 10 variáveis candidatas, serão formadas 1024 matrizes de entradas que darão origem a 1024 modelos de regressão a serem avaliados;
3. Entrada dos valores das matrizes \mathbf{X} e \mathbf{Y} :

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1N} \\ 1 & x_{21} & x_{22} & \cdots & x_{2N} \\ 1 & x_{31} & x_{32} & \cdots & x_{3N} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{m1} & x_{m2} & \cdots & x_{mN} \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix}.$$

4. Calcular a Matriz $\mathbf{X}^T\mathbf{X}$.

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} m & \sum_{i=1}^m x_{i1} & \sum_{i=1}^m x_{i2} & \cdots & \sum_{i=1}^m x_{iN} \\ \sum_{i=1}^m x_{i1} & \sum_{i=1}^m x_{i1}^2 & \sum_{i=1}^m x_{i1}x_{i2} & \cdots & \sum_{i=1}^m x_{i1}x_{iN} \\ \sum_{i=1}^m x_{i2} & \sum_{i=1}^m x_{i2}x_{i1} & \sum_{i=1}^m x_{i2}^2 & \cdots & \sum_{i=1}^m x_{i2}x_{iN} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sum_{i=1}^m x_{iN} & \sum_{i=1}^m x_{iN}x_{i1} & \sum_{i=1}^m x_{iN}x_{i2} & \cdots & \sum_{i=1}^m x_{iN}^2 \end{bmatrix}.$$

5. Obter a matriz inversa $(\mathbf{X}^T\mathbf{X})^{-1}$:

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{bmatrix} m & \sum_{i=1}^m x_{i1} & \sum_{i=1}^m x_{i2} & \cdots & \sum_{i=1}^m x_{iN} \\ \sum_{i=1}^m x_{i1} & \sum_{i=1}^m x_{i1}^2 & \sum_{i=1}^m x_{i1}x_{i2} & \cdots & \sum_{i=1}^m x_{i1}x_{iN} \\ \sum_{i=1}^m x_{i2} & \sum_{i=1}^m x_{i2}x_{i1} & \sum_{i=1}^m x_{i2}^2 & \cdots & \sum_{i=1}^m x_{i2}x_{iN} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sum_{i=1}^m x_{iN} & \sum_{i=1}^m x_{iN}x_{i1} & \sum_{i=1}^m x_{iN}x_{i2} & \cdots & \sum_{i=1}^m x_{iN}^2 \end{bmatrix}^{-1}.$$

6. Efetuar a multiplicação $\mathbf{X}^T\mathbf{Y}$:

$$\mathbf{X}^T\mathbf{Y} = \begin{bmatrix} \sum_{i=1}^m y_i & \sum_{i=1}^m x_{i1}y_i & \sum_{i=1}^m x_{i2}y_i & \cdots & \sum_{i=1}^m x_{iN}y_i \end{bmatrix}^T.$$

7. Encontrar os N coeficientes da regressão linear múltipla (β_N):

$$(\mathbf{X}^T \mathbf{X})^{-1} \cdot \mathbf{X}^T \mathbf{Y} = \beta = [\beta_0 \quad \beta_1 \quad \beta_2 \quad \dots \quad \beta_N]^T.$$

8. Obter os valores fornecidos pelo modelo:

$$\hat{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_N x_{iN}; \text{ para } i = 1, \dots, m.$$

9. Verificar através dos índices de desempenho (seção 4.2.2) os melhores ou o melhor modelo de inferência.

Vale ressaltar que, além desse algoritmo para seleção de variáveis, também foi desenvolvido um que utiliza a análise de componentes principais (PCA), por ser um dos métodos mais populares para seleção de variáveis e remoção de erros grosseiros (JOLLIFFE, 2002). A PCA foi empregada nesse trabalho com os seguintes objetivos: ser uma alternativa útil à análise de regressão na seleção das variáveis secundárias e obter o pré-processamento do conjunto de dados gerados pelas simulações. Logo, a PCA pode fornecer uma maior acurácia e confiabilidade na estimativa dos SV desenvolvidos.

4.3. REMOÇÃO DE ERROS GROSSEIROS (*OUTLIERS*)

Segundo Fortuna et al.(2007), dados coletados em uma planta podem estar corrompidos pela presença de erros “*outliers*”, isto é, observações não consistentes com a maioria dos dados e incluem pontos inexistentes, falhas de leituras, problemas de transmissão etc. Como citado anteriormente, nesse trabalho foram inseridos ruídos típicos de medições, cujos dados coletados das simulações poderiam apresentar *outliers*. A presença desses erros é um problema comum, onde os mesmos devem ser eliminados, pois um SV baseado em dados pode ter seus resultados distorcidos caso os dados possuam muitos *outliers* (LIN et al., 2007).

Uma técnica muito comum utilizada no pré-processamento de dados por estatística multivariada é a PCA. Essa técnica transforma o conjunto de dados originais em um novo sistema de coordenadas com eixos ortogonais.

4.3.1. Análise de Componentes Principais (PCA)

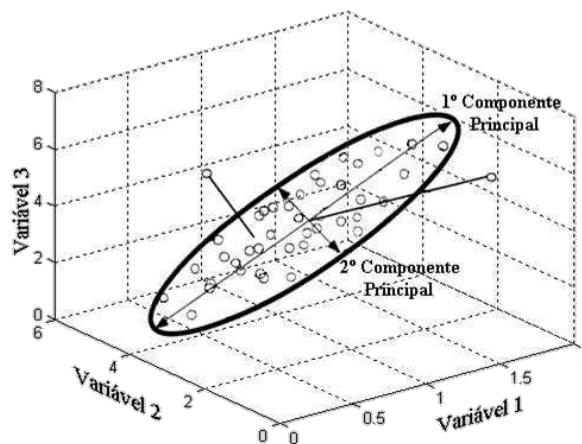
A PCA talvez seja a técnica mais antiga e mais conhecida de análise multivariada. Ela transforma um conjunto de dados originais em outro conjunto de dados o qual é chamado de

componentes principais (PC_i , com $i=1, 2, \dots, N$ componentes). Os PC_i são obtidos pela rotação dos eixos originais, de modo que os novos eixos são orientados visando descrever a maior variabilidade dos dados em uma determinada direção (AL-KANDARI e JOLLIFFE, 2001).

A justificativa do uso dessa técnica nesse trabalho parte do seguinte teorema: “vamos supor que se tenha um vetor \mathbf{x} de dimensão m e se deseje transmiti-lo usando l números, onde $l < m$. Caso o vetor \mathbf{x} possa ser truncado, um erro médio quadrático (MQE) que será igual à soma da variância (VAR) dos elementos eliminados de \mathbf{x} pode ser computado. Logo, existe uma transformação linear \mathbf{T} tal que o truncamento de $\mathbf{T}\mathbf{x}$ será ótimo no sentido do MQE, a transformação \mathbf{T} deve ter a propriedade que alguns de seus componentes tenham baixa variância. Assim, a PCA maximiza a taxa de redução da variância, sendo, portanto, ideal em um processo de seleção de regressores de redução de erros” (HAYKIN, 2001; JOLLIFE, 2002).

Considere uma matriz \mathbf{A} que possua três elementos $[x_1, x_2, x_3]$, os quais possam ser dispostos em um gráfico tridimensional, conforme ilustra a Figura 4.3.

Figura 4.3 Componentes Principais.



Fonte: Fortuna et al. (2007).

Supondo que a matriz de dados originais \mathbf{A} foi definida, uma PCA começa com dados de m variáveis para n indivíduos, cujos componentes principais estão representados no gráfico da Figura 4.4. O primeiro componente principal (PC_1) é então, a combinação linear das variáveis da matriz \mathbf{A} , elementos, $[x_1, x_2, \dots, x_m]$, fornecido pela Equação 4.14:

$$Z_1 = c_{11}X_1 + c_{12}X_2 + \dots + c_{1m}X_m. \quad (4.14)$$

A qual varia tanto quanto possível para os indivíduos, sujeitos à condição de que o segundo componente de principal (PC_2), possa ser representado pela Equação 4.15:

$$Z_2 = c_{21}X_1 + c_{22}X_2 + \dots + c_{2m}X_m. \quad (4.15)$$

Como consequência, a variância de Z_1 ($Var(Z_1)$) é tão grande quanto possível, dada esta restrição sobre as constantes c_{ij} . Essa restrição deve ser imposta, senão a $Var(Z_1)$ pode aumentar fazendo simplesmente crescer qualquer um dos valores das constantes c_{ij} . O segundo componente principal (Equação 4.15) é escolhido de modo que a $Var(Z_2)$ seja tão grande quanto possível, sujeito à restrição imposta na Equação 4.16 e à condição que Z_1 e Z_2 tenham correlação zero para os dados.

$$c_{21}^2 + c_{22}^2, \dots, c_{2m}^2 = 1. \quad (4.16)$$

O terceiro componente principal (PC_3) é fornecido pela Equação 4.17, de modo que a $Var(Z_3)$ seja tão grande quando possível e sujeito à restrição fornecida pela Equação 4.18.

$$Z_3 = c_{31}X_1 + c_{32}X_2 + \dots + c_{3m}X_m. \quad (4.17)$$

$$c_{31}^2 + c_{32}^2, \dots, c_{3m}^2 = 1. \quad (4.18)$$

É importante ressaltar que Z_3 é não correlacionada com Z_1 e Z_2 . Componentes principais (PC_i) posteriores são definidos da mesma maneira, ou seja, se existirem m variáveis, então existirão no máximo m componentes principais.

4.3.1.1. Cálculos da Análise de Componentes Principais

De fato, uma análise de componentes principais envolve encontrar os autovalores de uma matriz de covariâncias amostral. A matriz de covariâncias (COV) é simétrica e real, sendo apresentada na forma da Equação 4.19, de modo que é sempre possível encontrar um conjunto de autovetores ortonormais, onde a diagonal principal da matriz contém as variâncias e as demais posições a correlação entre as direções.

$$COV(A) = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & & c_{mm} \end{bmatrix}. \quad (4.19)$$

Sendo: o elemento c_{ii} na diagonal a variância de X_i e o termo fora da diagonal c_{ij} é a covariância

entre as variáveis X_i e X_j .

As variâncias dos PC_i são os autovalores da matriz COV. Existem m destes autovalores, alguns dos quais podem ser nulos. Autovalores negativos não são possíveis para uma matriz de covariância. Assumindo que os autovalores estão ordenados como $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$, então λ_i corresponde ao i -ésimo componente principal, representado na Equação 4.20:

$$Z_i = c_{i1}X_1 + c_{i2}X_2 + \dots + c_{im}X_m. \quad (4.20)$$

Em particular, $Var(Z_i) = \lambda_i$, e as constantes $c_{i1}, c_{i2}, \dots, c_{im}$ são os elementos do correspondente auto vetor, escalonado de modo que $c_{i1}^2 + c_{i2}^2, \dots, c_{im}^2 = 1$.

Uma propriedade importante dos autovalores é que a soma deles é igual à soma dos elementos da diagonal da matriz COV (Equação 4.21):

$$\lambda_1 + \lambda_2 + \dots + \lambda_m = c_{11} + c_{22} + \dots + c_{mm}. \quad (4.21)$$

O fato de que c_{ii} é a variância de X_i e λ_i é a variância de Z_i , significa que a soma das variâncias dos componentes principais é igual à soma das variâncias das variáveis originais. Portanto, os PC_i expressam toda variação dos dados originais.

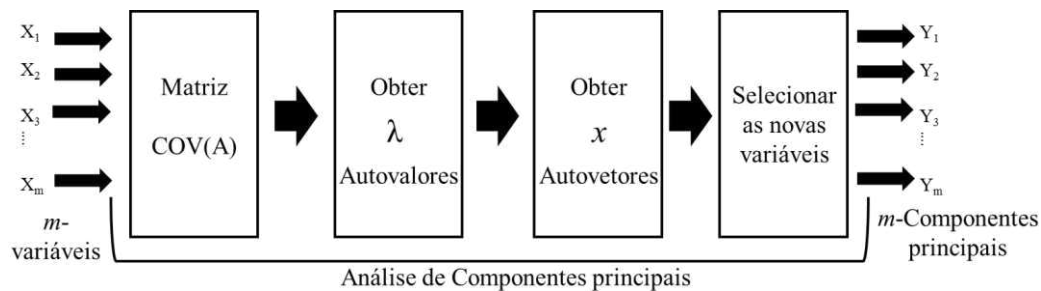
A fim de evitar que uma ou duas variáveis tenham uma influência indevida nos PC_i , é usual codificar (normalizar) as variáveis da matriz \mathbf{A} para terem médias zero e variâncias 1 (um) no início de uma análise de PCA. A matriz **COV** assume a forma da Equação 4.22.

$$COV(\mathbf{A}) = \begin{bmatrix} 1 & c_{12} & \dots & c_{1m} \\ c_{21} & 1 & \dots & c_{2m} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & & 1 \end{bmatrix}. \quad (4.22)$$

Onde: $c_{ij} = c_{ji}$ é a correlação entre X_i e X_j . Em outras palavras, a análise de componentes principais é feita sobre a matriz de correlação. Neste caso, a soma dos termos da diagonal, e, portanto, a soma dos autovalores, é igual a m , o número de variáveis da matriz \mathbf{A} .

Portanto, para determinar as componentes principais, é necessário calcular a matriz de variância-covariância, os autovalores, os autovetores e, por fim, escrever as combinações lineares, que serão as novas variáveis, denominadas de componentes principais. A Figura 4.4 apresenta o procedimento para se efetuar o cálculo análise de componentes principais.

Figura 4.4. Esquema do cálculo da análise de componentes principais.



Fonte: Adaptada de AL-KANDARI e JOLLIFFE (2001).

4.3.1.2. Metodologia da PCA para Seleção de Variáveis e Remoção de *Outliers*

Geralmente, a PCA é utilizada com os seguintes propósitos: exclusão de *outliers* e redução dimensional ou descarte de variáveis. Nesse trabalho a PCA foi utilizada com esses objetivos. Na remoção de *outliers*, efetuou-se uma análise criteriosa dos dados para identificar quais apresentam relevância/não redundância e feita uma análise gráfica dos PC_i (*scores*) selecionados. Na seleção de variáveis utilizou-se a metodologia proposta por Jolliffe (1972) e Al-Kandari e Jolliffe (2001), sendo analisadas as PC_i que têm autovalores (variâncias) maior que 70%. Assim, o descarte das variáveis baseia-se na PC_i selecionadas, onde é selecionada uma variável por componente que deve ser a variável de maior valor absoluto dos autovetores (*loading*). Para um melhor entendimento da metodologia da PCA, segue abaixo o algoritmo desenvolvido e implementado no *software* Matlab®:

1. Organizar os vetores de entrada. No primeiro passo, os dados são organizados em uma matriz $A_{n \times m}$, onde n representa o número de observações e m o número de variáveis independentes.
2. Normalizar os componentes da matriz A , variáveis X_1, X_2, \dots, X_p , visando obter médias zero e variâncias unitárias. Etapa importante para variáveis que possuem unidades diferentes (temperaturas, vazões, concentrações, etc.). Em alguns casos, essa etapa é omitida assumindo que a importância das variáveis reflete em suas variâncias;
3. Calcular a matriz de covariâncias $COV(A)$, uma matriz de correlações se o passo 2 foi efetuado;
4. Encontre os autovalores $\lambda_1, \lambda_2, \dots, \lambda_p$ e os seus correspondentes Autovetores v_1, v_2 e v_p . Os coeficientes do i -ésimo componente principal são então os elementos de v_i , enquanto que λ_i é sua variância;

5. Descartar quaisquer componentes que explicam somente uma pequena proporção da variação nos dados. Por exemplo, começando com sete variáveis, pode-se obter que os primeiros três componentes expliquem 90% da variância total. Com base nisso, os outros quatro componentes podem ser desprezados;
6. Efetuar a análise de variância explicativa e acumulada. Por fim, efetua-se a análise gráfica das PC_i para remoção de *outliers*.

4.3.2. Análise Residual para remoção de *Outliers*

A análise residual foi utilizada nesse trabalho como um critério alternativo à PCA na identificação de *outliers*. O resíduo (erro) de um modelo em uma regressão múltipla (linear ou não linear) é apresentado conforme a na Equação 4.23.

$$e_i = y_{i(\text{Real})} - y_{i(\text{Modelo})}. \quad (4.23)$$

O resíduo de um modelo de regressão múltipla tem um papel importante no julgamento sobre a adequação geral da estimativa. Análises gráficas de erros são frequentemente úteis, haja vista que os padrões de comportamento indicam que o modelo pode conter ou não amostras espúrias (MONTGOMERY e RUNGER, 2006 e 2014). Na análise de regressão múltipla, são encontrados alguns subconjuntos de observações excepcionalmente influentes, tais observações localizam-se relativamente, longe da vizinhança, onde o resto dos dados foi coletado.

A Figura 4.5 apresenta uma situação hipotética para duas variáveis. A disposição dos pontos no espaço x é importante na determinação das propriedades do modelo, haja vista que uma observação no ponto (x_{i1}, x_{i2}) está distante do resto dos dados. Assim, deseja-se examinar os pontos influentes de modo a determinar se eles controlam muitas propriedades do modelo, pontos ruins ou grosseiros (*outliers*) de algum modo, então eles devem ser eliminados.

Figura 4.5 Um ponto que está longe no espaço x .



Fonte: Adaptada de MONTGOMERY e RUNGE (2014).

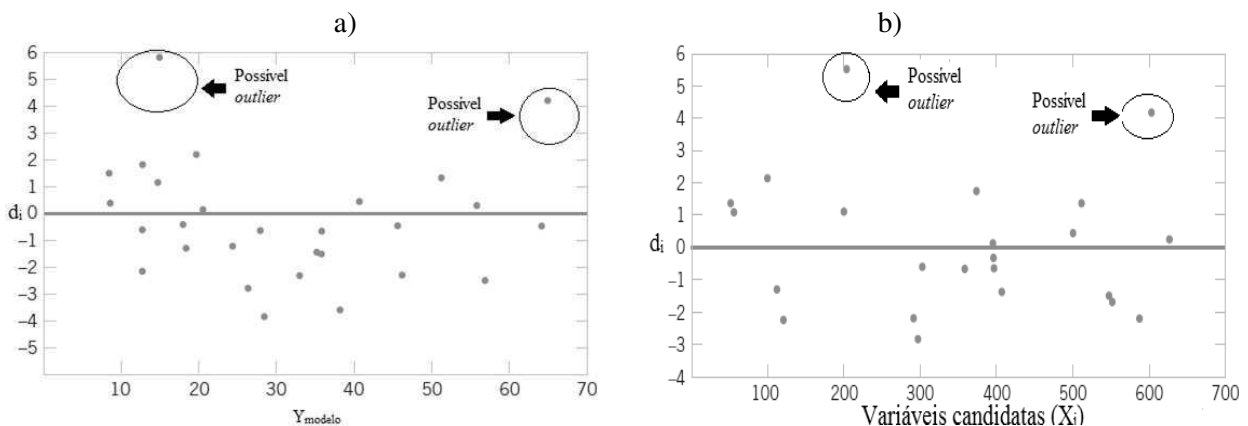
No mínimo, deve-se determinar se eles produzem ou não resultados consistentes com os restos dos dados, proporcionando resultados ruins na inferência. Alguns métodos de detecção de *outliers*, ao invés de se verificar o erro residual, são calculados outros tipos de resíduos escalonados, dentre os quais se destaca o erro residual padronizado (d_i) (Equação 4.24):

$$d_i = \frac{y_{i(\text{real})} - y_{i(\text{modelo})}}{\sqrt{\text{MQE}}} \quad (4.24)$$

O uso da equação 4.24 é frequentemente mais útil que a do erro normal, pois d_i é escalonado de modo que seus desvios-padrões sejam, aproximadamente, iguais a uma unidade. Resíduos grandes apresentam possíveis *outliers*, identificados facilmente a partir de observações dos gráficos de análises residuais das Figuras 4.6 a) e b).

Os gráficos das Figuras 4.6 a) e b) mostram, respectivamente, os erros padronizados residuais versus y_{modelo} e erros padronizados residuais versus as variáveis candidatas. Esses gráficos expressam o quanto os pontos distantes, referentes aos dados de entrada, estão do chamado agrupamento de dados coerentes (*clusters*).

Figura 4.6 a) Análise residual da saída fornecida pelo modelo versus o erro residual padronizado; e b) Análise residual das entradas X_i versus o erro residual.



Fonte: Adaptada de MONTGOMERY e RUNGE (2014).

Do mesmo modo que na PCA, o procedimento é totalmente subjetivo, haja vista que a remoção se baseia em um valor absoluto (modular) de d_i . O comum é tomar como base o mesmo valor dos limites inferior e superior, para que se possam remover os possíveis os erros grosseiros. Esse procedimento foi implementado nesse trabalho concomitante com PCA, de modo que a metodologia consiste em gerar um modelo de regressão com todas as variáveis candidatas e

assim, verificar a análise gráfica através do erro residual padronizado. Existem outros procedimentos que são baseados na reconciliação de dados e detecção dinâmica de *outliers*.

4.4. CONCLUSÕES DO CAPÍTULO

Nesse capítulo apresentou-se a metodologia de seleção de variáveis e remoção de erros *outliers*, aplicada nos dados de entrada dos sensores virtuais construídos. Sabendo que é de fundamental importância a seleção das variáveis de entrada, dentre as variáveis secundárias disponíveis, que sejam realmente relevantes para a inferência das variáveis principais (composição do CCl_4 e do CHCl_3 , corrente de base da torre de destilação do 1,2-DCE).

Para a seleção de variáveis nesse trabalho, foram desenvolvidos dois algoritmos implementados no *software* Matlab®. O primeiro consiste na análise de regressão multivariável através dos procedimentos sequenciais (*stepwise*), onde são selecionados os modelos de maior capacidade preditiva com base na minimização do MQE. O segundo algoritmo desenvolvido utiliza a técnica multivariável consolidada da análise de componentes principais, sendo uma alternativa útil caso o número de variáveis candidatas seja grande o suficiente ao ponto de inviabilizar o primeiro algoritmo.

Os dados amostrados do processo, geralmente, podem conter observações inconsistentes com as características estatísticas do restante dos dados. Essas observações relevantes podem ter uma influência indesejada nas estimativas do modelo. Desta forma, o primeiro passo no pré-processamento é identificar e remover os *outliers*. Nesse trabalho, foram aplicados os métodos da análise de componentes principais e da análise de erro residual padronizado, com a finalidade de remover erros grosseiros. Por fim, verificou-se quanto às melhorias significativas nos modelos de inferência após o tratamento dos dados originais.

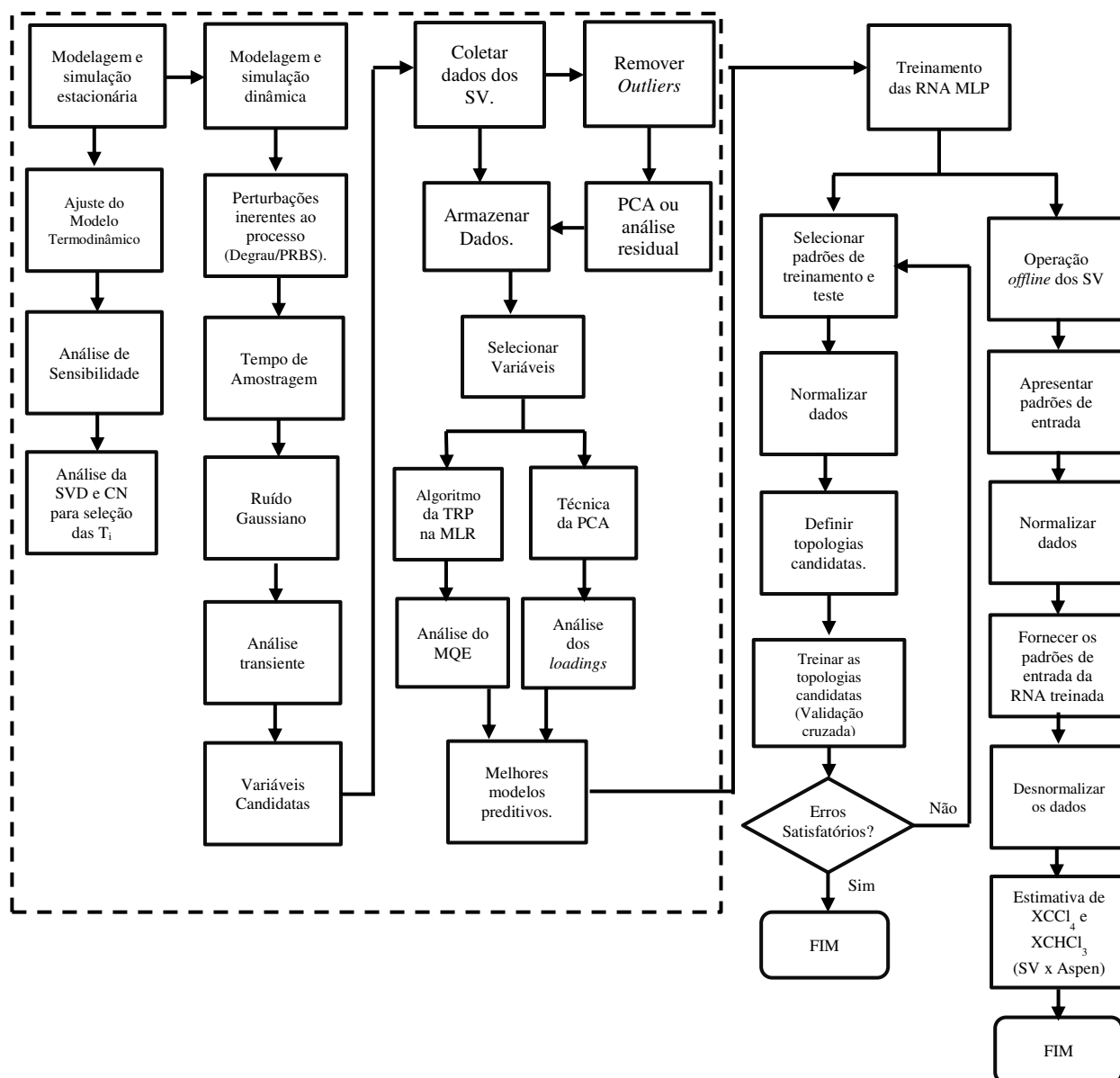
CAPÍTULO 5:

COLETA DE DADOS E AVALIAÇÃO DE MODELOS DE INFERÊNCIA

5.1. INTRODUÇÃO

Os resultados referentes ao desenvolvimento dos sensores virtuais, no processo de destilação do 1,2-DCE, são apresentados de acordo com as etapas desenvolvidas no fluxograma da Figura 5.1. Nesse capítulo são apresentados os resultados referentes à: modelagem e simulação em regime estacionário e dinâmico, seleção de variáveis e coleta de dados. Cada uma das etapas em destaque serão descritas nos próximos tópicos.

Figura 5.1. Metodologia desenvolvida para construção dos sensores virtuais.



5.2. SIMULAÇÃO DO PROCESSO

Em geral os dados de entrada de um sensor virtual (SV) são provenientes dos resultados da simulação ou das medições diretas instaladas na planta. Entretanto, o último tipo de configuração não é muito confiável, uma vez que problemas com as medições das variáveis físicas, como por exemplo, descalibração ou até a perda do instrumento, pode comprometer a estimativa de inferência. A simulação de processos em regime estacionário é utilizada rotineiramente na avaliação, seleção e desenvolvimento de novos processos. Hoje, no entanto, novos fatores devem ser considerados, relacionados principalmente com a operação e controle de plantas industriais (FORTUNA et al., 2007; LUYBEN, 2013).

Nesse contexto, a base de entendimento dinâmico e a forma de controlar as especificações exigem do engenheiro um conhecimento minucioso em regime estacionário. Além disso, no estudo do comportamento dinâmico de qualquer processo ou equipamento, faz-se necessário as condições iniciais as quais são obtidas através de simulações em regime estacionário.

A simulação em regime dinâmico deve representar, o mais fielmente possível, o comportamento da unidade de processo. Sendo de fundamental importância para os estudos específicos sobre o comportamento da torre 1 do 1,2 DCE. Levando-se em consideração o fato que os SV desenvolvidos nesse trabalho são dinâmicos, o uso de dados oriundos das simulações do modelo da torre 1 para geração das entradas requer uma avaliação criteriosa, tanto do ponto de vista de range de perturbações quanto da escolha do sinal de excitação, ruídos inseridos e sintonia dos controladores. Com as simulações, também são definidas as variáveis candidatas como entradas do modelo de inferência. Portanto, as simulações em regime estacionário e dinâmico proporcionaram a coluna diferentes condições de operação e em uma faixa na qual se espera que o sensor virtual possa operar. Logo, a representação real da torre 1 possibilita a geração do conjunto de dados para etapa de seleção de variáveis, fornecendo as entradas do sensor virtual e a topologia ideal das redes neurais a serem treinadas.

5.2.1. Simulação em regime estacionário

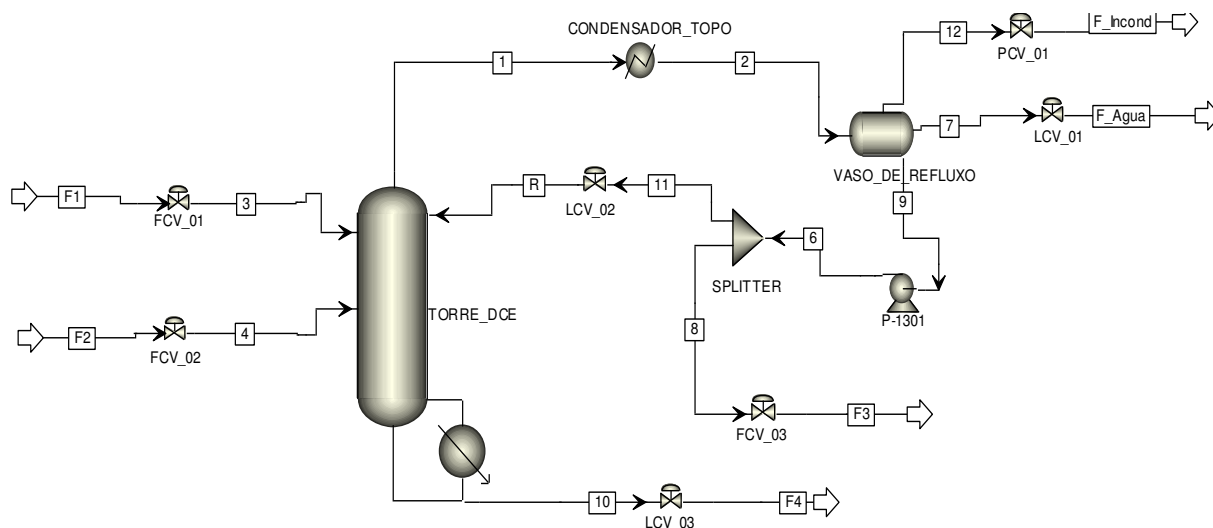
A simulação do processo em regime estacionário deste trabalho foi realizada no *software Aspen PlusTM*, na qual todos equipamentos e acessórios são instalados: coluna, refeedor, condensador, vaso de refluxo, divisor de corrente, válvulas de controle e bombas. Esses últimos,

nessa etapa, não são necessariamente obrigatórios, no entanto para conversão simultânea (*simultaneous design*) do regime estacionário para o transiente, esses itens devem ser incluídos para permitir o teste de pressão orientada (*pressure-driven*). Foram utilizados os seguintes equipamentos e especificações:

- ❖ Colunas: na rotina *RadFrac*, o modelo *strip2* para definir a coluna de destilação;
- ❖ Trocador de Calor: optou-se pela rotina *Heater*. A temperatura e a queda de pressão no condensador de topo da coluna foram fornecidas, respectivamente, com valores de 35° C e -0,1 atm.
- ❖ Variadores de Pressão: utilizou-se aqui o modelo *Pump*, para a bomba. Para a condição de operação da bomba foi definida uma pressão de descarga de 5,0 atm.

Segundo Schefflan (2011), modelos de colunas com a rotina *RadFrac*, como proposto para a torre 1, são rigorosos em se tratando de destilação e são baseados no método matemático *inside-out*. Essa rotina detecta e reconhece: uma fase de água livre ou fase de um segundo líquido em qualquer ponto da coluna, assume que em cada estágio o equilíbrio entre as fases é estabelecido e utiliza valores específicos para eficiências nos estágios. A Figura 5.2 apresenta o fluxograma em regime estacionário do processo desenvolvido no *Aspen Plus*TM.

Figura 5.2. Fluxograma da torre de purificação do 1,2-DCE desenvolvido no *Aspen Plus*TM.



Para fins de uma melhor representatividade da coluna real, através da rotina *RadFrac*TM, inseriu-se o sistema *overhead* (condensador e decantador), o modelo *reboiled absorption* que acrescenta um condensador parcial (modelo *Heater*TM) e um vaso de refluxo (modelo *Decanter*TM). Na torre 1 ocorre a condensação parcial dos vapores de topo da coluna, dando

origem a duas fases líquidas, uma orgânica saturada em H₂O e uma aquosa saturada em orgânicos, separadas no vaso decantador de refluxo. Logo, para melhor aproximação do sistema real, optou-se por utilizar dois blocos desacoplados: o da coluna e o do decantador. Foi adicionando também um separador (*Splitter*), cuja função é separar a corrente de leves (F₃) da corrente de refluxo do processo (R).

Os estágios da torre do 1,2-DCE são contados de forma descendente, sendo as duas alimentações (F₁ e F₂) inseridas, respectivamente, nos estágios 31 e 53. As eficiências dos estágios podem ser manipuladas para se adaptar aos dados reais. Nesse trabalho utilizou-se uma eficiência de Murphree de 70% para cada estágio, exceto no condensador e no refeedor, onde foi de 100%. Com os dois graus de liberdade para o processo devem ser fixadas duas variáveis. No entanto, com a coluna desacoplada do condensador perde-se um grau de liberdade, onde foi fixada a variável Q_R ou F₃.

As correntes de alimentação, F₁ e F₂, possuem um total de 18 componentes. No entanto, os que possuem quantidades significativas (maior ou igual a 10 ppm) são: 1,2-DCE (C₂H₄Cl₂), tetracloreto de carbono (CCl₄), clorofórmio (CHCl₃), água (H₂O), 1,1,2-tricloroetano (CH₃CCl₃), tetracloroetano (C₂H₂Cl₄), cloreto de etila (C₂H₅Cl), MVC (C₂H₃Cl), cloroetano (C₂H₅ClO) e cloral (C₂HCl₃O). Os dados operacionais foram fixados e as simulações foram efetuadas com as condições apresentadas na Tabela 5.1. Além disso, as unidades foram padronizadas, os parâmetros do processo pré-fixados e inseridos os valores das composições das correntes.

Tabela 5.1. Dados operacionais do Processo utilizados nas Simulações.

DESCRIÇÃO	UNIDADES	*VALOR
Vazão de alimentação da Oxidação (F _{EDCOXY})	t/h	21,00
Pressão da F _{EDCOXY}	kgf/cm ²	6,62
Temperatura da F _{EDCOXY}	°C	30,00
Prato de alimentação da Oxidação	-	31,00
Vazão de alimentação da Cloração Direta (F _{EDCDIRET})	t/h	24,75
Pressão da F _{EDCDIRET}	kgf/cm ²	7,00
Temperatura da F _{EDCDIRET}	°C	45,00
Prato de alimentação da Cloração Direta	-	53,00
Nº de Estágios (Incluindo Refeedor e Condensador)	-	72,00
Carga do Refeedor	Gcal/h	2,57
Pressão no Topo da Coluna	kgf/cm ²	1,03
Pressão na Base da Coluna	kgf/cm ²	1,60

Fonte: *Dados de um processo industrial.

5.2.1.1. Modelo Termodinâmico Ajustado

Durante as simulações estacionárias, observou-se algumas dificuldades na adequada reprodução da concentração dos compostos de interesse (CHCl_3 e CCl_4) na corrente de topo e base da torre do 1,2-DCE. O problema pode ser parcialmente explicado pelas baixíssimas concentrações envolvidas, da ordem de algumas dezenas de ppm. De fato, esse problema teve que ser corrigido, pois os sensores virtuais (um para estimativa da CHCl_3 e outro para CCl_4) devem fornecer essas respostas o mais próximo possível dos valores experimentais obtidos em análises laboratoriais. Sendo assim, foram utilizados diversos modelos termodinâmicos disponíveis no simulador *Aspen*TM. O procedimento teve o intuito de verificar e aproximar a solução do modelo quanto $X\text{CHCl}_3$, $X\text{CCl}_4$ e $X_{1,2}\text{-DCE}$, em seus valores experimentais.

As simulações que utilizaram os modelos de composição local UNIFAC forneceram valores baixos à composição dos compostos, tanto na base quanto no topo da coluna, em alguns casos o sistema não convergiu. Todos os modelos UNIQUAC mostraram-se incoerentes, apresentando valores de $X\text{CHCl}_3$ mais altos que os valores de $X\text{CCl}_4$ tetracloreto na base e o inverso no topo. Esta situação é oposta à observada na prática industrial (valores experimentais).

As simulações que utilizaram o modelo NRTL apresentaram resultados aceitáveis às concentrações, tanto para o produto de base quanto para o de topo da coluna. Bons resultados com esse modelo devem-se ao fato que o mesmo foi validado experimentalmente em sistemas de miscibilidade parcial (RENON e PRAUSNITZ, 1985), como no caso da mistura de compostos em estudo. O modelo NRTL-RK apresentou bons resultados, no entanto houve dificuldade para representar as composições de topo da coluna. O modelo NRTL com a equação de estado de Hayden-O'Connell (NRTL-HOC) apresentou valores mais coerentes para com os valores experimentais das concentrações, logo foi adotado à modelagem termodinâmica do processo. Portanto, para representar o equilíbrio entre as fases (Líquido-Vapor e Líquido-Líquido-Vapor) foi utilizada uma abordagem do tipo γ - ϕ . Onde o coeficiente de atividade da fase líquida (γ) foi determinado a partir do modelo NRTL (Equação 5.2) e o coeficiente de fugacidade da fase vapor (ϕ) foi calculado pela equação de estado de Hayden-O'Connell (Equação 5.3).

$$\ln \gamma_i = \frac{\sum_{j=1}^c (\tau_{ji} G_{ji} x_j)}{\sum_{k=1}^c (G_{ki} x_k)} + \sum_{j=1}^c \left[\frac{(x_j G_{ij})}{\sum_{k=1}^c (G_{kj} x_k)} \left(\tau_{ij} - \frac{\sum_{k=1}^c (x_k \tau_{kj} G_{kj})}{\sum_{k=1}^c (G_{kj} x_k)} \right) \right] \quad (5.2)$$

Onde: $G_{ji} = \exp(-\alpha_{ji}\tau_{ji})$.

$$\begin{aligned} \tau_{ij} &= a_{ij} + \frac{b_{ij}}{T} + e_{ij} \ln T + f_{ij} T. & a_{ij} &\neq a_{ji}. \\ \alpha_{ij} &= c_{ij} + d_{ij}(T - 273,15K). & b_{ij} &\neq b_{ji}. \\ \tau_{ii} &= 0. & c_{ij} &\neq c_{ji}. \\ G_{ii} &= 1. & d_{ij} &\neq d_{ji}. \end{aligned}$$

Os parâmetros binários $a_{ij}, b_{ij}, c_{ij}, d_{ij}, e_{ij}$ e f_{ij} podem ser determinados através de regressão de dados do ELV e/ou ELL.

$$Z_m = 1 + \frac{B_p}{RT}. \quad (5.3)$$

Onde: $B_p = \sum_i \sum_j x_i x_j B_{ij}(T)$.

$$B_{ij} = (B_{\text{free-nonpolar}})_{ij} + (B_{\text{free-polar}})_{ij} + (B_{\text{metastable}})_{ij} + (B_{\text{bound}})_{ij} + (B_{\text{chem}})_{ij}.$$

Os coeficientes $B_{\text{free-nonpolar}}, B_{\text{free-polar}}, B_{\text{metastable}}, B_{\text{bound}}$ e B_{chem} , são calculados em função da ligação química e da polaridade das espécies envolvidas.

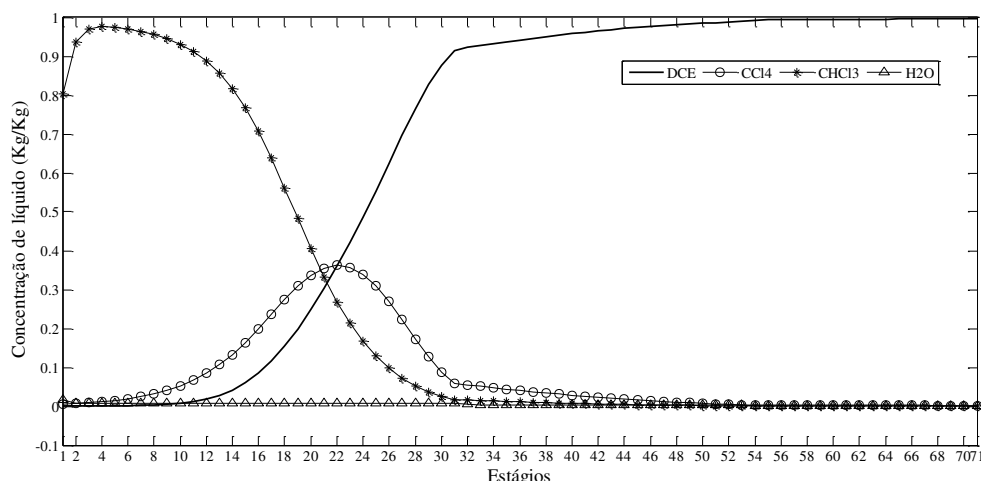
Para determinação da pressão de vapor, do calor latente de vaporização e da densidade dos componentes puros, o simulador utiliza a equação de Antoine modificada, Watson e Rackett modificada, respectivamente. As propriedades dos componentes puros e parâmetros binários foram obtidos através do banco de dados da *Aspen Plus*TM.

5.2.1.2. Resultados em Regime Estacionário

A Figura 5.3 apresenta o comportamento do perfil de concentração dos quatro principais compostos da torre, que são: o 1,2-DCE ($C_2H_4Cl_2$), CCl_4 , $CHCl_3$ e a H_2O . No gráfico da Figura 5.3, os estágios 1 e 71 representam, respectivamente, o topo e a base da coluna, percebe-se que o CCl_4 é o componente intermediário, pela inversão característica de concentração na zona de retificação.

Outra importante observação é quanto ao grau de pureza do 1,2-EDC nos estágios 31 e 53, correntes de alimentações, com aproximadamente 98,8 e 99,7%, respectivamente, o que caracteriza o processo como uma coluna destilação de alta pureza. Ou seja, a coluna utiliza 71 estágios para elevar, aproximadamente, em 0,3% a pureza da alimentação no estágio 31 e em 2% a alimentação no estágio 53, até atingir a concentração de 1,2-EDC na base da coluna.

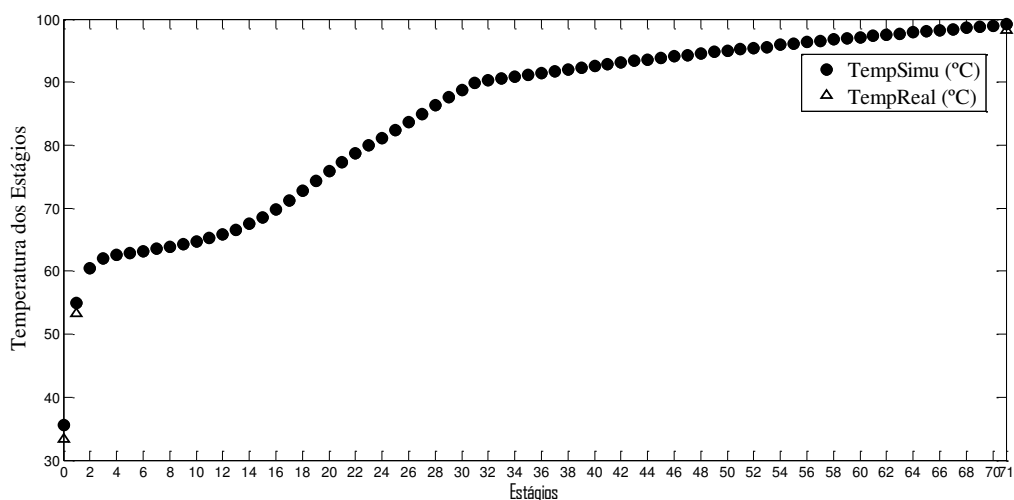
Figura 5.3. Perfil de composição líquida nos estágios da coluna.



O comportamento do CHCl_3 é caracterizado, na Figura 5.3, por uma curva acentuada e tendo como região de maior concentração o topo da coluna. A H_2O encontra-se em pequenas concentrações ao longo da coluna, aproximando-se de zero nos últimos estágios, o que confirma o comportamento azeotrópico que envolve a água e alguns componentes, referidos anteriormente. Entretanto, a concentração da H_2O foi prejudicada pela única escala (de zero a um) utilizada ao mostrar todos os compostos.

A Figura 5.4 apresenta a comparação entre os valores de temperatura encontrados na simulação e os dados reais da planta.

Figura 5.4. Perfil de temperaturas ao longo da coluna.



Conforme a Figura 5.4, verifica-se uma boa concordância entre estes pontos obtidos via simulação e àqueles fornecidos pela planta. Essa comparação somente foi possível para o prato de base (estágio 71), de topo (estágio 1) e o vaso de refluxo (estágio 0), pontos estes onde a coluna real possui elementos indicadores de temperatura.

Outra importante observação é referente à análise de sensibilidade, que fornece informações importantes sobre as respostas do processo frente às modificações em variáveis manipuladas e distúrbios. E pode ser definida da seguinte forma: “o estudo de como a incerteza presente na predição (variáveis dependentes- y) pode ser ocasionada por diferentes fontes de incerteza presentes nos diversos fatores de entrada (variáveis independentes- x e parâmetros- θ) de um modelo matemático do processo” (SALTELLI et al., 2004).

É então verificada na coluna a relação de causa-efeito, uma vez que no vaso de refluxo da coluna, praticamente, toda fase aquosa é separada, de forma que no regime estacionário as variáveis manipuladas mais relevantes são: carga do refeedor (Q_R) e vazão de retirada de leves (F_3). As variáveis de distúrbios são: vazões de alimentação, F_1 e F_2 , e as composições de alimentação. As vazões F_1 e F_2 são de fundamental relevância ao processo, pois têm influência significativa nos perfis de concentração. Considerando que F_1 e F_2 são compostas basicamente de 1,2-DCE, onde as demais concentrações estão diluídas, excluem-se os efeitos estacionários dos distúrbios composições de 1,2-DCE nessas correntes, durante a análise de sensibilidade.

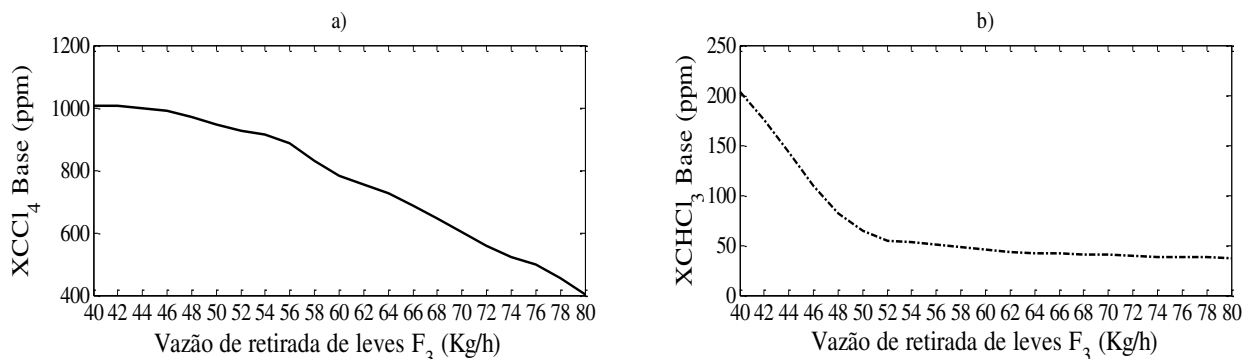
A análise de sensibilidade estacionária deve proporcionar uma relação de causa efeito de uma entrada $x_i(k)$ na saída $y_i(k)$. O estudo foi realizado modificando as variáveis F_1 , F_2 , F_3 e Q_R , verificando-se os efeitos nas concentrações de base $XCCl_4$ e $XCHCl_3$. Durante as simulações apenas uma variável foi alterada, enquanto as demais se mantiveram constantes. Vale ressaltar que as contribuições da modelagem vão além da relação causa efeito e do entendimento do problema. No entanto, a seleção das variáveis de entrada a serem efetivamente utilizadas pelo SV, em geral, não pode ser feita apenas baseada somente nesses aspectos teóricos, conforme abordagem da seção 4.2, fazendo-se necessário um estudo sobre seleção de variáveis que justifique a estrutura do modelo de inferência.

5.2.1.3. Efeito da vazão de leves (F_3)

A preocupação constante no processo de purificação do 1,2-DCE é com relação às oscilações das concentrações $XCHCl_3$ e $XCCl_4$ na base da coluna. Atualmente, na unidade industrial, para fins de controle, em malha aberta para a composição, dessas variáveis são realizadas manipulações em F_3 ou em Q_R . Logo, é essencial conhecer as consequências das modificações efetuadas na vazão de destilado F_3 . As Figuras 5.5 a) e b) apresentam, respectivamente, os comportamentos das concentrações (na fase líquida em ppm) do CCl_4 e do

CHCl_3 na base da coluna, quando a vazão F_3 é aumentada em um incremento de 2 kg/h, dentro do intervalo de 40 à 80 kg/h.

Figura 5.5. Efeito do aumento de F_3 em: a) XCCl_4 e b) XCHCl_3 .



De acordo com a Figura 5.5 a), o aumento da retirada de leves proporcionou uma diminuição significativa na concentração do componente intermediário, XCCl_4 , como consequência também há redução, menos acentuada, do componente mais volátil, XCHCl_3 , Figura 5.5 b). A redução simultânea desses compostos na base, mais acentuada para XCCl_4 , deve-se ao fato que os dois compostos possuem pontos de ebulição muito próximos.

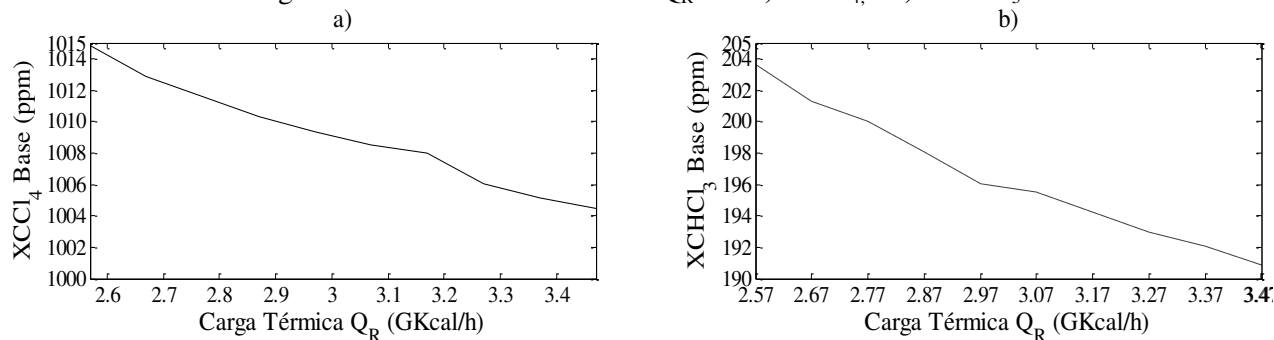
As composições, do XCCl_4 e XCHCl_3 , devem ser mantidas dentro de valores especificados, respectivamente, 3000 e 400 ppm na base da coluna. É desejado que a maior parte do XCHCl_3 seja retirado pelo topo, entretanto quando isto ocorre há também uma redução da XCCl_4 na base e de forma mais acentuada, conforme foi apresentado na Figura 5.5 a), fato indesejado no processo.

5.2.1.4. Efeito da Carga Térmica do Refervedor (Q_R)

Variações na carga térmica do refervedor (Q_R) são realizadas manipulando-se diretamente a vazão de vapor. Assim, os efeitos de Q_R sobre as concentrações dos compostos contaminantes na base da torre devem ser avaliados.

As Figuras 5.6 a) e b) apresentam, respectivamente, o comportamento da XCCl_4 e XCHCl_3 , quando Q_R é aumentada em um incremento de 0,1 Gcal/h e dentro do intervalo de 2,57 a 3,47 Gcal/h. O fato de ter escolhido esse range de perturbações, deve-se a limitação de convergência do sistema.

Figuras 5.6. Efeito do aumento de Q_R em: a) $XCCl_4$; e b) $XCHCl_3$.



De acordo com as Figuras 5.6 a) e b), um aumento progressivo em Q_R proporcionou um leve diminuição nas concentrações dos compostos da base da coluna. Essa diminuição é causada pela maior vaporização dos compostos mais voláteis nessa região, quando as demais variáveis se mantêm constantes. Em termos de escala, ocorreram variações em um range não superior a 15 ppm para ambos os compostos. Por fim, somente uma análise mais detalhada em termos de seleção de pares de variáveis como, por exemplo, a análise SVD e RGA pode fornecer uma indicação mais significativa quanto à manipulação de Q_R para controlar essas concentrações.

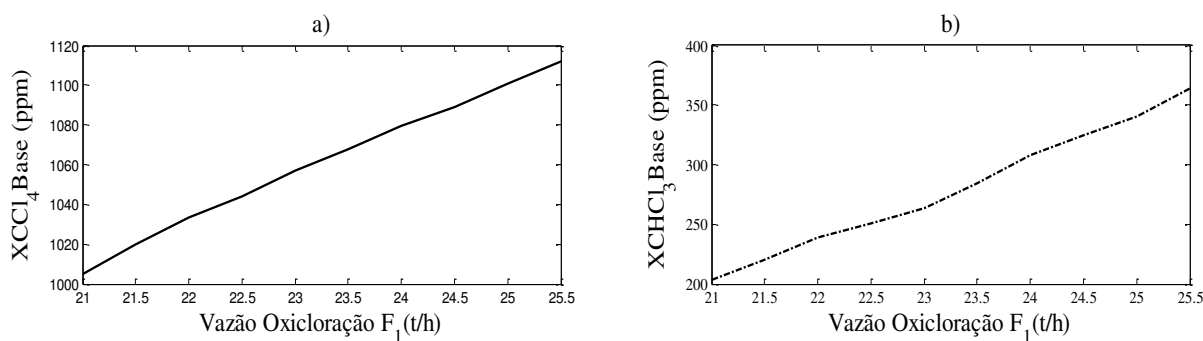
5.2.1.5. Efeito das Vazões de Alimentação da Torre (F_1 e F_2)

As correntes de alimentação da torre de 1,2-DCE, oriundas dos processos de oxicloração e cloração direta, respectivamente, F_1 e F_2 , são variáveis distúrbios na planta. Modificações nessas variáveis se devem às questões operacionais de produção. Assim, dependendo da demanda da matéria-prima (1,2-DCE), se faz necessário aumentar ou diminuir F_1 ou F_2 , essas modificações proporcionam alterações tanto no regime estacionário da coluna quanto no dinâmico.

A torre do 1,2-DCE é sensível à variação em F_1 ou F_2 , pois elas diminuem ou aumentam a eficiência de separação. Teoricamente um aumento das vazões, quando as demais variáveis se mantiverem constantes, proporcionaria uma menor vaporização na torre e, conseqüentemente, um aumento dos compostos mais voláteis na base da torre.

Assim, esses efeitos foram avaliados no estudo de análise de sensibilidade, variando uma das vazões de alimentação e mantendo-se a outra fixa. As Figuras 5.7. a) e b) apresentam, respectivamente, o efeito sobre $XCCl_4$ e $XCHCl_3$ quando F_1 é aumentada em um incremento de 0,5 t/h e dentro do intervalo de 21 até 25,5 t/h.

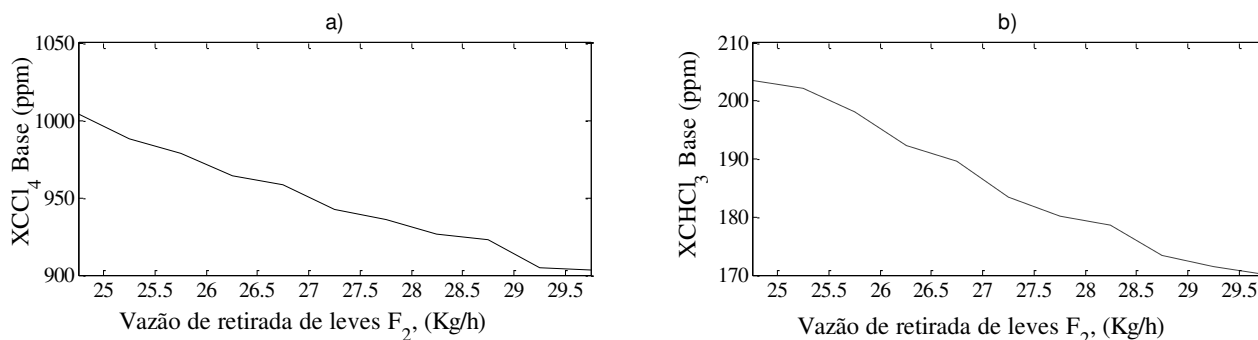
Figuras 5.7. Efeito do aumento de F_1 em: a) $XCCl_4$; e b) $XCHCl_3$.



De acordo com as Figuras 5.7 a) e b) houve um aumento progressivo e significativo das concentrações $XCCl_4$ (menos acentuada) e $XCHCl_3$, com o aumento da vazão F_1 . É possível verificar, que o componente mais volátil ($CHCl_3$) apresentou um aumento mais expressivo em termos de escala, Figura 5.7 b). O aumento dos componentes mais voláteis na base da coluna é verificado, pois com o aumento de F_1 ocorre menos vaporização da fase líquida, quando as demais variáveis são mantidas constantes.

As Figuras 5.8 a) e b) apresentam o comportamento das concentrações dos componentes de base, quando ocorre um aumento na vazão de alimentação F_2 . A variável F_2 é aumentada em um incremento de 0,5 t/h e dentro do intervalo de 24,75 para 29,75 t/h.

Figuras 5.8. Efeito do aumento de F_2 em: a) $XCCl_4$; e b) $XCHCl_3$.

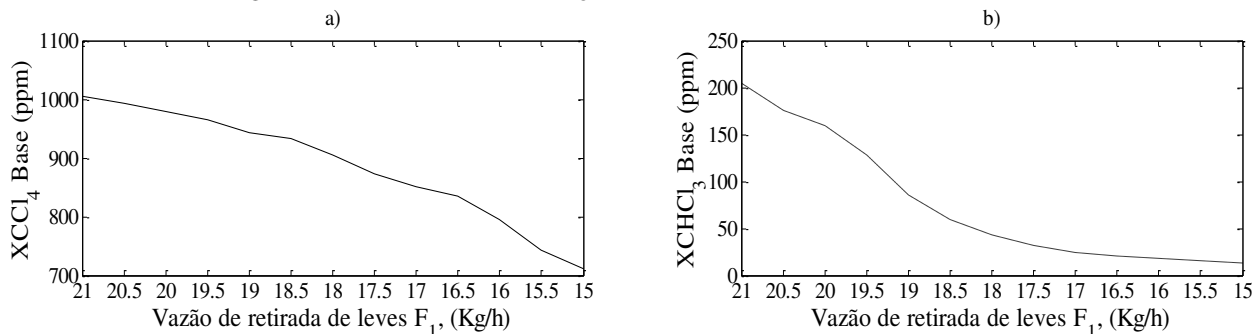


De acordo com as Figuras 6.8 a) e b), quando F_2 aumenta, progressivamente, ocorre a diminuição de $XCCl_4$ e de $XCHCl_3$. Esse efeito é o inverso ao aumento da corrente F_1 e ocorre por dois fatores: F_2 é maior 3,75 t/h que F_1 e é mais pura em 1,2-DCE (99,5 % kg/kg). Assim, proporciona uma maior diluição dos demais constituintes da mistura.

O aumento das vazões refletiu em alterações expressivas nos componentes mais voláteis da coluna, porque com o aumento de qualquer uma das vazões faz com que Q_R seja insuficiente para vaporizar a quantidade de líquido a mais que desce para a base da coluna.

Com intuito de uma maior avaliação, foram feitas simulações mediante a diminuição das referidas vazões. As Figuras 5.9 a) e b) apresentam, respectivamente, o comportamento de $XCCl_4$ e de $XCHCl_3$, quando ocorre uma diminuição progressiva em F_1 .

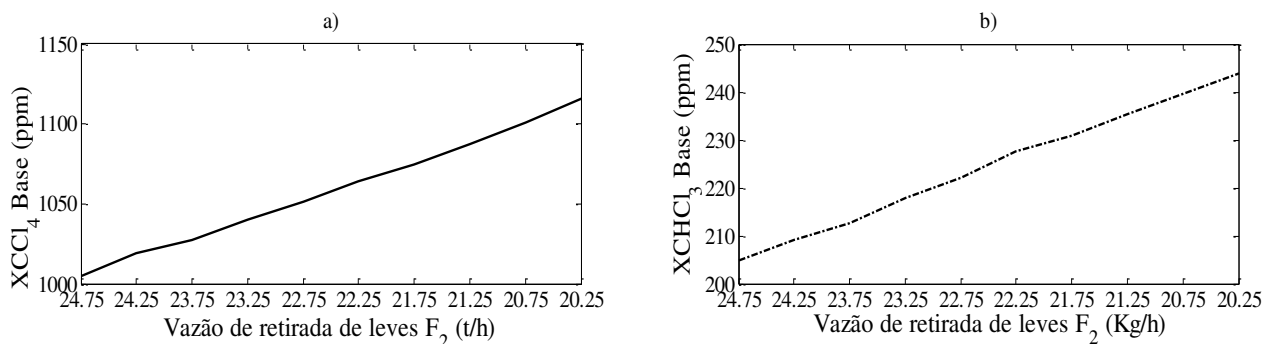
Figuras 5.9. Efeito da diminuição de F_1 em: a) $XCCl_4$; e b) $XCHCl_3$.



De acordo com os gráficos das Figuras 5.9 a) e b), confirmam que a diminuição de F_1 (de 21 para 15 t/h, passo de 0,5 t/h) proporciona uma diminuição nas concentrações $XCCl_4$ e $XCHCl_3$, o que é inverso dos efeitos proporcionados com o aumento dessa vazão.

Os testes também foram efetuados de forma avaliar o sistema com a diminuição da vazão F_2 (de 24,75 para 20,25 t/h, passo de 0,5 t/h). Assim, os gráficos das Figuras 5.10 a) e b) apresentam o comportamento das concentrações mediante tais modificações.

Figuras 5.10. Efeito da diminuição de F_2 em: a) $XCCl_4$; e b) $XCHCl_3$.



Os resultados obtidos nas Figuras 5.10 a) e b) confirmam as afirmações anteriores, já que apresentam resultados inversos aos de se aumentar F_2 . Com essas observações, percebe-se que houve um aumento tanto em $XCCl_4$ quanto $XCHCl_3$.

Os resultados obtidos nesta seção demonstram a importância da simulação em regime estacionário da coluna e, sobretudo, os efeitos produzidos nas concentrações de interesse mediante as modificações nas diversas variáveis de processos (análise de sensibilidade). Esses

resultados auxiliam tanto no entendimento do processo quanto na simulação dinâmica, além do conhecimento teórico à etapa de seleção das variáveis para a construção dos sensores virtuais.

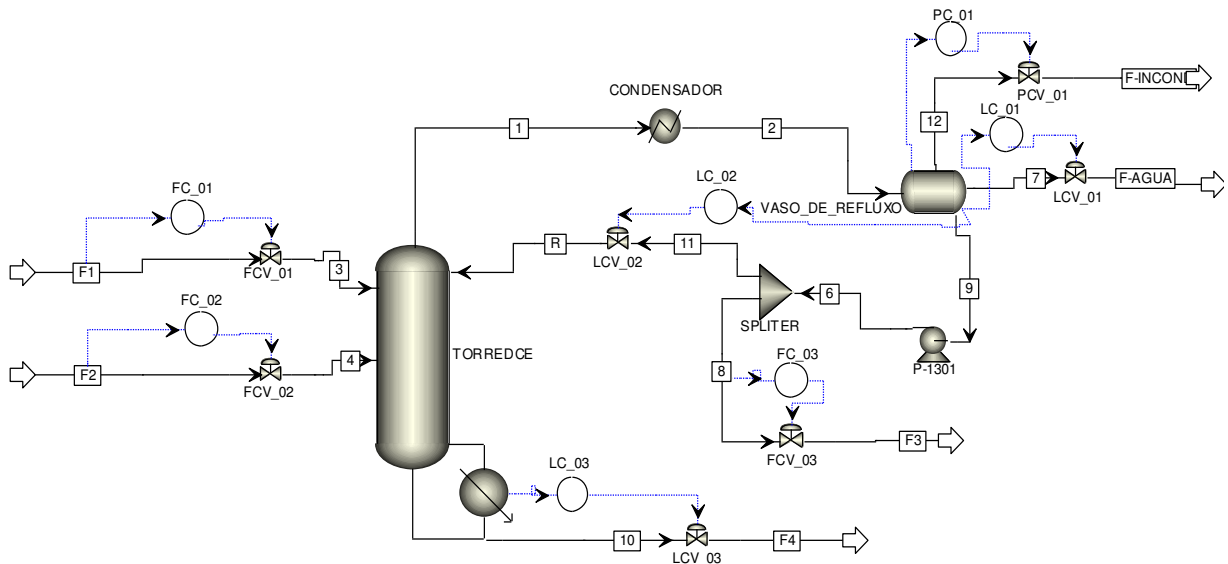
5.2.2. Simulação em Regime Dinâmico e Controle

O fato de se utilizar simulações dinâmicas no estudo da tecnologia, sensores virtuais, teve como principal motivação o fato de se ter o entendimento total do sistema. Com isso, as análises dinâmicas permitem avaliar os efeitos produzidos pelas entradas, em um range no qual se espera que o SV venha operar. O conjunto de dados gerados é discutido mediante as perturbações efetuadas em sua concepção, verifica-se aqui os efeitos produzidos nas respostas das variáveis $XCCl_4$ e $XCHCl_3$.

5.2.2.1. Estrutura de Controle

Durante o desenvolvimento do trabalho optou-se em utilizar os dados no regime dinâmico da coluna, para construção dos SV, obtidos através das simulações no *Aspen Dynamics*[®]. O fluxograma do esquema típico de controle da coluna do 1,2- DCE é apresentado na Figura 5.11. Por uma questão de sigilo industrial, foram alterados os TAG (nomes) dos instrumentos do sistema de controle, tais como, válvulas de controle e controladores.

Figura 5.11. Fluxograma de processo em regime dinâmico desenvolvido no *Aspen Dynamics*TM.



Em termos de estrutura de controle, a coluna possui um total de sete válvulas de controle e sete controladores, que ao serem implementados no simulador receberam a nomenclatura (TAGS) conforme a norma da *International Society of Automation* ISA-S-5.1 e da NBR-8190 da

ABNT. Na Tabela 5.2 são apresentadas as variáveis controladas e manipuladas, renomeadas para um melhor entendimento e sigilo do processo.

Tabela 5.2. Lista de variáveis controladas e manipuladas na Torre 1.

VARIÁVEL CONTROLADA	VARIÁVEL MANIPULADA
F_1	% de abertura da válvula da F_1
F_2	% de abertura da válvula da F_2
F_3	% de abertura da válvula da F_3
H_{H_2O}	F_{AGUA}
H_{org}	R
H_{referv}	F_4
P_{VR}	F_{INCOND}

A Tabela 5.3 apresenta os nomes dos controladores instalados no processo e o tipo de ação de controle. A ação de controle efetiva será direta quando o valor da PV↑ aumenta e o valor da MV/Sinal do controlador↑ aumenta, sendo reversa na situação contrária ao exposto.

Tabela 5.3. Lista de controladores instalados no processo e representados no simulador.

CONTROLADOR	DESCRIÇÃO	TIPO DE AÇÃO
FC_01	Controlador de vazão da corrente F_1	Reversa
FC_02	Controlador de vazão da corrente F_2	Reversa
FC_03	Controlador de vazão da corrente F_3	Reversa
LC_01	Controlador de nível da variável H_{H_2O}	Direta
LC_02	Controlador de nível da variável H_{org}	Direta
LC_03	Controlador de nível da variável H_{referv}	Direta
PC_01	Controlador de pressão da variável P_{VR}	Direta

A Tabela 5.4 apresenta os nomes das válvulas de controle do processo, inseridas ainda no regime estacionário, de suma importância no regime transiente e controle da unidade.

Tabela 5.4. Lista de válvulas de controle.

CONTROLADOR	DESCRIÇÃO
FCV_01	Válvula controladora de vazão da corrente F_1
FCV_02	Válvula controladora de vazão da corrente F_2
FCV_03	Válvula controladora de vazão da corrente F_3
LCV_01	Válvula controladora de nível H_{H_2O}
LCV_02	Válvula controladora de nível H_{org}
LCV_03	Válvula controladora de nível H_{referv}
PCV_01	Controlador de pressão da variável P_{VR}

Todos os controladores foram sintonizados com o método do relé de auto-sintonia. Mesmo as respostas das variáveis vazões, nível e pressão não sendo tão lentas quanto as de temperatura e composição, a sintonia dos parâmetros dos controladores baseada em métodos heurísticos pode comprometer os resultados nas estimativas dos sensores virtuais. Por isso, o

método do relé foi utilizado, o qual baseia-se no ganho final (K_{cu}) e no período final (T_U) da margem de estabilidade. As correções empíricas de Luyben-Tyreus (1997) foram utilizadas (Tabela 5.5), o que resulta em uma margem maior de controlabilidade para processos com constantes de tempo altas, como é o caso de colunas de alta pureza.

Tabela 5.5. Correlações para sintonia de controladores, TYREUS-LUYBEN (1997).

CONTROLADOR	K_c	τ_i	τ_D
PI	$0,31 K_{cu}$	$2,2 T_u$	–
PID	$0,45 K_{cu}$	$2,2 T_u$	$T_u/6,3$

5.2.2.2. Caracterização das Respostas em Regime transiente

A simulação dinâmica foi realizada com a estrutura de controle proposta na Figura 5.11, foram simuladas perturbações em degrau de $\pm 5\%$ em relação ao valor base das correntes F_1 e F_2 , e de $\pm 10\%$ para Q_R e F_3 . As simulações foram realizadas com um tempo final de 100 horas de operação. O objetivo de simular o processo em um tempo longo foi o de observar quando as respostas de $XCCL_4$ e $XCHC_3$ iriam alcançar o novo estado estacionário (estabilização).

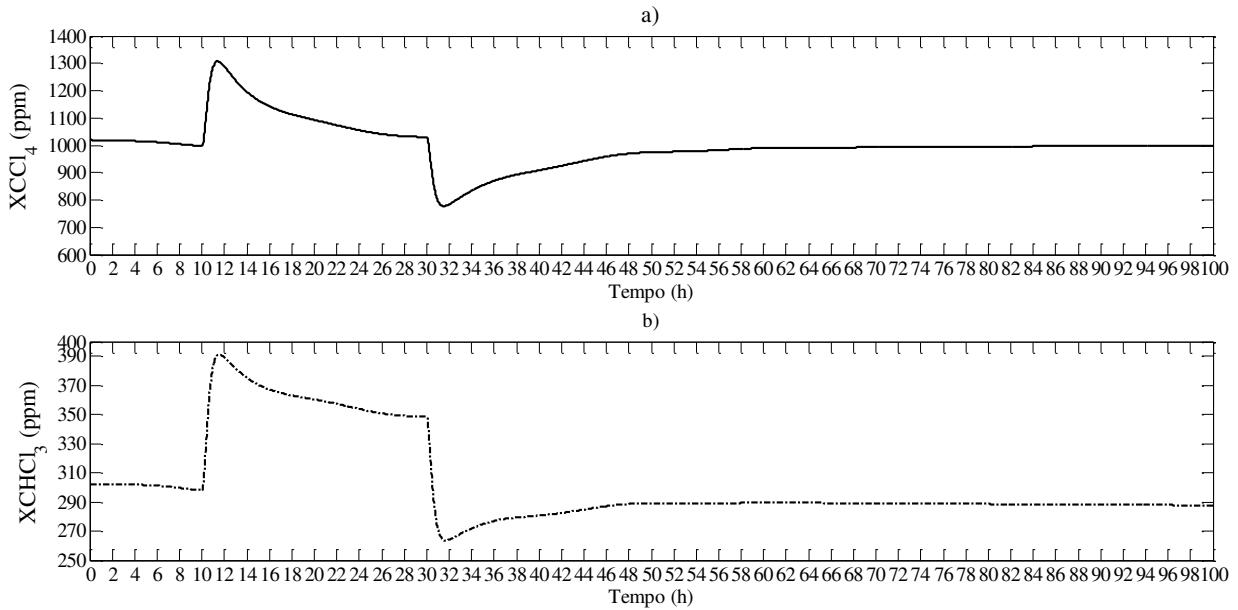
A simulação dinâmica é de extrema importância à construção dos SV, porque as observações transientes dizem a que patamar as respostas de concentrações podem alcançar. Além disso, os resultados dos sensores virtuais dinâmicos e uma futura implantação de controle por inferência dependem dessas observações. No entanto, a decisão final de quais variáveis irão compor o modelo de inferência somente será tomada na etapa de seleção de variáveis (item 5.5).

5.2.2.3. Dinâmica das Concentrações (Perturbações nas Vazões de Alimentação)

As vazões de alimentação, F_1 e F_2 , somente têm seus valores de referência (SP) alterados quando ocorrem situações anormais no processo como, por exemplo, partida/parada da planta ou demanda de produção. Essas variáveis caracterizam-se como grandes distúrbios no processo.

Os efeitos dinâmicos da variável F_1 sobre $XCCL_4$ e $XCHCl_3$ são verificados aqui. O procedimento deu-se da seguinte maneira: F_1 foi submetida às perturbações de $+5\%$ (no tempo de 10 h) e -5% (no tempo de 30 horas), percentuais em torno do SP do controlador. O último procedimento teve como objetivo retornar o processo as condições iniciais de operação. Assim, os gráficos das Figuras 5.12 a) e b) apresentam, respectivamente, o efeito de F_1 (SP=21000 kg/h) produzido em $XCCL_4$ e $XCHCl_3$.

Figura 5.12. Efeito das perturbações degrau efetuadas em F_1 sobre: a) X_{CCl_4} e b) X_{CHCl_3} .

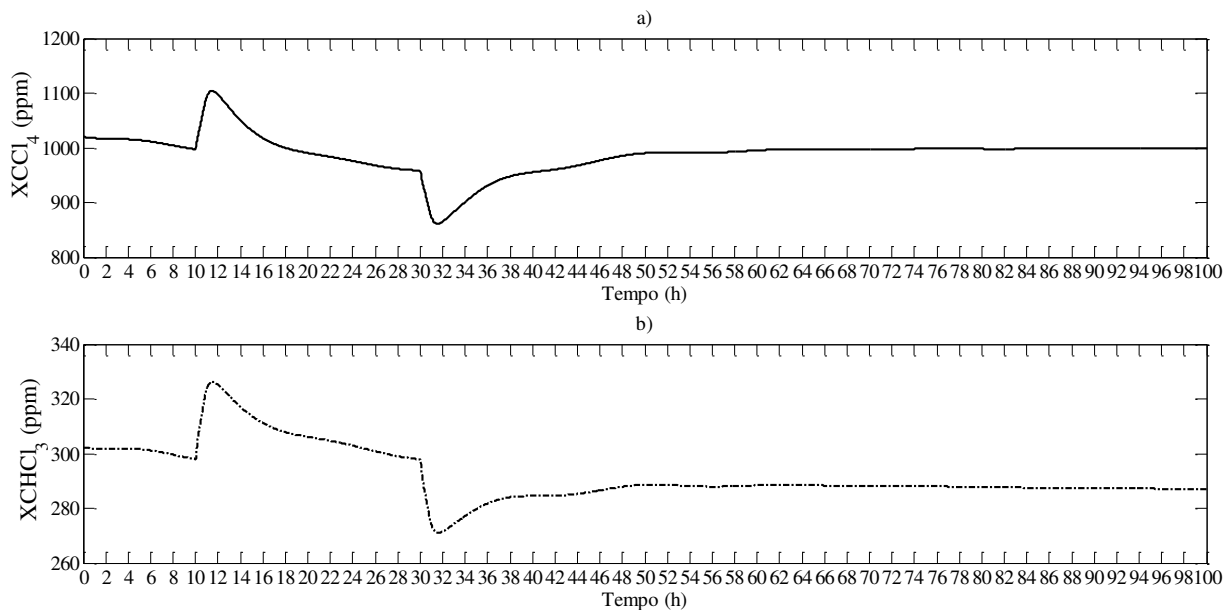


Nos gráficos das Figuras 5.12 a) e b) observar-se que ambas respostas de concentração estão em direções opostas das finais. Ou seja, X_{CCl_4} e X_{CHCl_3} tendem a aumentar (com +5% no SP de F_1 , no tempo de 10h) ou diminuir (com -5% no SP de F_1 , no tempo de 30h), logo após seguem direção contrária. O que caracteriza um transiente de respostas inversas de concentração. Em termos comparativos, a sobre-elevação (*overshoot*) percentual de ambas as respostas são bem similares, em torno de 30%. Essa última observação indica que X_{CCl_4} e X_{CHCl_3} apresentam tempos de primeiro pico (t_p) idênticos e são proporcionalmente sensíveis às perturbações em F_1 .

Do mesmo modo que para F_1 , foram efetuadas perturbações na vazão F_2 de +5% (no tempo de 10 h) e -5% (no tempo de 30 horas), percentuais em torno do SP do controlador. Os gráficos das Figuras 5.13 a) e b) apresentam, respectivamente, o efeito de F_2 (SP=24750 kg/h) produzido em X_{CCl_4} e X_{CHCl_3} .

De acordo com os gráficos das Figuras 5.13 a) e b), observar-se que ambas respostas de concentração estão em direções opostas das finais. Assim, o transiente de X_{CCl_4} e X_{CHCl_3} caracterizam-se como respostas inversas de concentração, pois tendem a aumentar (com +5% no SP de F_2 , no tempo de 10h) ou diminuir (com -5% no SP de F_2 , no tempo de 30h), logo após seguem direção contrária.

Figura 5.13. Efeito de perturbações degrau efetuadas em F_2 sobre: a) X_{CCl_4} e b) X_{CHCl_3} .



Os efeitos produzidos por F_2 , Figuras 5.13 a) e b), proporcionaram um *overshoot* de 10% para X_{CCl_4} , que é ligeiramente maior do que o da resposta X_{CHCl_3} , cerca de 8,33%. Essa última observação indica que X_{CCl_4} é um pouco mais sensível dinamicamente que X_{CHCl_3} , ao serem inseridos degrau na variável F_2 .

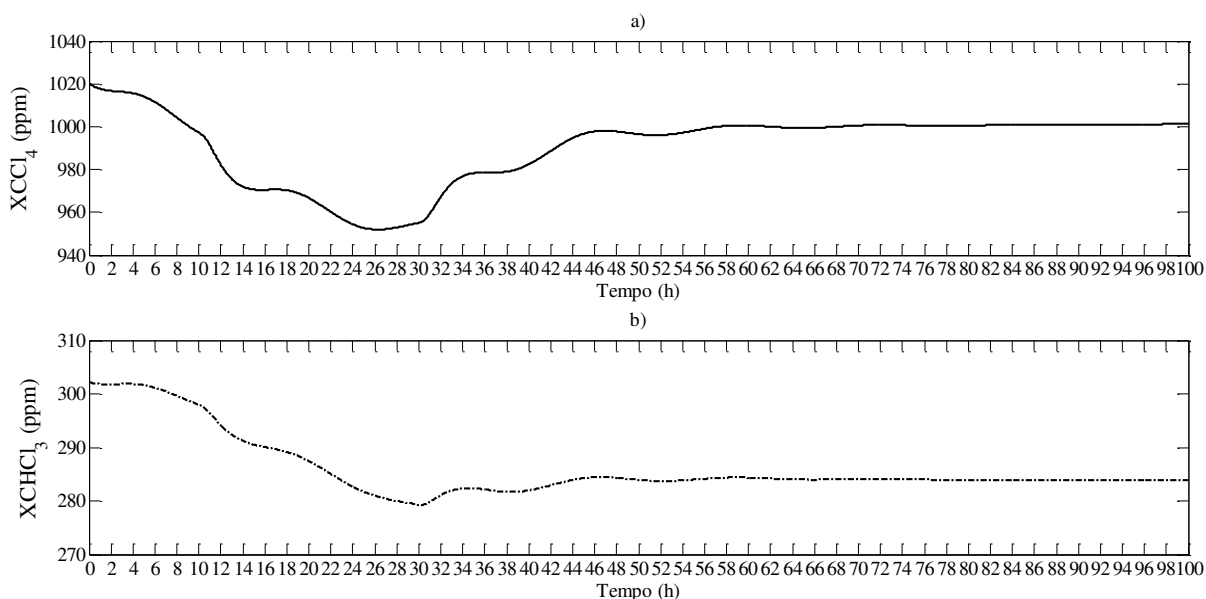
Por fim, de acordo com as análises transientes efetuadas nessa sessão, verifica-se que as concentrações dos compostos mais voláteis na base torre de 1,2-DCE apresentam respostas reversas e são mais sensíveis a modificações na vazão proveniente do processo de oxicloração (F_2). Essa observação é relevante e justificável, principalmente, por essa corrente ser maior em unidade (Kg/h) e menos rica nessas impurezas.

5.2.2.4. Dinâmica das Concentrações (Perturbações na Vazão de Destilado, F_3)

A vazão de retirada de leves é uma variável que é utilizada com frequência pela operação da planta como variável manipulada e assim, alcançar o objetivo de manter dentro dos limites de especificações as concentrações do X_{CHCl_3} e do X_{CCl_4} . Logo, a avaliação transiente das concentrações de base mediante distúrbios inseridos em F_3 deve ser verificada.

Aqui a coluna é submetida à perturbação degrau em F_3 , modificações de +10% (no tempo de 10 h) e -10% (no tempo de 30 horas), percentuais em torno do SP do controlador. Os gráficos das Figuras 5.14 a) e b) apresentam, respectivamente, o efeito de F_3 (SP=40 kg/h) produzido em X_{CCl_4} e X_{CHCl_3} .

Figura 5.14. Efeito de perturbações degrau efetuadas em F_3 sobre: a) X_{CCl_4} e b) X_{CHCl_3} .



Conforme apresentado nos gráficos das Figuras 5.14 a) e b), as concentrações dos compostos CCl_4 e $CHCl_3$ comportamento inverso às modificações efetuada em F_3 . Verifica-se que a resposta da variável X_{CCl_4} apresentou uma sobre-elevação percentual menor (aproximadamente 7%) em comparação com a resposta de X_{CHCl_3} (aproximadamente 8%). Logo, percebe-se que a concentração do X_{CHCl_3} é mais sensível às perturbações em F_3 .

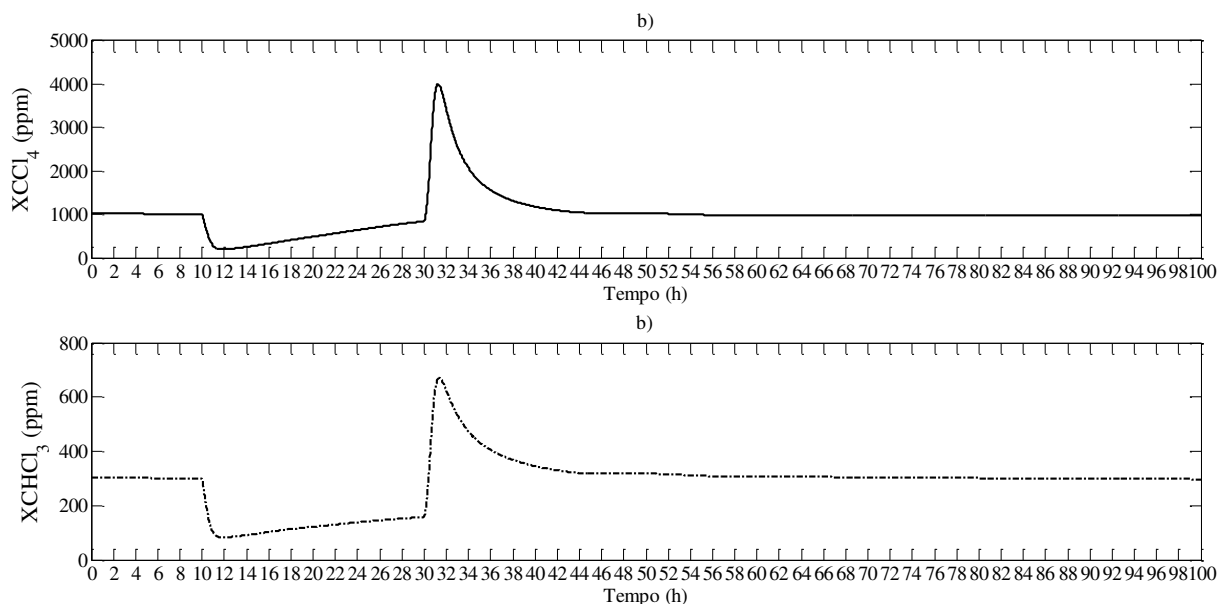
Como dito anteriormente, F_3 é uma variável comumente empregada para fins de manutenção das concentrações na base da coluna, haja vista que os maiores distúrbios da planta se concentram nas duas vazões de alimentação. Sendo assim, essa variável é incluída como uma variável candidata no estudo de inferência.

5.2.2.5. Dinâmica das Concentrações (Perturbações na carga térmica, Q_R)

Outra variável de processo, também avaliada em regime estacionário, utilizada com frequência pela operação para fins de controle de X_{CCl_4} e X_{CHCl_3} é a carga térmica do refeedor, na prática manipulada através da vazão de vapor para o refeedor da coluna. Desse modo, os efeitos transientes que essa variável produz em X_{CCl_4} e X_{CHCl_3} são avaliados aqui.

O processo é então submetido às perturbações degrau em Q_R de +10% (no tempo de 10 h) e -10% (no tempo de 30 horas), percentuais em torno do valor estacionário de $Q_R = 10,7601$ GJ/h. Os gráficos das Figuras 5.15 a) e b) apresentam, respectivamente, os efeitos transientes produzidos por Q_R nas saídas X_{CCl_4} e X_{CHCl_3} .

Figura 5.15. Efeito de perturbações degrau efetuadas em Q_R sobre: a) X_{CCl_4} e b) X_{CHCl_3} .



Tanto no gráfico da Figura 5.15 a) quanto o da Figura 5.15 b) verificam-se respostas inversas de concentração. Pois, X_{CCl_4} e X_{CHCl_3} tendem a diminuir (com +10% em Q_R , no tempo de 10h) ou aumentar (com -10% em Q_R , no tempo de 30h), logo após seguem direção contrária. Em termos comparativos, a sobre-elevação percentual de ambas as respostas são similares e bastante relevantes, em torno de 75% para a modificação de +10% em Q_R .

A última observação indica uma situação em que a variável carga térmica é importante para o controle das composições de base, haja vista tamanha sensibilidade produzida. Essa variável pode ser uma boa indicação para o sensor virtual ou controle, pois como o processo é de alta pureza, as vazões F_1 e F_2 também apresentam alta significância nas composições, entretanto elas não podem ser utilizadas como manipuladas porque são distúrbios na planta.

5.2.2.6. Dinâmica das Concentrações (Perturbações nas concentrações das correntes de alimentação).

As concentrações dos compostos CCl_4 e $CHCl_3$ das correntes de alimentação, F_1 e F_2 , são distúrbios não diretamente mensuráveis da planta, sendo necessárias análises de laboratórios para obtenção dos valores dessas variáveis. No entanto, essas variáveis constituem distúrbios de interesse no processo e têm sido utilizadas como entradas de sensores virtuais em colunas de alta pureza, como nas propostas dos trabalhos de Luo et al. (1995) e Kano et al. (2000).

No processo em estudo, os valores em fração mássica dos compostos CCl_4 e CHCl_3 são menores na vazão F_2 , pois nessa corrente o composto 1,2-DCE é mais concentrado (99,51% kg/kg). Mesmo estando em pequenas proporções (em torno de ppm), as concentrações de CCl_4 e CHCl_3 na alimentação têm um efeito dinâmico significativo sobre XCCl_4 e XCHCl_3 .

Diante dessas informações, nesse trabalho é proposta uma avaliação dinâmica investigativa sobre as respostas de concentrações de interesse, XCCl_4 e XCHCl_3 . Para tal objetivo, foram inseridos sinais binário pseudo-randômico (PRBS) com amplitudes de $\pm 20\%$ e com tempo total de simulação de 100 horas. O sinal PRBS foi inserido em cada uma das concentrações dos compostos CCl_4 e CHCl_3 das alimentações F_1 e F_2 , vale ressaltar que os distúrbios foram inseridos individualmente e não concocoritamente.. A Tabela 5.6 apresenta a amplitude do sinal PRBS inserido e os valores do caso base das concentrações.

Tabela 5.6. Rotina de perturbações nas concentrações de alimentação.

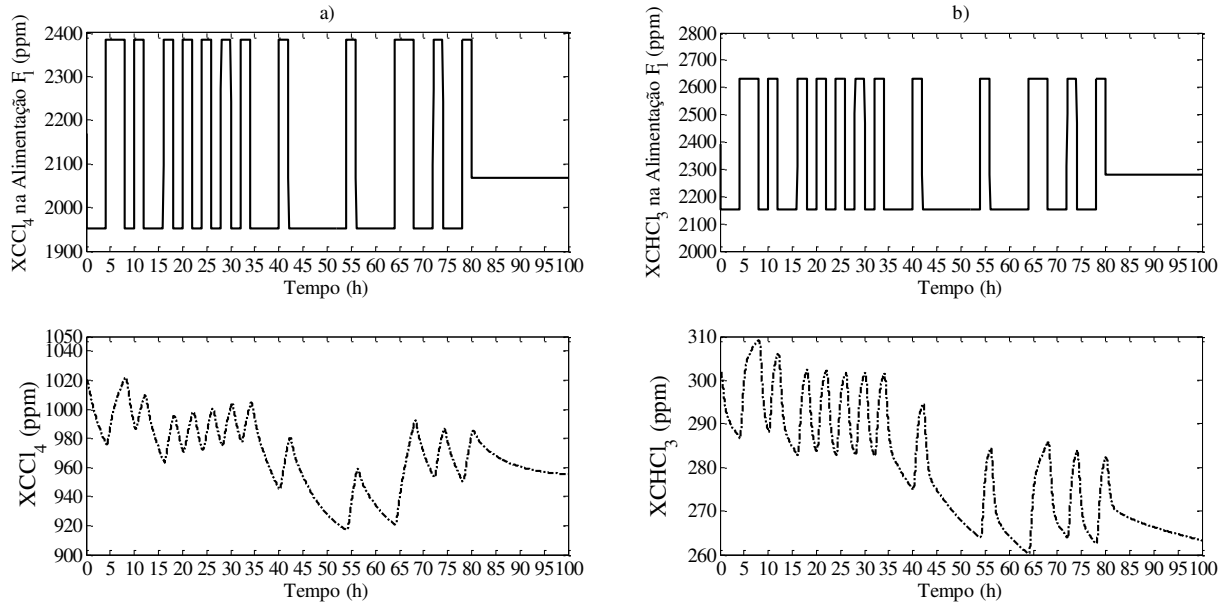
CONCENTRAÇÕES(Kg/Kg)	ALIMENTAÇÃO	CASO BASE	(+20%)	(-20%)
XCCl_4	F_1	0,002169	0,0026028	0,0017352
XCHCl_3		0,002392	0,0028704	0,0019136
XCCl_4	F_2	$2,9 \times 10^{-05}$	$3,19 \times 10^{-05}$	$2,62 \times 10^{-05}$
XCHCl_3		$2,2 \times 10^{-05}$	$2,42 \times 10^{-05}$	$1,98 \times 10^{-05}$

A finalidade de efetuar tais perturbações aleatórias é para verificar se as concentrações das impurezas, XCCl_4 e XCHCl_3 , acompanhariam o patamar de inserção das mesmas. É importante notar que essas modificações transientes ocorreram quando o processo se encontrava em malha fechada, ou seja, os controladores das vazões F_1 , F_2 e F_3 estavam atuando, haja vista a necessidade de somente serem avaliados os efeitos dos distúrbios inseridos.

O gráfico da Figura 5.16 a) apresenta o comportamento da XCCl_4 , quando a concentração do CCl_4 da alimentação F_1 está submetido ao sinal PRBS e as demais concentrações das alimentações se mantêm constantes. Do mesmo modo, o gráfico da Figura 5.16 b) apresenta o comportamento da XCHCl_3 , quando a concentração do CHCl_3 da alimentação F_1 está submetido ao sinal PRBS e as demais concentrações das alimentações se mantêm constantes.

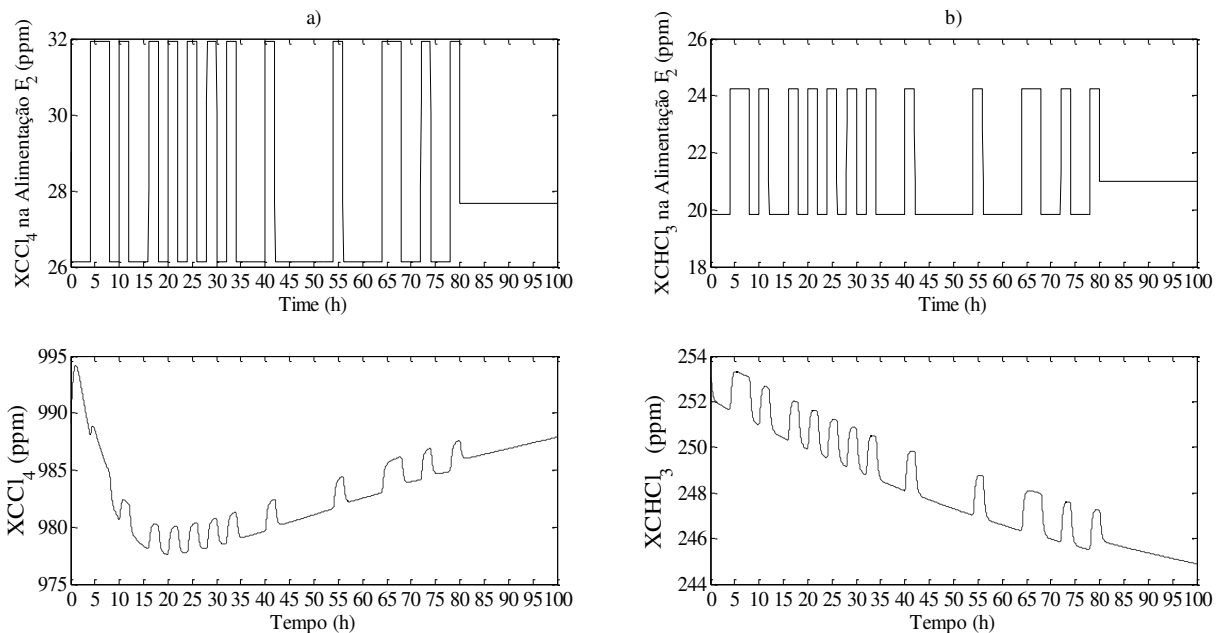
De acordo com a Figura 5.16 a), as modificações efetuadas na concentração do CCl_4 da alimentação proporcionaram oscilações em XCCl_4 , amplitude máxima de 100 ppm. A concentração de saída XCHCl_3 também desmostrou acompanhar às perturbações aleatórias, sendo proporcionalmente mais sensível, ver Figura 5.16. b).

Figura 5.16. a) PRBS na concentração do CCl_4 da alimentação F_1 e seu efeito sobre XCCl_4 ; b) PRBS na concentração do CHCl_3 da alimentação F_1 e seu efeito sobre XCHCl_3 .



Conforme ilustrado no gráfico da Figura 5.17 a), a concentração do CCl_4 na alimentação F_2 está submetida ao sinal PRBS, com demais concentrações das alimentações mantendo-se constantes, o comportamento da XCCl_4 acompanha o transiente da perturbação. A resposta de XCHCl_3 aparenta ser mais sensível, conforme é ilustrada na Figura 5.17 b), às alterações da sequência de sinais PRBS inseridos na concentração do CHCl_3 da alimentação F_2 , estando as demais concentrações nas alimentações inalteradas.

Figura 5.17. a) PRBS na concentração do CCl_4 da alimentação F_2 e seu efeito sobre XCCl_4 ; b) PRBS na concentração do CHCl_3 da alimentação F_2 e seu efeito sobre XCHCl_3 .



Conforme apresentado na Figura 6.21 a), quando a concentração do CCl_4 da alimentação F_2 é submetida às perturbações PRBS (amplitude máxima de $\pm 20\%$), nota-se que $X\text{CCl}_4$ aparenta inversão em seu comportamento. Do mesmo modo, a resposta de $X\text{CHCl}_3$ acompanha o patamar da sequência de PRBS inseridos na concentração de alimentação do CHCl_3 . Entretanto, as respostas $X\text{CCl}_4$ e $X\text{CHCl}_3$ demonstram ser menos sensíveis aos sinais das composições em F_2 do que em F_1 . Logo, diante dessas observações, as composições do CCl_4 e CHCl_3 na alimentação F_2 são proporcionalmente maiores que em F_1 , motivo pelo qual $X\text{CCl}_4$ e $X\text{CHCl}_3$ serem mais sensíveis tais modificações.

A importância de avaliar os efeitos das impurezas na alimentação sobre as saídas na base da coluna vai além das perturbações efetuadas, porque esses distúrbios podem contribuir diretamente na robustez dos sensores virtuais, aumentando assim o grau de confiança do modelo. Por fim, o teor do CCl_4 e CHCl_3 nas alimentações F_1 e F_2 irão compor o grupo de variáveis candidatas, verificando quanto a necessidade ou não de uso dessas variáveis pelos SV.

5.2.2.7. Temperaturas dos Estágios

As temperaturas dos estágios são variáveis que têm estreita relação com as composições de destilado e de base em um processo de destilação. No entanto, é característica dos processos de alta pureza as mínimas variações de temperaturas, sendo assim o uso somente dessas variáveis na estimativa de composição irrelevante ou, em alguns casos, impraticáveis para SV. No entanto, o uso das temperaturas junto com outras variáveis secundárias do processo, pode ou não trazer uma melhor acurácia para os sensores virtuais.

Verificou-se no estudo aqui desenvolvido, durante as simulações do processo, que ocorrem também mínimas variações de pressão entre a base e o topo da coluna, por isso o estudo referente aos efeitos das pressões nas concentrações do CCl_4 e CHCl_3 foram considerados como irrelevantes. Nesse ponto de vista, existe uma relação direta entre as temperaturas e as concentrações líquidas nos estágios, que deve ser avaliado como relevante ou irrelevante à perspectiva de construção dos SV.

Como descrito anteriormente, o processo possui apenas três sensores de temperatura: no estágio de base (estágio 71), de topo (estágio 1) e o vaso de refluxo (estágio 0) da coluna. No entanto, o uso de apenas um valor de temperatura pode não condizer com a real necessidade do sensor virtual. Em trabalhos como os de Kano et al. (2009) e (2000), as entradas do SV foram, além de outras variáveis, valores de temperaturas distribuídas ao longo da coluna.

Tomando como base esses trabalhos, o objetivo é localizar e selecionar as temperaturas dos estágios que possuem influência direta nas concentrações da base da coluna. O estudo foi realizado com base na matriz de ganho estacionário (SGA). Com a matriz SGA de sensibilidade da coluna, foram realizados os cálculos da decomposição de valores singulares (SVD), mencionados na sessão 2.6.1. A sensibilidade de cada temperatura dos estágios em relação a um par de variáveis manipuladas (MV) pode ser representada em forma de uma matriz, conforme Equação 5.3. Aqui T_j indica temperatura no estágio j e MV_i são as i variáveis manipuladas avaliadas no estudo.

$$K_{2 \times 71} = \begin{bmatrix} \left. \frac{\partial T_1}{\partial MV_1} \right|_{MV \neq 1} & \dots & \left. \frac{\partial T_j}{\partial MV} \right|_{MV \neq i} & \dots & \left. \frac{\partial T_{71}}{\partial MV} \right|_{MV \neq i} \\ \left. \frac{\partial T_1}{\partial MV_2} \right|_{MV \neq 1} & \dots & \left. \frac{\partial T_j}{\partial MV} \right|_{MV \neq i} & \dots & \left. \frac{\partial T_{71}}{\partial MV} \right|_{MV \neq i} \end{bmatrix}. \quad (5.3)$$

A partir dos resultados das simulações em regime estacionário, deu-se início à metodologia da SVD. Na análise da SVD foram selecionadas algumas variáveis manipuladas, são elas: F_1 , F_2 , F_3 , Q_R e R . Uma observação importante deve ser feita em torno do refluxo da coluna (R), pois com os dois graus de liberdade do processo fixou-se a variável Q_R ou F_3 . No entanto, durante as observações da SVD mantinha-se fixa Q_R e (F_3 ou R), tendo em vista de se observar o efeito do refluxo na coluna.

As análises da SVD foram feitas combinando os pares das variáveis (Variável Manipulada- MV_i) com as temperaturas dos estágios (Variável de Processo- PV_j). Para fins de se implantar a sensibilidade da malha, foram efetuadas perturbações de magnitude +10% no valor estacionário individual de cada uma das MV_i e observado as alterações nas PV_j . A Tabela 5.7 apresenta as rotinas das perturbações das variáveis em comparação com o valor base.

Tabela 5.7. Rotina de perturbações para o cálculo da SVD.

VARIÁVEIS	CASO BASE	PERTURBAÇÃO (+10%)
R (Kg/h)	25906,108	28496,7188
Q (Gcal/h)	2,570	2,827
F_3 (Kg/h)	40,000	44,000
F_1 (t/h)	21,000	23,100
F_2 (t/h)	24,750	27,225

Os valores dos ganhos estáticos foram obtidos para cada um dos pares combinados em relação a T_j da coluna. Com esses valores foi gerada a $K_{n \times m}$, Equação 5.3, decomposta em três

outras matrizes, que são: U , S e V^T . A partir dos resultados da SVD, observou-se os valores dos números condicionais (NC) para cada par combinado. A Tabela 5.8 apresenta os resultados das combinações e seus respectivos valores de NC, em destaque os melhores pares combinados frente às variáveis controladas (temperaturas).

Tabela 5.8. Resultados do Método SVD.

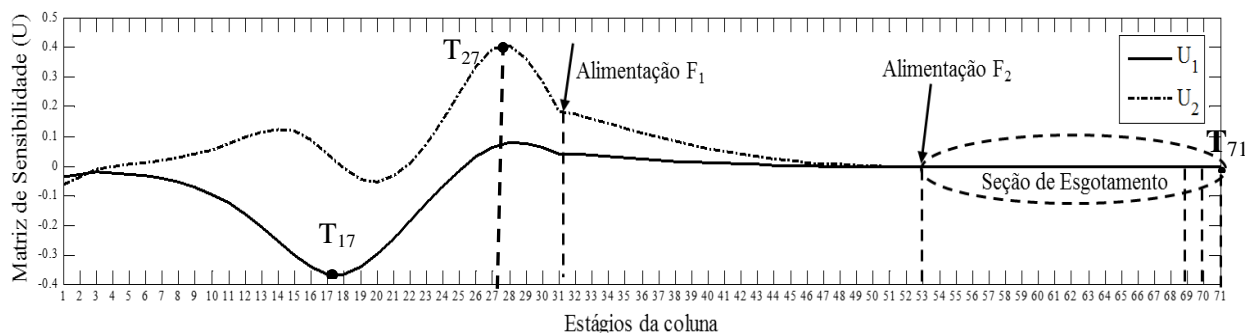
PARES COMBINADOS	NÚMERO CONDICIONAL (CN)
$F_1 - F_2$	10,31390
$F_1 - F_3$	4,2030
$F_1 - R$	222,8583
$F_1 - Q$	26,1591
$F_2 - F_3$	9,9220
$F_2 - Q$	441,3733
$F_2 - R$	115,3110
$F_3 - Q$	35,3539
$F_3 - R$	342,5086
$R - Q$	12565,00

Quanto menor o valor do número condicional significa uma maior interação na malha, enquanto um grande valor para o NC indica que os graus de liberdade são de tal modo que o número de objetivo de controle necessita ser reduzido. Observa-se pela Tabela 5.8, que os valores dos quatro menores NC, em destaque na mesma, estão ligados com pelo menos uma vazão da coluna: vazão de oxicloração, vazão de cloração direta e vazão de retirada de leves, respectivamente, F_1 , F_2 e F_3 . A análise de SVD confirmam na verdade o que foi mostrado, anteriormente, nos resultados das simulações estacionárias e dinâmicas, que F_1 , F_2 e F_3 realmente tem uma interação direta com as controladas temperaturas, ou seja, também com as concentrações de base da torre.

Adotando como a malha de maior interação como sendo aquela que tem o menor NC, ou seja, os resultados são avaliados para o par F_1 - F_3 que apresentou menor valor de NC. A matriz de ganho singular esquerda (U) fornece um sistema de coordenadas mais adequado para visualização da sensibilidade do processo (Luyben, 2006 e 2013).

Com os resultados, obtidos nos *softwares Aspen PlusTM* e da SVD no Matlab[®], são plotados os dados de U pelos estágios, tem-se uma análise onde os estágios que apresentam as maiores magnitudes de U , ou seja, maior diferença do valor 0 de U , indicando os locais da coluna onde possivelmente a medição ou controle seria mais efetivo. O gráfico da Figura 5.18 apresenta a matriz de sensibilidade U , para o caso do par com maior interação na malha (menor NC).

Figura 5.18. Representação gráfica do método SVD, matrizes de sensibilidade U_1 (F_1) e U_2 (F_3).



De acordo com o gráfico da Figura 5.18, que na região de esgotamento (*stripping*), abaixo do estágio de alimentação 53, não ocorrem variações das temperaturas em termos de escala gráfica, o que caracteriza mais uma vez o processo como de alta pureza, haja vista o alto teor do 1,2-DCE nessa região. Verifica-se também na Figura 5.18, que os estágios que apresentam maior variação de temperatura são os estágios 17 e 27. Esses pontos, apesar de maior variabilidade, não apresentam sensores de temperatura na coluna real.

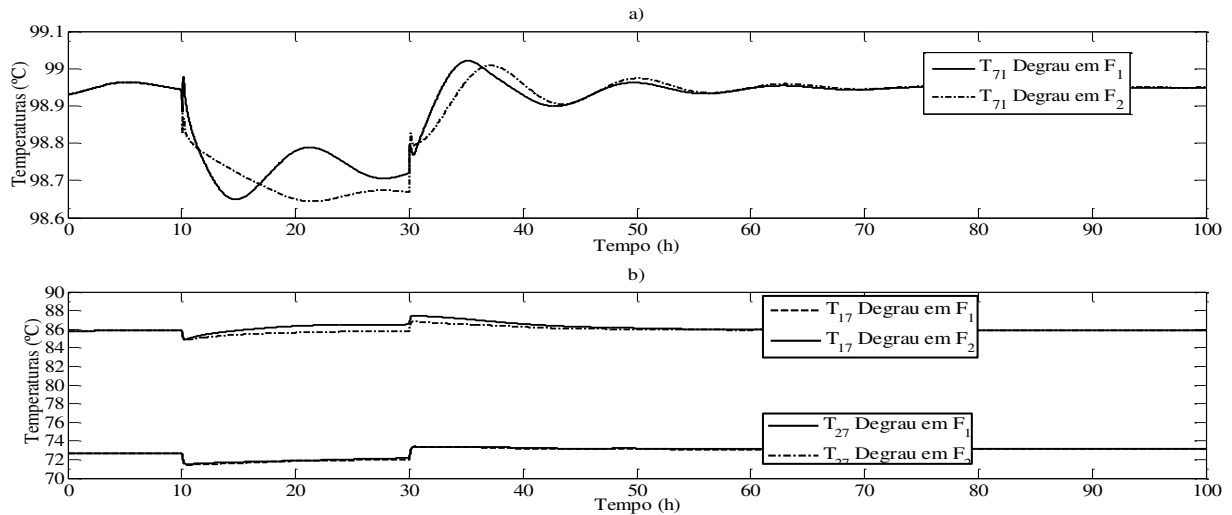
De acordo com as afirmações anteriores, a primeiro momento pode parecer surgir problemas no desenvolvimento dos SV, mas na verdade nos trazem informações importantíssimas à sua construção e controle dessa coluna de colunas de alta pureza, as que seguem:

- ❖ O uso somente das temperaturas da base da coluna para controle de composição torre do 1,2-DCE é impraticável ou inviabilizado, haja vista a pouca variação das T_j . Assim, um controle *single* ou *dual composition* que usa T_j (sessão esgotamento) pode não ser eficiente;
- ❖ Recomenda-se a instalação de sensores de temperaturas nos estágios 17 e 27, para suprir medições que somente são efetuadas na base, topo e vaso de refluxo;
- ❖ O uso somente de temperaturas para estimativa de composições, em processos de alta pureza, através de sensores virtuais pode trazer resultados insatisfatórios. O uso das temperaturas com demais variáveis de entradas, vazões, etc., será aqui avaliado.

Com base nessas informações, para que possamos ter uma visão mais abrangente do comportamento dessas temperaturas, optou-se em verificar o comportamento dinâmico dessas temperaturas frente perturbações nas vazões. As perturbações em F_1 e F_2 foram efetuadas de amplitudes de +20% (no tempo de 10 h) e -20% (no tempo de 30 horas), percentuais em torno dos valores de referência dessas variáveis, respectivamente, $F_{1SP} = 21000$ kg/h e $F_{2SP} = 247500$

kg/h. As Figuras 5.19 a) e b) apresentam, respectivamente, o comportamento de T_{71} e T_{27} e T_{17} diante das modificações degrau no processo.

Figura 5.19. Variação degrau nas vazões efeito sobre as temperaturas: a) T_{71} e b) T_{17} e T_{27} .



Verifica-se de fato que ao modificar os valores das vazões, ocorre uma mínima variação transiente entre as temperaturas. No entanto, mesmo a variação sendo mínimas as temperaturas T_{17} e T_{27} (seção de retificação da coluna) apresenta uma sobre-elevação percentual maior em termos de unidade de escala gráfica do que T_{71} . Essa observação confirma o que a análise SVD indicou: T_{17} e T_{27} são as temperaturas de maior variabilidade. Além disso, percebe-se que a temperatura de base (T_{71}) apresenta variação inferior a $0,5\text{ }^{\circ}\text{C}$.

Em virtude dessas discussões, por uma questão de tornar o modelo de variáveis candidatas mais minucioso, optou-se em utilizar apenas os seguintes valores de temperaturas: as duas mais sensíveis na avaliação da SVD, respectivamente, T_{17} e T_{27} , e a da base coluna (temperatura do estágio 71, T_{71}). A principal justificativa para avaliação dessa última variável é o fato de existir um sensor físico nesse ponto da coluna, onde os valores da mesma podem ser captados de forma direta pelo instrumento.

5.3. CONJUNTO DE DADOS

Os resultados das RNA dependem da quantidade de dados, ou seja, bons treinamentos de uma RNA somente podem ser alcançados quando uma quantidade grande de dados estão disponíveis. Analogamente ao que ocorre com modelos lineares, é importante que os dados contenham informações suficientemente relevantes para permitir a reprodução do processo pelo modelo (BRAGA et al., 2000).

Como verificado, as vazões de alimentação (F_1 e F_2) são distúrbios de alta grandeza para às impurezas da base. Contudo, elas somente são alteradas por uma demanda necessária da planta. Deseja-se que os dados para as estimativas de inferência representem uma amplitude na qual se espera que o sensor virtual venha operar, situações reais de operação do processo. Desta forma, os dados foram gerados com as seguintes considerações:

- ❖ As vazões F_1 , F_2 e F_3 sofreram perturbações degrau no tempo. As variáveis F_1 e F_2 , tiveram seus valores de referência modificados duas vezes ao dia. É importante notar que essas vazões sofreram perturbações degrau em instantes de tempos distintos;
- ❖ Foi inserido um ruído Gaussiano nos sinais de medição transmitido aos controladores de F_1 , F_2 e F_3 ;
- ❖ O sinal utilizado em Q_R e nas concentrações de alimentação foi o sinal PRBS, em todas as variáveis a amplitude máxima do sinal foi de +10%.

Espera-se que esses dados sejam representativos e em quantidades suficientes, para que o modelo de identificação possa gerar um padrão de reconhecimento futuro. É fato que sistemas reais apresentam certo grau de incerteza nos valores dos parâmetros, das medições efetuadas, das entradas esperadas e dos distúrbios medidos e/ou não medidos. Na análise dinâmica é comum desprezar tais incertezas e ruídos, procedendo como se todas as quantidades tivessem valores bem definidos, conhecidos precisamente. Entretanto, os ruídos fazem com que a variável de processo pareça desviar-se do ponto de ajuste, sem qualquer perturbação real está ou não atuando no sistema. Esses efeitos são resultados da tecnologia utilizada para detectar ou medir a variável de processo. Em sinais elétricos, o ruído de medição é muitas vezes devido à interferência de outras fontes elétricas. Na prática, o ruído pode também ser causado pelo desgaste no sensor ou alguma obstrução física que faz com que o sensor emitir uma leitura para imprecisa para o controlador (COOPER, 2004; GARCIA, 2005).

Assim, os dados reais dos processos químicos contêm componentes aleatórios (ruídos), que influenciam no comportamento do sistema e que podem ser a origem dos erros grosseiros das medições (*outliers*). Nesse trabalho, foram simulados ruídos inerentes a interferências de leituras de sensores e de transmissão de dados, os ruídos brancos Gaussianos. Os ruídos foram inseridos nas medições das vazões F_1 , F_2 e F_3 , com um desvio padrão (σ) de 2% dos valores de referência dessas variáveis. Segundo Bega et al.(2011), o valor de σ depende da precisão (*accuracy*) do instrumento especificada pelo fabricante em cada ponto ou toda faixa medida.

Além disso, em modelos de identificação não lineares, como é caso da rede neural artificial, geralmente utiliza-se como sinal de excitação das entradas um sinal binário randômico ou pseudo-randômico, haja vista que sinais binários são fáceis de gerar. O tipo mais comum de sinal binário pseudo-randômico é o PRBS (do inglês *pseudo random binary signal*), que é uma sequência de comprimento máximo ou simplesmente sinais de sequência m .

Para que esse efeito fosse introduzido na análise dinâmica, as variáveis de entrada do estudo (exceto F_1 , F_2 e F_3 , que foram modificadas com sinal degrau) foram perturbadas com o sinal PRBS, cujo sinal aleatório é mantido constante. As perturbações inseridas obedeceram às amplitudes apresentadas na Tabela 5.9.

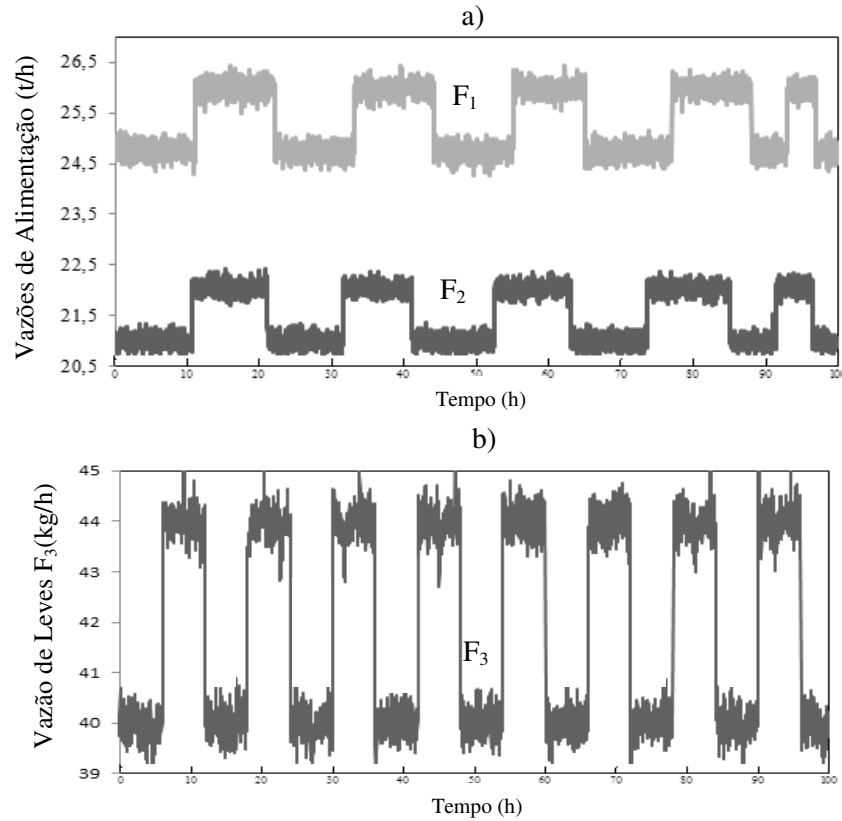
Tabela 5.9. Caracterização das perturbações utilizadas nas simulações para coleta dos dados.

VARIÁVEIS DE ENTRADA	AMPLITUDES DAS PERTURBAÇÕES
F_1	Degrau $\pm 5\%$
F_2	Degrau $\pm 5\%$
F_3	Degrau $\pm 10\%$
Q_R	PRBS $\pm 10\%$
$CCl_4(F_1)$ e $CCl_4(F_2)$	PRBS $\pm 10\%$
$CHCl_3(F_1)$ e $CHCl_3(F_2)$	PRBS $\pm 10\%$

Como verificado na Tabela 5.9, para um melhor critério de avaliação, as simulações foram desenvolvidas através da criação de sequências de perturbações, ou seja, transientes de situações em que ocorrem máximas e mínimas variações das variáveis de entradas, consequente efeito nas saídas. A amplitude dos distúrbios inseridos baseou-se na análise dos resultados, ou seja, análise transiente e oscilações típicas da unidade. A metodologia de coleta de dados foi efetuada durante 100 horas de operação, gerando 5001 pontos para cada uma das saídas, $XCCl_4$ e $XCHl_3$. Uma informação importante é sobre o tempo de amostragem, definido aqui por 1,2 minutos. Esse valor foi calculado como sendo um décimo da menor constante de tempo do sistema.

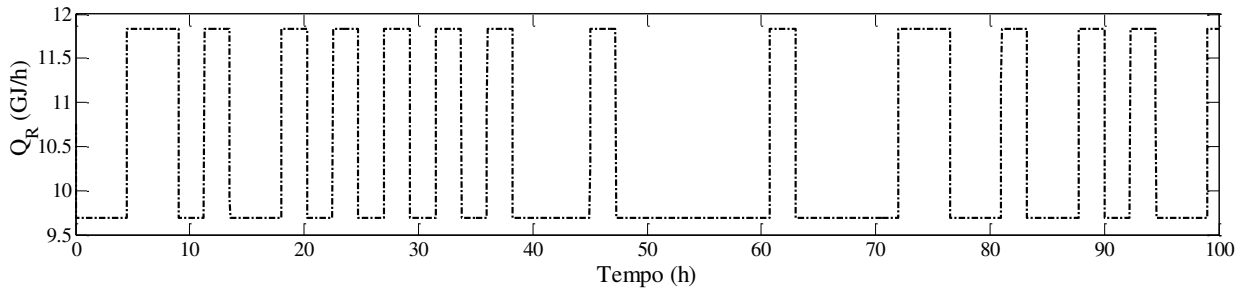
As perturbações e ruídos inseridos nas vazões F_1 , F_2 e F_3 do processo, procedimento definido anteriormente, são apresentados, respectivamente, conforme apresenta as Figuras 5.20 a) e b). De acordo com os gráficos das Figuras 5.20 a) e b), os ruídos inseridos nas medições refletiram nos valores das variáveis controladas (vazões). As perturbações nas variáveis F_1 , F_2 e F_3 foram inseridas, respectivamente, a cada 10,5, 11 e 6 horas.

Figura 5.20. Procedimento para coleta dos dados a) Dinâmica de F_1 ; e b) Dinâmica de F_2 .



O gráfico da Figura 5.21 apresenta o comportamento do sinal PRBS implementado na carga térmica do refeedor da coluna, para gerar os dados de entrada e saída dos SV.

Figura 5.21. Variações PRBS na variável da carga térmica Q_R .



As perturbações nas concentrações dos compostos das alimentações (F_1 e F_2), não foram alteradas de forma simultânea. Ou seja, ao avaliar o comportamento da saída X_{CCl_4} , apresentado nos gráficos das Figuras 5.22 a) e b), foram alteradas apenas as concentrações do CCl_4 em F_1 e F_2 . Do mesmo modo, para a saída X_{CHCl_3} , cujos comportamentos são apresentados nos gráficos das Figuras 5.23 a) e b). Dessa forma, foram efetuados dois experimentos distintos para coleta dos dados.

Figura 5.22. a) PRBS na concentração do CCl_4 na alimentação F_1 ; e b) PRBS na concentração do CCl_4 na alimentação F_2 .

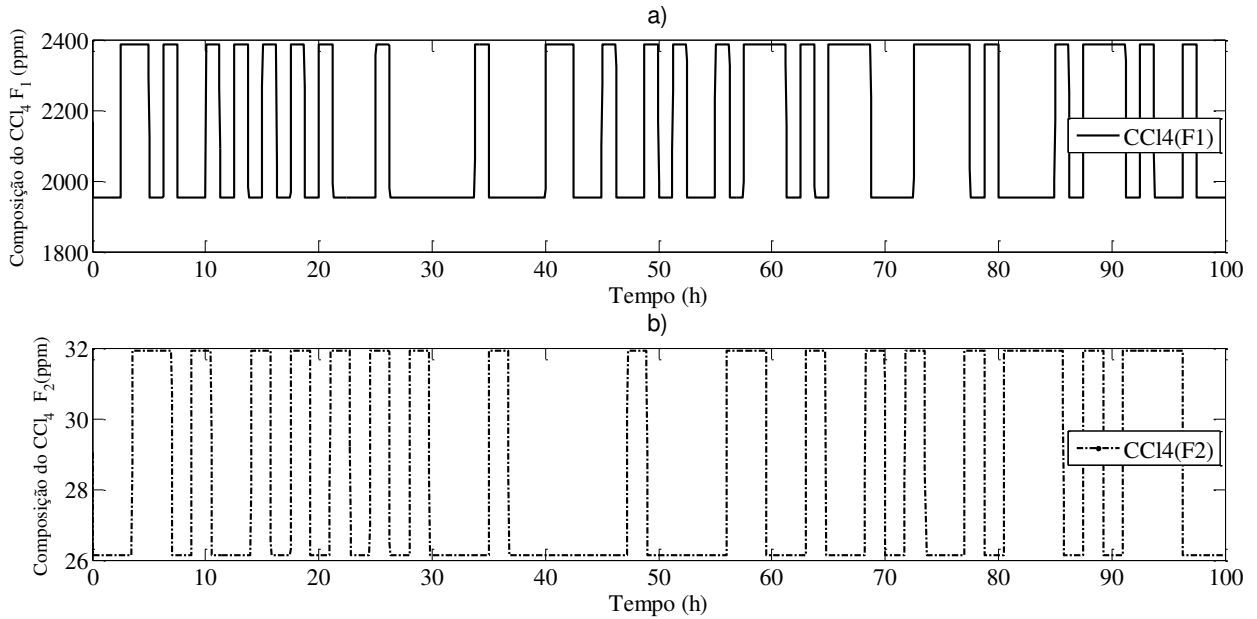
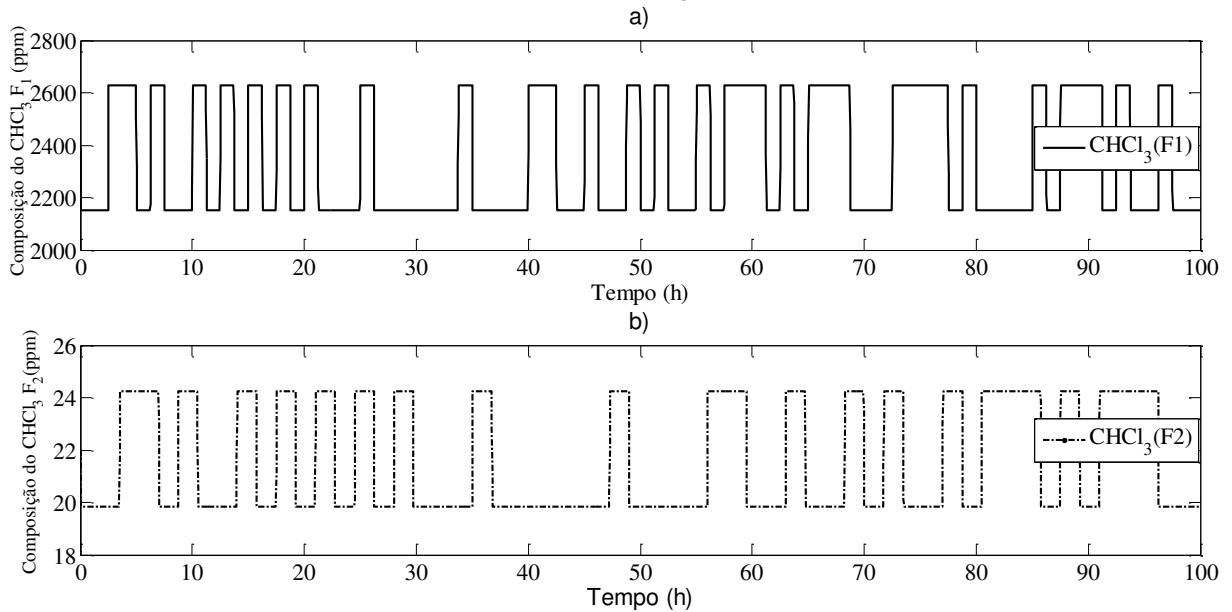


Figura 5.23. PRBS na concentração do CHCl_3 na alimentação F_1 ; e b) PRBS na concentração do CHCl_3 na alimentação F_2 .

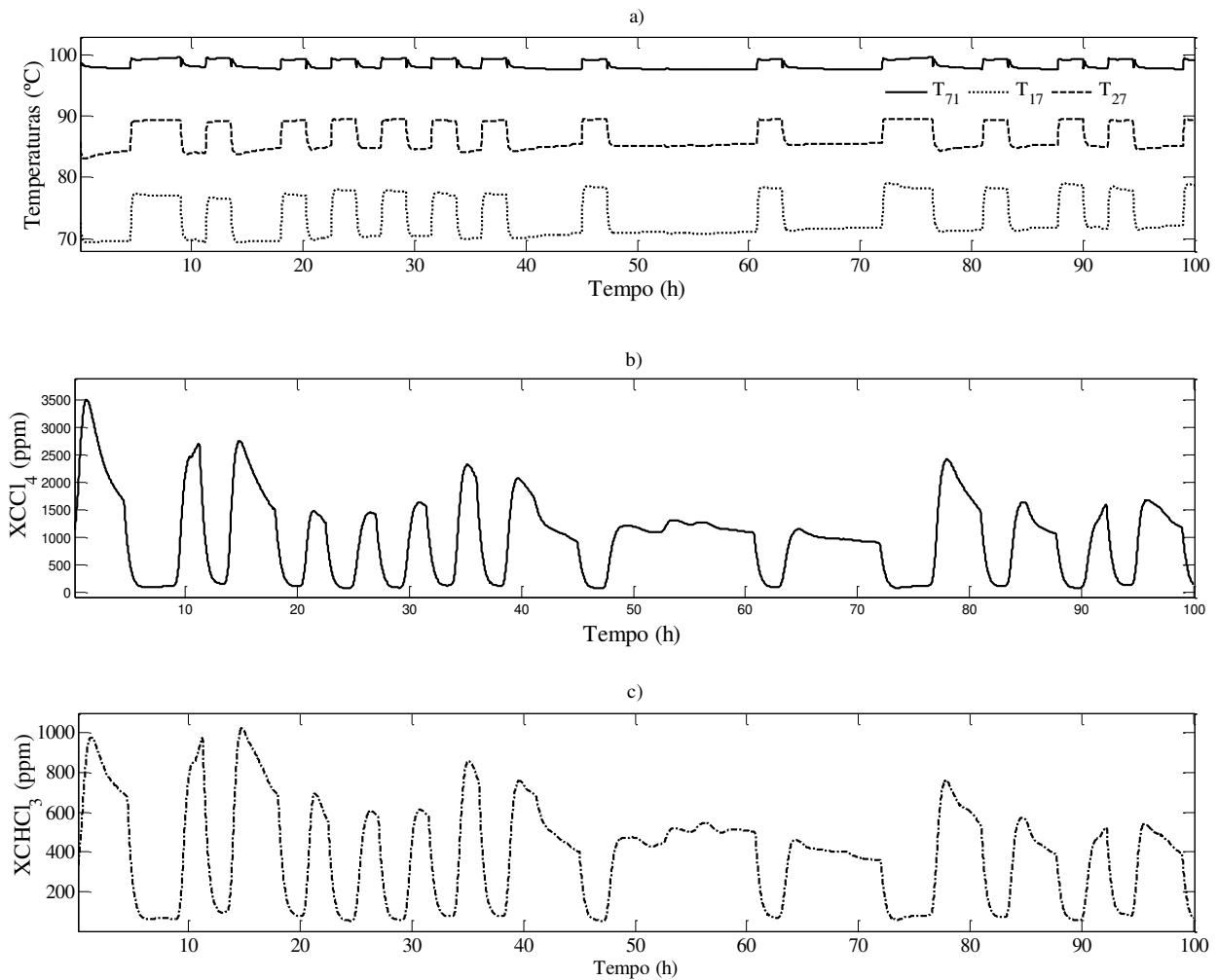


É importante observar que tanto nas Figuras 5.22 a) e b) quanto nas Figuras 5.23 a) e b), as concentrações tiveram seus valores reestabelecidos após 97 horas de operação. Esse procedimento teve como objetivo interromper as perturbações e retornar as condições iniciais de operação da coluna.

As perturbações e os ruídos inseridos nas variáveis: F_1 , F_2 , F_3 , Q_R , concentrações de alimentação do CCl_4 (F_1 e F_2) e do CHCl_3 (F_1 e F_2), produziram efeitos diretos nas variáveis de

saída do processos. O gráfico da Figura 5.24 a) apresenta os comportamentos das temperaturas dos estágios 17, 27 e 71. Enquanto que os gráficos da Figura 5.24 b) e c) apresentam, respectivamente, o comportamento transiente das composições de saídas.

Figura 5.24. Dinâmica das saídas frente as perturbações e ruídos: a) Temperaturas selecionadas no estudo; b) X_{CCl_4} e c) X_{CHCl_3} .



Como já mencionado anteriormente e de acordo com a Figura 5.24 a), nota-se que a T_{71} é menos sensível as perturbações inseridas no sistema, em comparação com outras duas temperaturas da seção de retificação da coluna, T_{17} e T_{27} . O comportamento das concentrações X_{CCl_4} e X_{CHCl_3} (Figuras 5.24 b) e c)) são semelhantes em termos gráfico, entretanto elas apresentam valores na escala gráfica distintos. Nos dois últimos gráficos, os valores gerados são os que devem ser comparados com as estimativas dos sensores virtuais.

Podem ser armazenados quatro conjuntos de dados, sendo os mesmos salvos no excel em arquivos *.xls*, gerados a partir da planta do *Aspen Dynamics*TM e com a plataforma de

comunicação OLE (*Object Linking and Embedding*), via conexão do VBA (*Visual Basic Application*). Esses arquivos são lidos por *m-files* (extensão *.m*) do Matlab®. Os dados que podem utilizados são descritos abaixo:

- ❖ Dados 1: valores da saída (X_{CCl_4}) e das entradas, incluindo as amostras das concentrações do CCl_4 (F_1 e F_2) e sem os valores das amostras das temperaturas;
- ❖ Dados 2: valores da saída (X_{CCl_4}) e das entradas, incluindo as amostras das temperaturas e sem os valores das concentrações do CCl_4 (F_1 e F_2);
- ❖ Dados 3: valores da saída (X_{CHCl_3}) e das entradas, incluindo as amostras das concentrações do $CHCl_3$ (F_1 e F_2) e sem os valores das temperaturas;
- ❖ Dados 4: valores da saída (X_{CHCl_3}) e das entradas, incluindo as amostras das temperaturas e sem os valores das concentrações do $CHCl_3$ (F_1 e F_2).

Por fim, os dados foram armazenados visando ter uma dimensão específica (5001×7), de modo que os modelos contendo as temperaturas e concentrações de alimentação sejam avaliados na etapa de seleção de variáveis. Assim, para cada saída foram geradas duas matrizes de dados, cujo conjunto final possui um total 70014 dados.

5.4. DADOS PRÉ-PROCESSADOS

Após a metodologia de coleta de dados da planta, o próximo passo consiste em verificar a qualidade dos dados obtidos. Assim, a inferência deve ser capaz de representar o processo, no qual ela está associada em todo o espaço operacional no qual a unidade do processo trabalha. O objetivo aqui é gerar outros conjuntos de dados alternativos aos “Dados 1, 2, 3, e 4”, novos conjuntos de dados onde os *outliers* são removidos.

Dados oriundos de sensores físicos como, por exemplo, sensores de temperatura, pressão, nível, vazão, etc., e analisadores em linha, são passíveis de fatores dinâmicos não lineares que afetam diretamente a medição/inferência, são eles: variabilidade, ruído, repetibilidade, histerese, folga (*backslach*), agarramento de válvulas, saturação, etc. Mas, a possibilidade de incertezas não é exclusiva de dados obtidos dos sensores físicos.

No caso de dados oriundos de modelos fenomenológicos, como é o caso desse trabalho, este também está condicionado à limitação de fatores ou de hipóteses simplificadoras efetuadas durante o processo de modelagem ou até a imprecisão de dados obtidos na planta, para validação estacionária e dinâmica (FORTUNA et al., 2007).

Assim, dados coletados em uma planta ou oriundos de um modelo matemático podem estar corrompidos pela presença de erros grosseiros “*outliers*”, isto é, observações não consistentes com a maioria dos dados e incluem pontos inexistentes, falhas de leituras, problemas de transmissão, etc. Com base na metodologia da seção 4.3, os possíveis *outliers* foram avaliados aqui de duas maneiras distintas: análise da PCA por meio de avaliação dos *loadings* (Autovetores) ou gráfico dos “*scores*”, abordagem citada por Jolliffe (2002), e a análise residual da MLR, através dos erros residuais padronizados (d_i).

5.4.1. Remoção de *Outliers* por meio da PCA

Os autovalores (*scores*) são as projeções das amostras na direção das componentes principais, suas coordenadas são obtidas pelo produto da matriz de dados e a matriz de autovetores. Os conjuntos de dados 1, 2, 3 e 4, armazenados inicialmente, são submetidos às análises dos autovalores à remoção de *outliers*. Nas Tabelas 5.10 a), b), c) e d) são apresentados os valores das variâncias (autovalores) e das variâncias explicativas e acumuladas das PC_i .

De acordo com Jolliffe (2002) as principais PC, aquelas com variância maior que 0,7 correspondem, respectivamente: cinco primeiras PC das Tabelas 5.10 a) e c), e as três primeiras PC das Tabelas 5.10 b) e d).

Tabela 5.10. Variância explicativa e acumulada das PC do conjunto de: a) Dados 1; b) Dados 2; c) Dados 3; e d) Dados 4.

a)

Componentes Principais	Variância (Autovalores)	Variância Explicativa (%)	Variância Acumulada (%)
PC1	1,9761	28,23	28,23
PC2	1,6497	23,57	51,80
PC3	1,1527	16,47	68,26
PC4	0,9233	13,19	81,46
PC5	0,8704	12,43	93,89
PC6	0,3661	5,23	99,12
PC7	0,0617	0,88	100,00

b)

Componentes Principais	Variância (Autovalores)	Variância Explicativa (%)	Variância Acumulada (%)
PC1	3,8408	54,87	54,87
PC2	1,6366	23,38	78,25
PC3	0,9750	13,93	92,18
PC4	0,3733	5,33	97,51
PC5	0,1180	1,69	99,20
PC6	0,0439	0,63	99,82
PC7	0,0123	0,18	100,00

c)

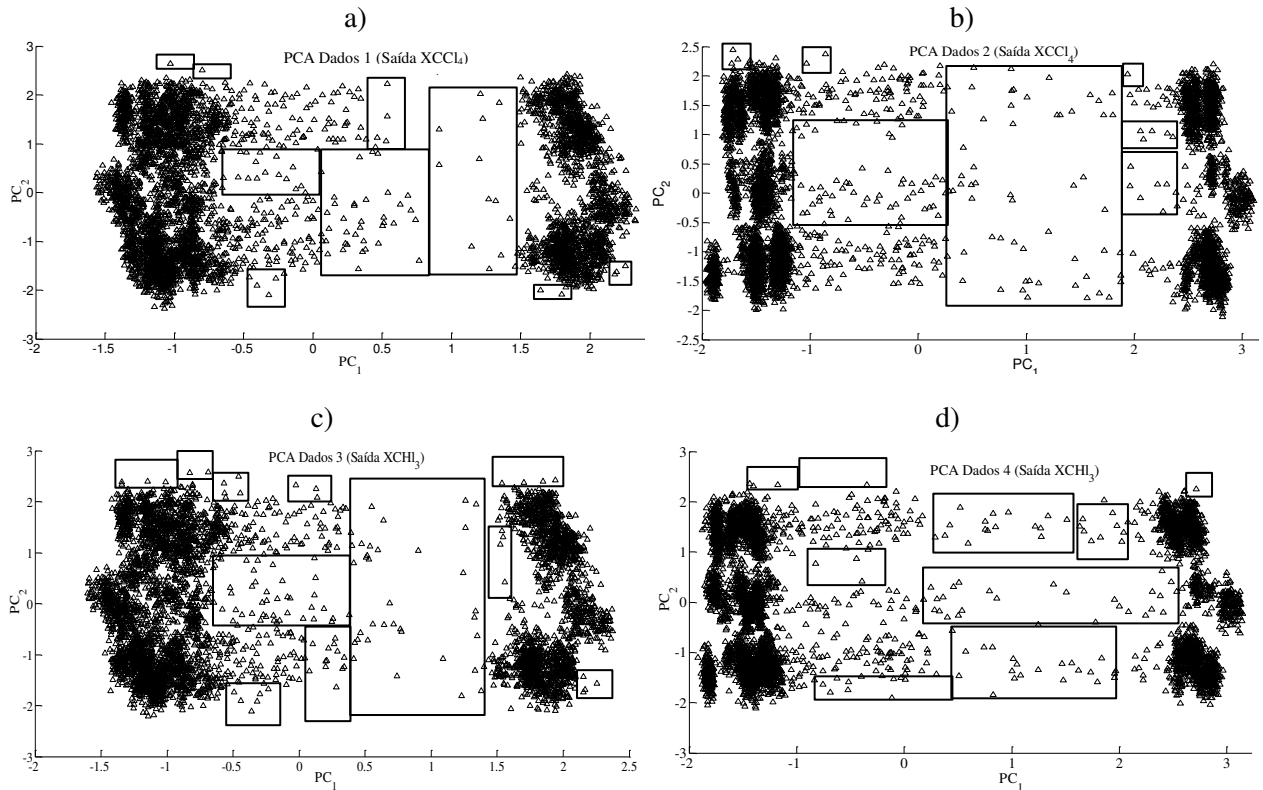
Componentes Principais	Variância (Autovalores)	Variância Explicativa (%)	Variância Acumulada (%)
PC1	1,9896	28,42	28,42
PC2	1,6006	22,87	51,29
PC3	1,1518	16,45	67,74
PC4	0,9237	13,20	80,94
PC5	0,8798	12,57	93,51
PC6	0,3925	5,61	99,12
PC7	0,0619	0,88	100,00

d)

Componentes Principais	Variância (Autovalores)	Variância Explicativa (%)	Variância Acumulada (%)
PC1	3,8524	55,03	55,03
PC2	1,5777	22,54	77,57
PC3	0,9797	14,00	91,57
PC4	0,4163	5,95	97,51
PC5	0,1171	1,67	99,19
PC6	0,0442	0,63	99,82
PC7	0,0127	0,18	100,00

Logo, a partir da Tabela 5.10, as sete variáveis originais dos conjuntos de dados 1, 2, 3 e 4 podem ser analisadas a partir das PC selecionadas, as quais expressam, respectivamente, 93,89%, 93,51%, 92,18% e 91,57% da variância preservada dos dados originais (reprodução da variabilidade). O uso do gráfico dos autovalores é uma alternativa viável na identificação de *outliers*. Analisam-se apenas as primeiras componentes principais (PC₁ e PC₂) de cada conjunto de dados, porque estas retêm individualmente maior variabilidade explicativa dos dados. Assim, os dados questionáveis sobre possíveis *outliers* são identificados por retângulos nas Figuras 5.25 a), b), c) e d), que correspondem, respectivamente, as análises da PCA dos conjuntos de dados 1, 3, 2 e 4.

Figura 5.25. Análise PCA: a) Conjuntos de Dados 1 (Saída $XCCl_4$); b) Conjuntos de Dados 2 (Saída $XCCl_4$); c) Conjuntos de Dados 3 (Saída $XCHCl_3$); e e) Conjuntos de Dados 4 (Saída $XCHCl_3$);



Conforme apresentam os gráficos das Figuras 5.25 a), b), c) e d), o total de *outliers* é identificado por retângulos. As matrizes originais de dados foram analisadas, sendo excluídos os dados que se encontram nessa faixa de observabilidade, formando um novo conjunto de dados com as mesmas 7 (sete) variáveis originais e contendo um menor número de amostras. Assim, através da seleção das principais PC, os conjuntos de dados 1 e 3, com sete variáveis sintéticas (PC_i) e 5001 amostras, foram reduzidos a cinco variáveis sintéticas. De maneira análoga, os conjuntos de dados 2 e 4, com sete variáveis sintéticas (PC_i) e 5001 amostras, foram reduzidos a três variáveis sintéticas.

A identificação das amostras em observância às PC_i envolve um alto grau de subjetividade, haja vista que os dados considerados como *outliers* apresentam-se relativamente distantes da massa de dados, o que pode indicar ou não a presença de amostras influentes nos conjuntos de dados originais. Por esse motivo, nesse trabalho optou-se por adotar uma metodologia complementar à PCA visando à remoção de *outliers*, o critério da análise residual (erros padronizados, d_i). Desse modo, ter uma maior confiabilidade na remoção de erros grosseiros dos dados amostrais.

5.4.2. Remoção de *Outliers* por meio de Análise Residual

Como descrito na seção 4.2.3, a análise residual de erros consiste em um procedimento para encontrar observações excepcionalmente influentes. Essas observações influentes estão relativamente longe da vizinhança, onde o resto dos dados foi coletado.

O procedimento aqui implementando consiste em implementar os conjuntos de dados originais em um modelo de regressão múltipla (MLR) e avaliar o resíduo (erro), $e_i = y_i - y_{modelo}$. Entretanto, como já citado, em muitas situações é mais conveniente utilizar os erros residuais padronizados (d_i), Equação 4.24, pois esses são escalonados de modo que seus desvios-padrão sejam aproximadamente iguais a um (MONTGOMERY e RUNGE, 2014).

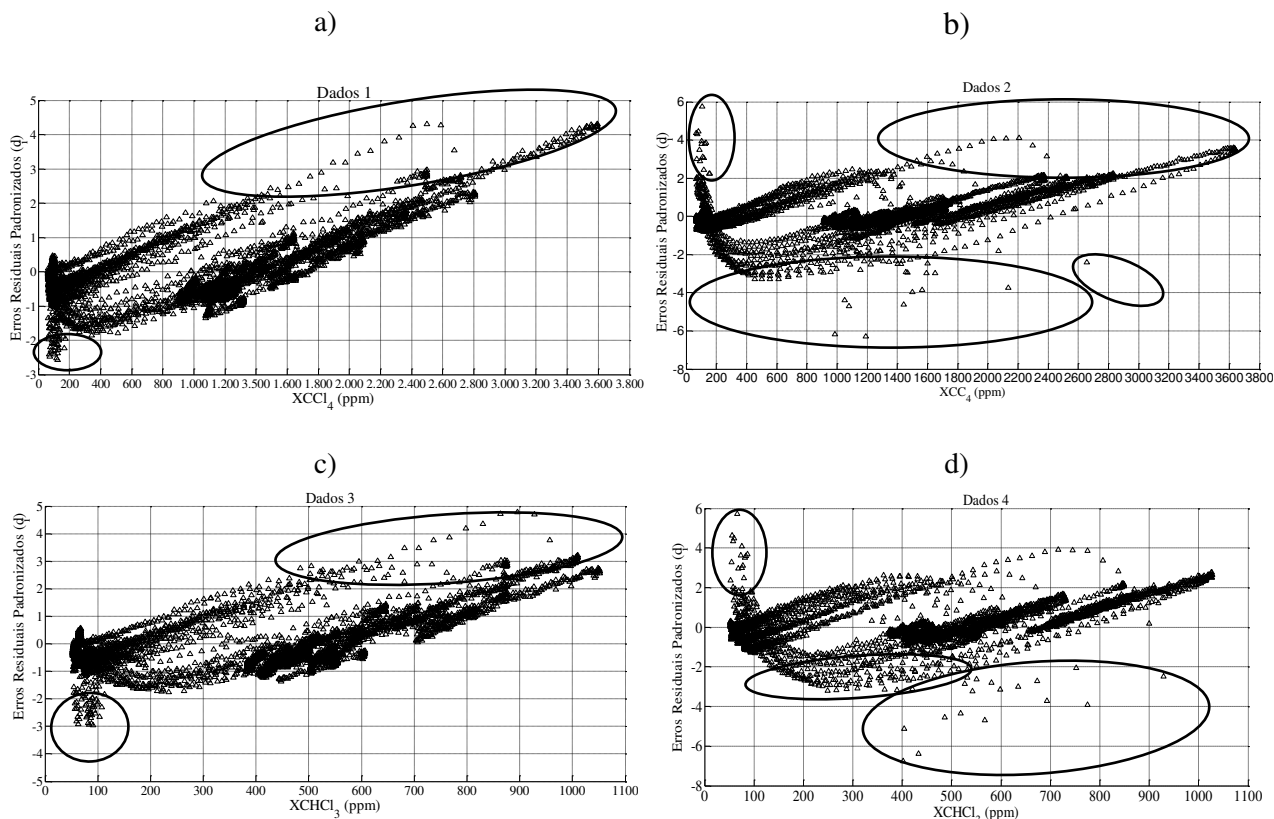
Consequentemente, d_i grandes podem indicar possíveis *outliers*, facilmente interpretados através da inspeção dos gráficos residuais. Sendo assim, são plotados os valores de d_i versus os valores de $XCCl_4$ e $XCHCl_3$, para os dados 1, 2, 3 e 4 obtidos nos modelos de regressão, conforme apresentam os gráficos das Figuras 5.26 a), b), c) e d).

Conforme é verificado nos gráficos das Figuras 5.26 a), b), c) e d), para todos conjuntos de dados, a maioria dos resíduos padronizados encontram-se dentro do intervalo de ± 2 desvios padrões, indicando a normalidade dos resíduos e, portanto, valores contidos fora desse intervalos podem ser considerados erros grosseiros. É importante citar, que a avaliação com base nessa metodologia também foi efetuada para as entradas, visando encontrar os possíveis pontos *outliers*.

Do mesmo modo que as observações através da PCA, as análises residuais padronizadas são meramente subjetivas, haja vista que um ponto que apresenta um erro modular alto não necessariamente é pode ser considerado um *outlier*. Sendo assim, durante três simulações consecutivas, efetuou-se observações nos conjuntos de dados iniciais gerados (Dados 1, 2, 3 e 4), com base nas duas técnicas aqui mencionadas.

Verificou-se que em alguns casos não ocorriam dados fora dos padrões e em outros casos, os dados apresentavam razoável variabilidade. Neste caso, tais pontos foram removidos, dando origem a quatro novos conjuntos de dados pré-processados: “Dados 5, 6, 7 e 8”; que apresentam uma remoção cerca de 7% de *outliers*.

Figura 5.26. Gráfico dos erros padronizados residuais (d_i) versus saídas: a) Dados 1; b) Dados 2; c) Dados 3; e d) Dados 4.



Assim, a identificação de *outliers* é uma abordagem importante na modelagem empírica de conjunto de dados de alta dimensão. Logo, é possível encontrar informações que podem levar às incertezas no desempenho dos sensores virtuais. Portanto, as duas técnicas aqui empregadas contribuíram de forma extremamente construtivas (criteriosas) para remoção de erros grosseiros em dados dinâmicos, obtidos durante a simulação ordenada do processo.

5.5. ETAPA DE SELEÇÃO DE VARIÁVEIS

É fato que o aumento no número de variáveis de entradas resulta em melhoria do ajuste dos dados de treino usados na construção de um modelo de inferência. No entanto, a inclusão de muitas variáveis frequentemente causa uma redução substancial na habilidade preditiva dos sensores virtuais em relação aos dados de teste (validação). De fato, é necessário selecionar o conjunto de variáveis que produza o melhor modelo preditivo. No caso da coluna de destilação do 1,2-DCE, objeto de estudo, o modelo construído permite avaliar as variáveis secundárias mediante as duas saídas: composição líquida do tetracloreto de carbono ($XCCl_4$) e do clorofórmio ($XCHCl_3$).

5.5.1. Seleção de Variáveis Utilizando a Técnica de TRP

Como discutido anteriormente, o número de equações a ser examinado cresce rapidamente à medida que o número de variáveis candidatas aumenta. No entanto, a proposta do algoritmo computacional da seção 4.2.3.5, implementado no *software* Matlab[®], torna o objetivo eficiente para todas as regressões possíveis (TRP).

O procedimento se torna prático à medida que o conjunto de variáveis candidatas e o de amostra estão bem definidos. Desta forma, na etapa 8 do algoritmo da seção 4.2.3.5, implementou-se o algoritmo da TRP, haja vista que com esse procedimento é possível verificar todas as combinações possíveis das variáveis candidatas, onde cada um dos modelos e seus respectivos regressores são avaliados por meio de um critério de desempenho. Além disso, segundo Montgomery e Runger (2014), se o número de regressores candidatos não for muito grande, como é caso aqui nesse estudo, o método da TRP é o preferencial. Recomenda-se, geralmente, usar critérios de avaliação para minimização do erro médio quadrático (MQE), em conjunto com esse procedimento.

Essa abordagem possa encontrar o(s) “melhor(es)” modelo(s) para inferência relativamente ao critério mínimo MQE (ou outros critérios), enquanto os métodos tipo etapas (Adição sequencial, Seleção por Eliminação e da Regressão por Etapas) não oferecem essa segurança. Além disso, eles podem ser distorcidos pelas dependências entre as variáveis candidatas. Nessa metodologia, as entradas, u_i , foram renomeadas para facilitar o estudo de seleção de variáveis de acordo com a Tabela 5.11.

Tabela 5.11. Variáveis candidatas a serem entradas dos sensores virtuais.

ENTRADAS (VARIÁVEIS CANDIDATAS)										
u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}
F ₁	F ₂	F ₃	Q _R	CCl ₄ (F ₁)	CCl ₄ (F ₂)	CHCl ₃ (F ₁)	CHCl ₃ (F ₂)	T ₁₇	T ₂₇	T ₇₁

Com já discutido anteriormente, a maioria dos trabalhos sobre sensores virtuais incluem nas entradas a concentração do composto a ser inferido na(s) corrente(s) de alimentação. Nesse trabalho é proposta uma avaliação quanto à significância dessas variáveis junto aos modelos dos SV construídos. Nessa metodologia, a Tabela 5.12 apresenta as variáveis candidatas avaliadas nos quatro conjuntos, definidos anteriormente, cada um contém sete variáveis e uma saída.

Tabela 5.12. Combinação das variáveis candidatas nos modelos de inferência.

SAÍDA INFERIDA	CONJUNTO DE DADOS	ENTRADAS (VARIÁVEIS CANDIDATAS)						
$y_1(\text{XCCL}_4)$	1	u_1	u_2	u_3	u_4	u_5	u_6	u_{11}
	2	u_1	u_2	u_3	u_4	u_9	u_{10}	u_{11}
$y_2(\text{XCHCl}_3)$	3	u_1	u_2	u_3	u_4	u_7	u_8	u_{11}
	4	u_1	u_2	u_3	u_4	u_9	u_{10}	u_{11}

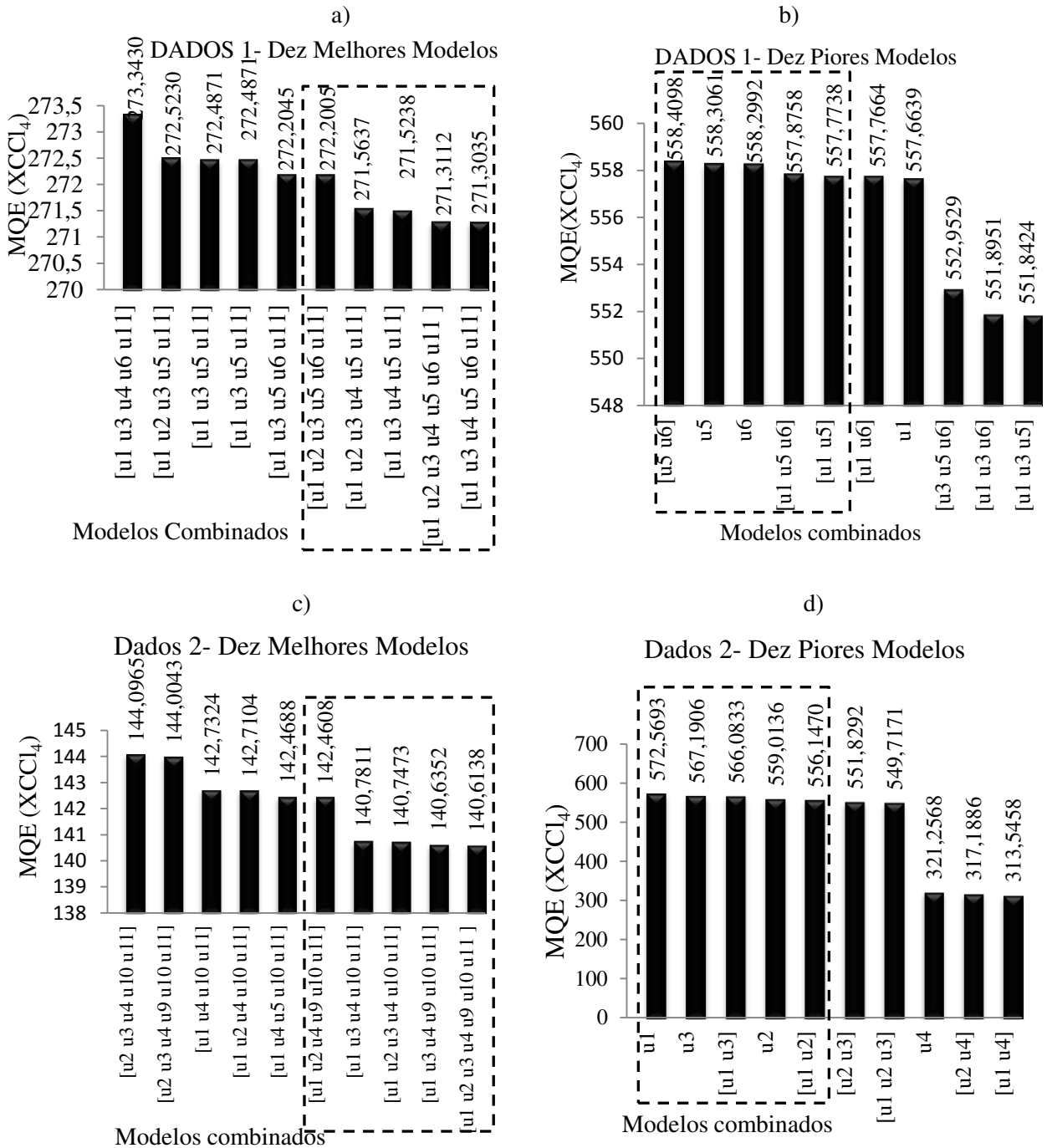
De acordo com a Tabela 5.15, todos os conjuntos de dados possuem as variáveis u_1 , u_2 , u_3 e u_4 . Nos conjuntos de dados 2 e 4, as concentrações de alimentação foram substituídas pelas temperaturas T_{17} e T_{27} . Assim, o procedimento visa avaliar se a substituição das concentrações de alimentação pelas temperaturas traria melhora significativa aos modelos gerados.

Sendo assim, tem-se um total de 7 (sete) regressores candidatos para os dados 1, 2, 3 e 4. O número total de combinações entre as variáveis de cada conjunto de dados é: $NC = (2^7 - 1) = 127$. Logo, será necessário avaliar os 127 modelos em cada conjunto de dados, equações em forma de regressão linear múltipla, calcular seus coeficientes, avaliar os modelos gerados e encontrar o(s) melhor(es) modelo(s) para serem treinados como entradas da rede neural artificial.

O algoritmo de todas as combinações possíveis para saídas XCCL_4 e XCHCl_3 , frente aos 508 modelos gerados (127 para cada conjunto de dados), bem como o programa de seleção de variáveis, foram implementados no *software* Matlab[®].

O número de variáveis inseridas nos modelos varia de 1 a 8 (contando com o vetor coluna identidade u_0). Visando tornar mais clara as interpretações da etapa de seleção de variáveis, foram gerados gráficos histogramas interpretativos. Assim, nas Figuras 5.27 a), b), c) e d) são apresentados, respectivamente, os dez melhores e dez piores modelos para os conjuntos de dados 1 e 2 (critério do valor do MQE), dados esses que estão relacionados com a saída XCCL_4 .

Figura 5.27. Inferência de $XCCl_4$: a) Dez melhores modelos Dados 1; b) Dez piores modelos Dados 1; c) Dez melhores modelos Dados 2; e d) Dez piores modelos Dados 2.



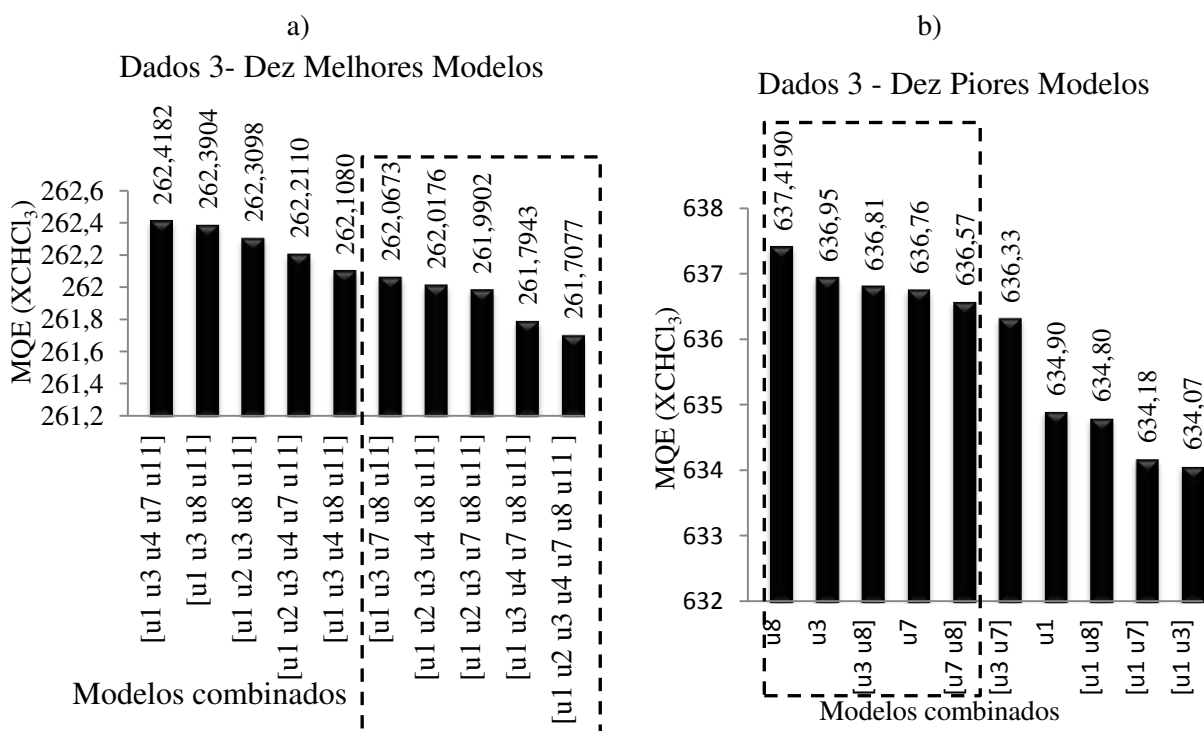
Em uma análise dos gráficos das Figuras 5.27 a) e c) nos responde quanto ao questionamento da habilidade preditiva dos modelos finais para a saída $XCCl_4$. É notório verificar, que os melhores modelos gerados para o conjunto de dados que possuem as concentrações de alimentação (u_5 e u_6), dados 1, tiveram desempenho inferior em comparação com os modelos que contêm as temperaturas selecionadas no estudo (u_9 e u_{10}), dados 2. Verifica-

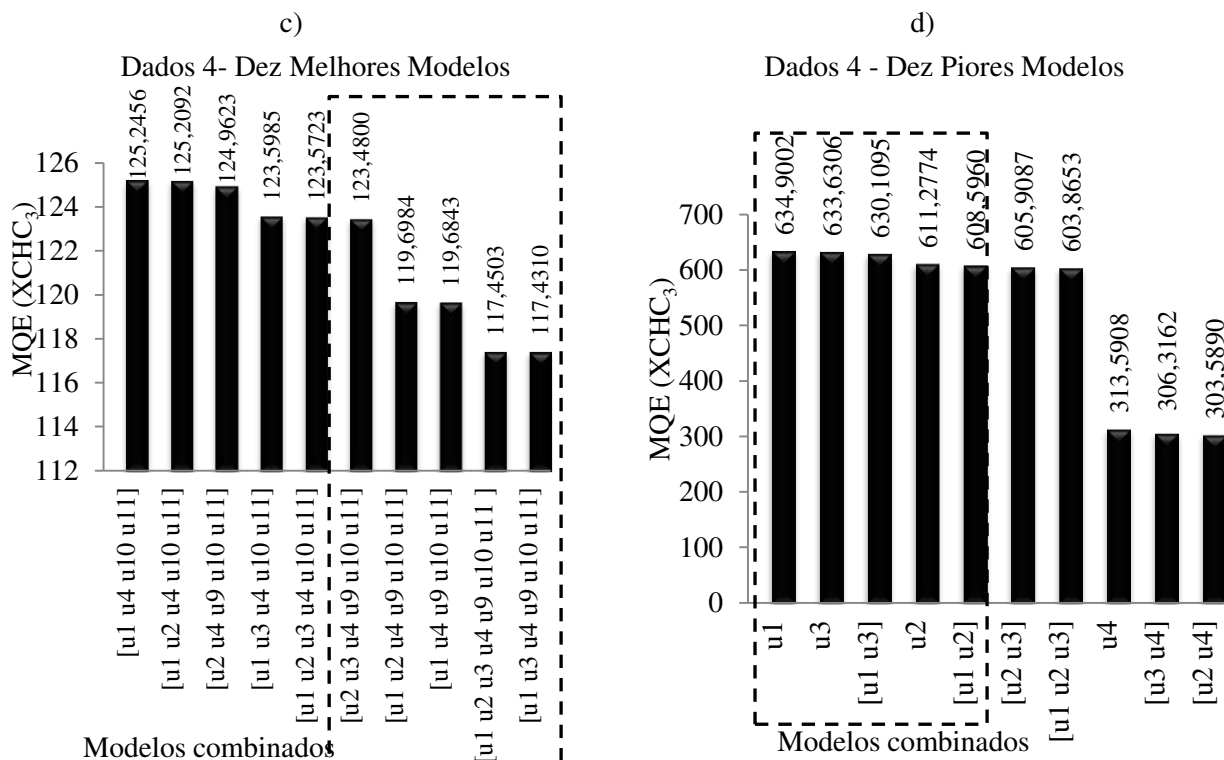
se também, que os cinco melhores modelos desses conjuntos apresentam valores de MQE bem próximos, o que pode indicar que esses levam a bons resultados na inferência da saída $XCCl_4$.

Diante do exposto, nesse estudo optou-se em utilizar os cinco melhores modelos dos conjuntos de dados 1 e 2, destacados por retângulos nas Figuras 5.27 a) e c). Por fim, esses modelos serão avaliados à construção do sensor virtual do composto CCl_4 . Ainda sobre a saída $XCCl_4$, os modelos aqui apresentados como piores, Figura 5.27 b) e d), possuem valores de MQE muito acima dos considerados como melhores, portanto são descartados do estudo.

Nas Figuras 5.28 a), b), c) e d) são apresentados, respectivamente, os dez melhores e os dez piores modelos para os conjuntos de dados 3 e 4 (critério do valor do MQE), dados esses que estão relacionados com a saída $XCHCl_3$.

Figura 5.28. Inferência de $XCHCl_3$: a) Dez melhores modelos Dados 3; b) Dez piores modelos Dados 3; c) Dez melhores modelos Dados 4; e d) Dez piores modelos Dados 4.





De forma análoga, as mesmas conclusões sobre os melhores e piores modelos, citadas anteriormente para a saída $XCCl_4$, são válidas para a saída $XCHCl_3$. Sendo assim, foram selecionados os cinco melhores modelos à $XCHCl_3$ dos conjuntos de dados 3 e 4, em destaque nas Figuras 5.28 a) e c).

Em todos os conjuntos de dados, os piores modelos são aqueles que contêm no mínimo uma e no máximo três variáveis candidatas. Sobre os modelos que possuem melhor capacidade preditiva, esses têm no mínimo cinco ou todas as variáveis avaliadas. Nesse contexto, a técnica aponta que os modelos que contêm duas ou as três temperaturas (u9, u10 e u11) são os mais viáveis para o treinamento das RNA, tal fato é verificado nas Figuras 5.27 c) e 5.28 c).

É importante citar que os modelos foram verificados como estatisticamente significativos, ou seja, pode-se verificar o grau de correlação entre as entradas e saídas. No entanto, a título de processos de destilação, e conhecidas as correlações não lineares inerentes ao processo, não é objetivo deste trabalho obter a estimativa das variáveis de saída por meio de um modelo linear e sim, apenas foi verificado estatisticamente a significância de cada uma das variáveis de entradas, com as saídas $XCCl_4$ e $XCHCl_3$.

5.5.2. Seleção de Variáveis Utilizando a PCA

Com já citado anteriormente, o uso da análise de componentes principais para fins de selecionar variáveis é uma rota alternativa na metodologia de sensores virtuais. Por outro lado, a técnica pode confirmar realmente se todas as variáveis contida(s) no(s) melhor(es) modelo(s), selecionado(s) no algoritmo de regressão múltipla (TRP), devem permanecer ou serem removidas do conjunto original. Para facilitar o entendimento e cálculos, são utilizados os mesmos modelos analisados na remoção de *outliers* (seção 5.4.1), indicados na Tabela 5.15 e contidos entre os melhores modelos das Figuras 5.27 a) e 5.28 c), referentes, respectivamente, às saídas $XCCl_4$ e $XCHCl_3$.

Em uma análise fatorial, considerando-se 7 (sete) variáveis dos modelos em questão, têm-se 7 (sete) fatores que corresponderiam às variáveis originais. A escolha do número de fatores pode levar em conta diferentes critérios. Um deles está em incluir, na análise, aqueles componentes principais (PC_i) que conseguem sintetizar uma variância acumulada em torno de 70%. Nessa metodologia, as sete variáveis originais dos conjuntos de dados 1, 2, 3 e 4 são analisadas a partir das PC selecionadas, as quais expressam, respectivamente, 93,89%, 93,51%, 92,18% e 91,57% da variância dos dados originais. Sendo assim, a seleção de variáveis dos conjuntos de dados 1 e 3 é efetuada através das cinco primeiras PC_i . De maneira análoga, a seleção de variáveis dos conjuntos de dados 2 e 4 é efetuada através das três primeiras PC_i .

Através das PC_i selecionadas são extraídas as principais variáveis originais pelo valor absoluto dos autovetores (*loadings*). As Tabelas 5.16 a) e b) apresentam os autovetores das variáveis originais referentes aos dados 1 e 2, respectivamente, para saídas $XCCl_4$ e $XCHCl_3$. Enquanto as Tabelas 5.16 c) e d), apresentam os autovetores das variáveis originais referentes aos dados 3 e 4, saída $XCHI_3$.

Tabela 5.16. a) Autovetores (*Loadings*) das PC selecionados com relação à saída $XCCl_4$; b) Autovetores (*Loadings*) das PC selecionados com relação à saída $XCHCl_3$.

a)

Componentes	F ₁	F ₂	F ₃	Q _R	CCl ₄ (F ₁)	CCl ₄ (F ₂)	T ₇₁
PC ₁	0,08759	0,66676	0,21704	-0,1597	0,0199685	-0,687689	-0,042999
PC ₂	-0,1038	0,67101	0,1811	-0,0187	-0,159502	0,6926011	0,026313
PC ₃	-0,0505	0,22886	-0,5625	-0,1393	0,7750196	0,0920526	0,011678
PC ₄	0,69682	0,00427	0,01344	0,02668	0,0531129	0,1363459	-0,701514
PC ₅	-0,0752	-0,0498	0,60948	0,56258	0,5509159	0,0196112	0,003623

b)

Componentes	F ₁	F ₂	F ₃	Q _R	T ₁₇	T ₂₇	T ₇₁
PC ₁	0,05956	0,69599	-0,0855	0,705	-0,087186	-0,010509	-0,0014
PC ₂	-0,0503	0,68195	-0,2253	-0,6936	0,0227068	-0,006528	-0,00325
PC ₃	0,00975	0,22152	0,96908	-0,0964	0,0427759	0,0207631	-0,013068

c)

Componentes	F ₁	F ₂	F ₃	Q _R	CCl ₄ (F ₁)	CCl ₄ (F ₂)	T ₇₁
PC1	0,128417	0,66615	-0,206632	-0,1575126	-0,002162	-0,686015	-0,040314
PC2	-0,09638	0,68323	-0,141677	-0,0094096	-0,167311	0,6892968	0,024996
PC3	-0,07084	0,22195	0,5443353	-0,1546138	0,7876478	0,0706308	0,011719
PC4	0,693119	-0,0117	0,0049301	0,02955024	0,0649098	0,1511083	-0,701074
PC5	0,070857	0,01468	0,6041837	-0,5906482	-0,529679	-0,016967	-0,003537

d)

Componentes	F ₁	F ₂	F ₃	Q _R	T ₁₇	T ₂₇	T ₇₁
PC1	0,059559	0,69599	-0,085528	0,7049977	-0,087186	-0,010509	-0,0014
PC2	-0,050318	0,68195	-0,225263	-0,6936118	0,0227068	-0,006528	-0,00325
PC3	0,009746	0,22152	0,9690795	-0,0963724	0,0427759	0,0207631	-0,013068

De acordo com os resultados apresentados na Tabela 5.16 a), as variáveis selecionadas são aquelas que apresentaram os maiores valores absolutos dos autovetores, correspondente a cada componente principal (PC), em ordem foram selecionadas: F₂, F₃, CCl₄(F₁), CCl₄(F₂), e T₇₁. Ou seja, o modelo final descartaria as variáveis F₁ e Q_R. A mesma análise foi efetuada para os resultados da Tabela 6.13 b), onde o modelo final deve conter: F₂, F₃ e Q_R. Esses mesmos modelos foram selecionados para compor o modelo final da saída XCHCl₃, de acordo com os resultados apresentados nas Tabelas Tabela 6.13 c) e d).

De fato, os modelos finais produzidos pela PCA não comprovam os resultados da análise dinâmica e o conhecimento fenomenológico do processo, haja vista que a técnica descarta variáveis de extrema relação com as composições de base da torre como, por exemplo, Q_R e F₁. Além disso, em comparação com os resultados da seleção de variáveis por todas as regressões possíveis, esses modelos não estão contidos entre os 20 (vinte) melhores modelos selecionados na TRP, contidos apenas entre os cinquenta melhores.

Portanto, a análise de regressão multivariável demonstra ser mais eficiente e minuciosa em termos de escolha dos regressores candidatos. A PCA não falhou na escolha dos modelos e no

descarte de variáveis, talvez ela seja mais efetiva quando se têm um maior número de regressores candidatos ou quando exista pouco entendimento fenomenológico do problema a ser resolvido.

5.6. CONCLUSÕES DO CAPÍTULO

Com as diversas observações dos resultados oriundos das simulações em regime estacionário e dinâmico, frente às modificações das variáveis do processo e diante das avaliações em $XCCl_4$ e $XCHCl_3$, selecionou-se o conjunto de variáveis candidatas a serem avaliadas como entradas dos SV para estimativa de $XCCl_4$ e outro para $XCHCl_3$. Ao final, obteve-se um total de 11 (onze) variáveis secundárias (entradas) candidatas a compor os sensores virtuais, que são: F_1 , F_2 , F_3 , Q_R , $CCl_4 (F_1)$, $CCl_4 (F_2)$, $CHCl_3 (F_1)$, $CHCl_3 (F_2)$, T_{17} , T_{27} e T_{71} .

É importante informar que outras variáveis foram avaliadas, como, por exemplo, a vazão de refluxo da coluna que resultou em alterações insignificativas em $XCCl_4$ e $XCHCl_3$, o que foi função dos valores altos das correntes de alimentação (F_1 e F_2). A sensibilidade baixa dessa variável também foi confirmada através dos altos valores do NC da análise SVD, quando essa variável era utilizada como possível manipulada para as temperaturas dos estágios da coluna.

As variáveis candidatas foram avaliadas na etapa de seleção de variáveis (seção 5.5). Nessa etapa foram selecionados os melhores modelos compostos pelo conjunto de variáveis ótimas finais, onde cada regressor (entrada) é avaliada quanto a robustez ao modelo final. Assim, o modelo final contemplará as respectivas entradas selecionadas indicadas pela mais alta correlação com cada uma das saídas ($XCCl_4$ e $XCHCl_3$).

O método sequencial TRP foi o escolhido para seleção dos melhores modelos para inferência, ao se utilizar duas abordagens distintas de seleção de variáveis, verificou-se que a PCA é mais efetiva quando se têm um maior número regressores. As conclusões indicam que devem ser utilizadas as 11 (onze) variáveis inicialmente candidatas, que são: F_1 , F_2 , F_3 , Q_R , $CCl_4 (F_1)$, $CCl_4 (F_2)$, $CHCl_3 (F_1)$, $CHCl_3 (F_2)$, T_{17} , T_{27} e T_{71} para construção dos sensores virtuais, para inferência de $XCCl_4$ e $XCHCl_3$. Uma vez que os regressores candidatos foram selecionados, a etapa subsequente é a de coleta e armazenamento dos dados. Em um caminho alternativo, os conjuntos de dados iniciais foram submetidos à remoção de *outliers*, sendo avaliados pela análise da PCA e análise residual.

Por fim, para o caso dessa coluna de destilação de alta pureza do 1,2-DCE, as observações dos resultados das simulações, tanto em regime estacionário quanto transiente,

foram de suma importância nos resultados aqui avaliados na etapa de seleção de variáveis, haja que a escolha final dos regressores e dos melhores modelos aqui avaliados, tomou como base as variáveis inicialmente propostas como candidatas.

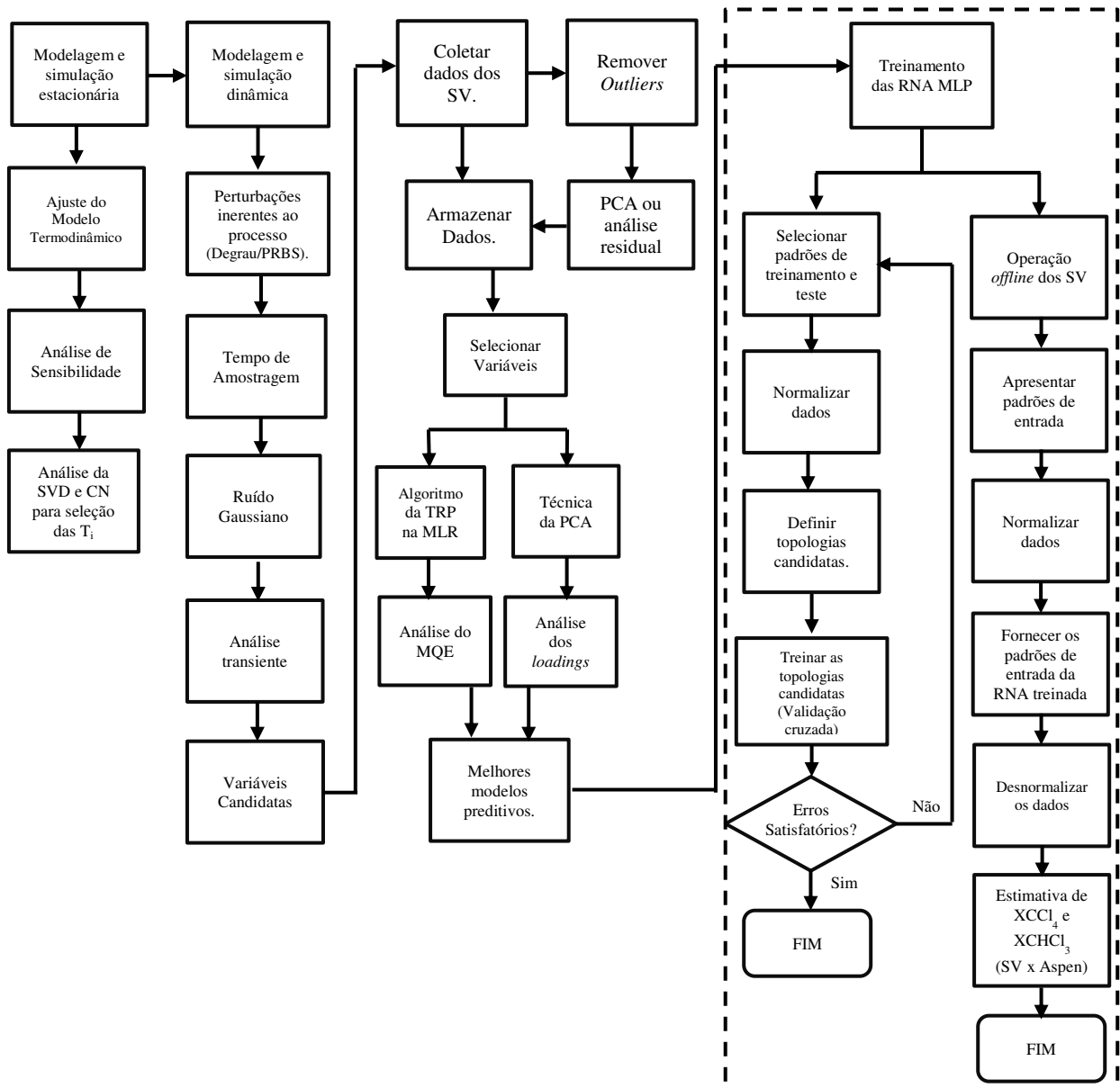
CAPÍTULO 6:

SENSORES VIRTUAIS PARA INFERÊNCIA DAS IMPUREZAS

6.1. INTRODUÇÃO

Baseando-se nos estudos realizados nas etapas anteriores, o objetivo desse capítulo é apresentar os resultados obtidos às saídas (XCCl_4 e XCHCl_3), através dos sensores virtuais desenvolvidos. As etapas desenvolvidas no treinamento das RNA, visando tal objetivo, estão em destaque no fluxograma da Figura 6.1.

Figura 6.1. Metodologia para construção dos sensores virtuais (em Destaque Treinamento dos SV).



6.2. SENSORES VIRTUAIS NEURAIIS (SVN)

6.2.1. Modelos de Inferência Avaliados para Treinamento dos SNV

Conforme definido na etapa de seleção de variáveis, aqui é proposto utilizar os dez melhores modelos (com as u_i entradas) de inferência às saídas: XCCL_4 (Dados 1 e 2) e o XCHCl_3 (Dados 3 e 4). Esses modelos estão dispostos em ordem crescente do erro médio quadrático, respectivamente, nas Tabelas 6.1 a) e b).

Tabela 6.1. Melhores modelos de inferência a serem estimados: a) Saída XCCL_4 (Dados 1 e 2); e o b) XCHCl_3 (Dados 3 e 4).

a)

MODELOS	VARIÁVEIS	MQE
M1	[u1 u3 u4 u5 u6 u11]	271,3035
M2	[u1 u2 u3 u4 u5 u6 u11]	271,3112
M3	[u1 u3 u4 u5 u11]	271,5238
M4	[u1 u2 u3 u4 u5 u11]	271,5637
M5	[u1 u2 u3 u5 u6 u11]	272,2005
M6	[u1 u2 u3 u4 u9 u10 u11]	140,6138
M7	[u1 u3 u4 u9 u10 u11]	140,6352
M8	[u1 u2 u3 u4 u10 u11]	140,7473
M9	[u1 u3 u4 u10 u11]	140,7811
M10	[u1 u2 u4 u9 u10 u11]	142,4608

b)

MODELOS	VARIÁVEIS	MQE
M11	[u1 u2 u3 u4 u7 u8 u11]	261,71
M12	[u1 u3 u4 u7 u8 u11]	261,79
M13	[u1 u2 u3 u7 u8 u11]	261,99
M14	[u1 u2 u3 u4 u8 u11]	262,02
M15	[u1 u3 u7 u8 u11]	262,07
M16	[u1 u3 u4 u9 u10 u11]	117,43
M17	[u1 u2 u3 u4 u9 u10 u11]	117,45
M18	[u1 u4 u9 u10 u11]	119,68
M19	[u1 u2 u4 u9 u10 u11]	119,70
M20	[u2 u3 u4 u9 u10 u11]	123,48

Na Tabela 6.1 a), os modelos de M1 a M5 são os melhores modelos que contêm as temperaturas e os modelos de M6 a M10 são aqueles que contêm a composição do CCl_4 nas alimentações, ambos serão utilizados aqui pelas RNA na inferência do XCCL_4 . De maneira

análoga, na Tabela 6.1 b) são apresentados os melhores modelos que contêm as temperaturas (M11 a M15) e os que contêm as composições do CHCl_3 nas alimentações (M16 a M20), ambos serão utilizados aqui pelas RNA na inferência do XCHCl_3 .

6.2.2. Topologia de Construção dos SNV

A definição do tipo de rede neural mais adequada para cada um dos sensores virtuais foi feita como base nos seguintes aspectos (SILVA et al. 2010; WERLE, 2012; HAYKIN, 2001):

i) Conjunto ideal de dados para o treinamento das RNA

Essa etapa consiste no armazenamento e na divisão dos dados das variáveis selecionadas como entradas e saídas. A definição das variáveis de entrada e saída, foi discutida na etapa de seleção de variáveis, o que proporcionou a escolha dos melhores modelos das entradas para cada uma das saídas, XCCL_4 e XCHCl_3 . Os sinais de entrada foram normalizados para um melhor condicionamento da inferência (intervalo de normalização dependente do tipo de função de ativação, seção 3.3.2, a ser utilizada), haja vista que os vetores de entrada possuem unidades diferentes. Neste caso, ao se utilizar o algoritmo de treinamento da retropropagação do erro (*error back-propagation*), foram testadas diversas combinações de funções de ativação para as camadas ocultas e de saídas, tais como: tangente hiperbólica (normalização no intervalo [-1 1]), linear (normalização no intervalo [-1 1]), logaritma-sigmóide (normalização no intervalo [0 1]) e positiva linear (normalização no intervalo [-1 1]).

A última fase dessa etapa consiste na divisão dos dados (matrizes de entradas e saídas). Como vimos na seção 3.3.3, durante o processo de aprendizagem da rede de múltiplas camadas, costuma-se dividir os dados disponíveis. Neste caso, os dados foram divididos em 3 (três) conjuntos: um conjunto de dados de treinamento, com um total de 70% dos dados, um conjunto para teste e outro para validação, cada um contendo um total de 15% dos dados. Outros percentuais foram avaliados para tais fins.

ii) Topologia ideal ou melhor estrutura dos sensores virtuais neurais

Quanto à topologia das redes neurais, Perceptron Multicamadas (MLP), empregadas para construção dos modelos de inferência, foram utilizadas diversas arquiteturas e alguns algoritmos de treinamento. Foram testadas RNA cujas camadas intermediárias (ocultas) variaram de 1 a 3 e cujo número de neurônios variaram de 5 a 30 em cada uma das camadas.

O teste das arquiteturas das RNA é feito aplicando um critério que indique o melhor conjunto dos resultados. Aqui foi adotado o critério do menor Erro Médio Quadrático (MQE), ou seja, a topologia dos sensores virtuais neurais que produziu um menor MQE de predição em comparação com o valor gerado pelo modelo dinâmico do processo (implementado no simulador *Aspen Plus Dynamics*TM).

iii) Validação dos SVN

O tipo de validação mais comum de um modelo de identificação neural é o de utilizar um percentual do total dos dados disponíveis (um novo grupo de dados, dados de teste) que relacionam as entradas com as saídas (HAYKIN, 2001). Logo, a eficiência do sensor virtual neural pode ser avaliada com relação aos dados de teste, tanto quanto a sua capacidade de generalização durante o treinamento, quanto na adoção desses dados como critério de parada na fase de treinamento. O desempenho dos SVN foi avaliado por outro conjunto de dados diferente dos utilizados no treinamento e no teste, o conjunto de validação cruzada (citado na seção 3.3.6).

Com base nesses três aspectos, foi pré-estabelecido o algoritmo de construção das redes neurais avaliadas, dividido em duas etapas:

a) Algoritmo da fase de treinamento da RNA. Etapas: 1-selecionar padrões de treinamento da RNA; 2-normalizar os dados de entrada; 3-definir topologia candidatas. 4-treinar as topologias candidatas; 5-escolha da melhor topologia (validação cruzada); 6-Erro MQE satisfatório? (Sim, seguir para etapa 7; Não, retornar para 3); 7-Fim.

b) Algoritmo da fase de operação da RNA. Etapas: 1-Aplicação da RNA; 2-Apresentar padrões de entrada; 3-Normalizar dados; 4-Aplicar padrões da RNA nas entradas das RNA treinadas; 5-Estimativa de $XCCl_4$ e $XCHCl_3$ (SV x Aspen) 6-Desnormalizar os dados; 7-Fim.

6.2.3. Ambiente de Treinamento dos SVN

O tempo de processamento de um neurônio artificial é da ordem de grandeza de 10^{-9} s (*clock* na ordem de GHz), enquanto o do neurônio biológico é de 10^{-3} s. O parâmetro de velocidade de uma RNA (acrônimo do processamento cerebral) está basicamente relacionado com o número de operações por segundo dos computadores (SILVA et al., 2010).

Verifica-se então, que o tempo de processamento do neurônio artificial é muito superior ao neurônio biológico. Entretanto, o processamento cerebral é muitas vezes mais rápido que uma

RNA, porque os neurônios da rede neural biológica trabalham com um alto grau de paralelismo entre si, enquanto que nos neurônios artificiais o nível de paralelismo é bem limitado, pois a maioria dos computadores é constituída de máquinas tipicamente sequenciais (HAYKIN, 2001). Visando acelerar a convergência, durante os treinamentos das diversas topologias das RNA, foram utilizados dois computadores: o primeiro é um AMD Phenom™ II-X4-B95 de 3,0 GHz e com 8GB de memória RAM e o segundo é um PC Intel® Core™ i7-3632-QM de 2,2 GHz e com 8GB de memória RAM. O uso de um computador mais robusto, com uma maior capacidade de processamento e de memória proporcionou melhorias no tempo de treinamento da rede, dependendo da definição da arquitetura e da dimensão do conjunto de dados.

Diversos *softwares* são utilizados para treinamento de redes neurais artificiais, dentre os quais destaca-se o uso do *software* Matlab® devido ao seu ótimo ambiente computacional de visualização e com uma computação numérica de alto desempenho. Foram utilizadas várias funções padrões deste programa e códigos fontes de programação desenvolvidos no mesmo *software* para definição das topologias candidatas.

6.2.4. Tipo de RNA e Especificações da Arquitetura Empregada

Nesse trabalho foram utilizadas as redes neurais artificiais multicamadas como modelos de identificação dos SVN, ou seja, a rede alimentada adiante do tipo MLP (*feedforward, Multi-layer Perceptron*) com algoritmo de treinamento de retropropagação do erro (*error back-propagation*).

As redes *perceptron* multicamadas são caracterizadas pela presença de pelo menos uma camada intermediária (oculta) de neurônios, situada entre a camada de entrada e a respectiva camada neural de saída. A especificação da topologia de uma rede multicamada mais apropriada para mapeamento de um problema específico é usualmente efetuada de forma empírica, pois tal dimensionamento depende, dentre outros fatores, do algoritmo de aprendizagem empregado, da forma de iniciação das matrizes de pesos, da complexidade do problema a ser solucionado, da maneira como os dados estão dispostos espacialmente (disposição espacial das amostras) e da qualidade do conjunto de dados disponíveis para treinamento (TAYLOR et al., 2006).

Mesmo não tendo uma definição precisa quanto à quantidade de camadas ocultas e o número específico de neurônios dessas camadas, o treinamento da rede do tipo MLP é sempre realizado em um único sentido, ou seja, da camada de entrada em direção à camada neural de saída. Dessa forma, devemos encontrar um ajuste ótimo tanto na definição de quantidade de

camadas escondidas quanto ao número de neurônios. A Tabela 6.2 fornece o número de camadas, neurônios da camada intermediária e as funções de ativação empregadas durante o treinamento das RNA usadas nos SVN.

A Tabela 6.2. Topologias das RNA avaliadas para construção dos sensores virtuais.

Número de Camadas Ocultas	Nº de Neurônios	Funções de Ativação Avaliadas
1	3 a 40	Tangente hiperbólica, logaritmo-sigmóide, linear e positiva linear.
2	3 a 40 (cada camada)	
3	3 a 40 (cada camada)	

Deve-se considerar o fato que as RNA do tipo MLP com uma camada escondida, em alguns casos, quando comparadas com aquelas que tenham duas ou mais camadas escondidas são normalmente menos propensas a estacionar em mínimos locais, ou seja, sua estrutura mais compacta reduz a complexidade geométrica da função que mapeia o erro médio quadrático. Por outro lado, verificou-se que os dados produzidos durante as simulações continham algumas descontinuidades, nesse caso uma ou duas camadas escondidas adicionais podem ajudar no treinamento (CURRY e MORGAN, 2006)

Verificou-se que as RNA que utilizavam poucos neurônios nas camadas ocultas (3 e 5) geravam um erro maiores em relação às amostras dos subconjuntos de teste, em virtude da insuficiência de neurônios (*underfitting*). Isso pode ser justificado pelo fato que, quando são utilizados poucos neurônios nas camadas ocultas as RNA perdem a sua habilidade suficiente à modelagem e aprendizado dos sensores virtuais (HAYKIN, 2001; PIAIA et al., 2008). Em síntese, quando era utilizado um número maior de neurônios nas camadas ocultas (30 e 40) foram proporcionados menores erros na etapa de treinamento (melhoras nessa etapa). No entanto, resultavam em erros altos na etapa de validação (resultados insatisfatórios), visto que são necessários ajustes nos pesos e nos parâmetros das RNA, o que gera problemas de sobre treinamento.

A função de transferência logarítmica e tangente sigmoide utilizadas nas camadas ocultas e de saída, foram aquelas que produziram os melhores resultados. A função logarítmica apresentou erros menores quando se utiliza número de neurônios iguais ou similares nas camadas ocultas (por exemplo: [15 15] ou [15 10]) e erros maiores quando se utiliza número de neurônios muito distintos. A Tabela 6.3 apresenta as arquiteturas finais utilizadas nas avaliações dos SV.

A Tabela 6.3. Arquiteturas finais das RNA avaliadas para construção dos sensores virtuais.

Número de Camadas Ocultas	Nº de Neurônios
1	3, 5, 10, 15, 20, 25 e 30.
2	[3 3], [5 5], [5 10], [10 5], [10 10], [15 10], [15 15], [20 15], [20 20], [25 25] e [30 30].
3	[3 3 3], [5 5 5], [5 10 5], [10 5 5], [10 10 10], [10 15 10], [15 10 5], [15 15 15], [20 15 10], [20 20 20], [25 25 25] e [30 30 30].

Assim, tomando como base a Tabela 6.3, foram avaliadas 30 RNA para cada um dos modelos finais selecionados, referentes às saídas XCCl_4 e à XCHCl_3 . Para cada modelo das saídas, foram combinadas e geradas as seguintes quantidades de redes: 7 (sete) RNA com uma camada escondida, 11(onze) RNA com duas camadas escondidas e 12 (doze) RNA com três camadas escondidas.

6.2.5. Algoritmos de Treinamento utilizados nas redes MLP

O ajuste dos pesos e dos limiares de ativação (bias) do *Perceptron* é realizado por meio da regra de aprendizagem de Hebb, cuja finalidade é a classificação de padrões que podem pertencer a uma das duas classes possíveis (HAYKIN, 2001). Ou seja, se a variável composição produzida pela rede MLP está coincidindo com a composição dinâmica produzida pelo simulador *Aspen Dynamics*TM, então os pesos sinápticos e os limiares da RNA serão incrementados (terão ajuste excitatório) proporcionalmente aos valores dos seus sinais de entrada. Caso contrário, se a composição produzida pela rede for diferente do valor desejado, os pesos sinápticos e os limiares serão então inibidos (TAYLOR et al., 2010).

Esse processo de treinamento é repetido de forma sequencial para todas as amostras de dados da etapa do treinamento, até que a composição das impurezas (XCCl_4 e XCHCl_3) produzida pela RNA MLP seja similar à desejada pela amostra (a gerada pelo simulador). Conforme já apresentado, vários algoritmos de treinamento podem ser utilizados para ajustar os pesos sinápticos e a *bias*, mas nenhum deles pode ser considerado como melhor, mas sim o mais adaptado para o objetivo específico.

Durante o treinamento das redes neurais utilizadas para construção dos SVN, o algoritmo de treinamento *Levenberg-Marquart* (LM) mostrou-se mais eficaz do que outros três empregados (ver seção 3.6). A produção de melhores resultados com o algoritmo LM, deve-se ao fato que ele consegue minimizar funções que são somas quadráticas de outras funções não

lineares, sendo considerado rápido e eficaz para RNA de tamanho moderado, como as RNA aqui construídas, utilizando o MQE como critério de desempenho (HAGAN, 1996).

6.3. RESULTADOS DAS INFERÊNCIAS DAS IMPUREZAS

Como já mencionado, a estimativa de composição em colunas de destilação de alta pureza é um tema pouco empregado. E quando aplicado refere-se ao componente-chave leve no produto de base ou do componente-chave pesado no produto de topo, ou em ambos. Entretanto, a estimativa de composição nesse trabalho, é considerada não convencional pelos seguintes fatores:

- ❖ O componente-chave de base, 1,2-DCE, possui ponto de ebulição intermediário em relação aos demais componentes presentes;
- ❖ A estimativa de composição que se deseja obter é dupla, porque a inferência das concentrações de dois componentes na mesma corrente (CCl_4 e CHCl_3). Não foi verificado algo similar na literatura consultada sobre *soft sensor* ou em problemas de natureza similar.

Sendo assim, as estimativas dos dois sensores virtuais construídos devem garantir a previsibilidade e representatividade do comportamento do sistema. Para garantir uma boa previsibilidade na inferência foram testadas diversas topologias de RNA, com diferentes funções de ativação e diferentes algoritmos de treinamento.

A especificação da configuração topológica de uma rede multicamada *Perceptron* (MLP), tais como a quantidade de camadas ocultas (intermediária) e seus respectivos números de neurônios, depende de diversos fatores a serem avaliados. Mas especificamente, a disposição espacial das amostras de treinamento e os valores iniciais atribuídos tanto aos parâmetros de treinamento como às matrizes de pesos que o *software* Matlab® assume são elementos que auxiliam na definição da melhor topologia. Como mencionado anteriormente, o ajuste dos pesos e do limiar (*bias*) de cada um dos neurônios das redes MLP avaliadas é efetuado utilizando-se o processo de treinamento supervisionado, isto é, para cada amostra dos dados de entrada obtêm-se a respectiva resposta desejada.

Das 30 RNA (para cada modelo) do tipo MLP avaliadas, foi selecionada a melhor RNA dos vinte modelos pré-selecionados. Foram avaliadas 3 (três) topologias distintas: com 1 (uma),

2 (duas) e 3 (três) camadas ocultas. Resultados mais satisfatórios foram alcançados ao utilizar-se o algoritmo de treinamento *Levenberg-Marquardt* (LVM), sendo apresentados no próximo tópico.

6.3.1. Avaliação de Desempenho e Modelos Finais para Inferência

Nesta seção, são apresentadas as melhores RNA dos melhores modelos selecionados no estudo. Essas redes foram avaliadas por meio do erro médio quadrático (MQE) nas etapas de treinamento (70% dos dados totais), validação e teste (ambas correspondem a 15% dos dados totais).

Os primeiros resultados aqui apresentados são referentes à saída XCCL_4 , onde os melhores modelos possuem combinações entre variáveis as quais estão incluídas as concentrações do CCl_4 nas alimentações são: M₁, M₂, M₃, M₄ e M₅ (Dados 1). Já os modelos os quais as concentrações das alimentações foram substituídas pelas temperaturas são: M₆, M₇, M₈, M₉ e M₁₀ (Dados 2). Logo, as Tabelas 6.4 a) e b) apresentam os valores das melhores RNA para cada um dos dez melhores modelos correlacionados com a saída XCCL_4 .

Tabela 6.4 Melhores RNA do tipo MLP (*Levenberg-Marquardt*) para inferência da composição do XCCL_4

a) M₁, M₂, M₃, M₄ e M₅; e b) M₆, M₇, M₈, M₉ e M₁₀.

a)

Número de Camadas Ocultas	Número de Neurônios da(s) Camada(s) Oculta(s)	Funções de Ativação	Número de Épocas	MQE (TREIN)	MQE (VAL)	Modelo
1	25	<i>logsig-logsig</i>	54	4,5776E-03	4,1850E-03	M1
	20	<i>logsig-logsig</i>	35	3,7794E-03	4,4585E-03	M2
	15	<i>logsig-logsig</i>	61	6,3918E-03	5,8576E-03	M3
	25	<i>logsig-logsig</i>	30	6,3365E-03	6,2236E-03	M4
	20	<i>logsig-logsig</i>	48	4,6222E-03	3,9196E-03	M5
2	[20 20]	<i>logsig-logsig</i>	54	4,8709E-03	3,4030E-03	M1
	[20 20]	<i>logsig-logsig</i>	26	3,6807E-03	3,8043E-03	M2
	[15 15]	<i>logsig-logsig</i>	57	5,0569E-03	6,1182E-03	M3
	[15 15]	<i>logsig-logsig</i>	45	5,6357E-03	4,5954E-03	M4
	[15 15]	<i>logsig-logsig</i>	51	3,2640E-03	3,5216E-03	M5
3	[25 25 25]	<i>logsig-logsig</i>	24	3,8527E-03	3,2979E-03	M1
	[15 15 15]	<i>logsig-logsig</i>	52	2,8845E-03	2,8983E-03	M2
	[15 15 15]	<i>logsig-logsig</i>	31	4,1850E-03	3,7182E-03	M3
	[20 20 20]	<i>logsig-logsig</i>	31	4,7111E-03	4,7711E-03	M4
	[10 15 10]	<i>logsig-logsig</i>	41	4,8208E-03	4,4009E-03	M5

b)

Número de Camadas Ocultas	Número de Neurônios da(s) Camada(s) Oculta(s)	Funções de Ativação	Número de Épocas	MQE (TREIN)	MQE (VAL)	Modelo
1	25	<i>logsig-logsig*</i>	66	1,5940E-03	1,0869E-03	M6
	30	<i>logsig-logsig*</i>	125	1,4853E-03	1,3545E-03	M7
	30	<i>logsig-tansig**</i>	50	2,0749E-03	1,9464E-03	M8
	30	<i>logsig-logsig*</i>	25	1,9601E-03	2,5365E-03	M9
	10	<i>logsig-logsig*</i>	64	1,6298E-03	1,9816E-03	M10
2	[10 10]	<i>logsig-logsig*</i>	42	1,1766E-03	1,1357E-03	M6
	[20 20]	<i>logsig-logsig*</i>	44	1,7475E-03	1,1401E-03	M7
	[20 20]	<i>logsig-logsig*</i>	67	2,6430E-03	1,5353E-03	M8
	[5 10]	<i>logsig-tansig**</i>	57	1,7397E-03	1,8273E-03	M9
	[20 20]	<i>logsig-logsig*</i>	66	1,8670E-03	1,1912E-03	M10
3	[10 15 10]	<i>logsig-logsig*</i>	57	7,0387E-04	6,7113E-04	M6
	[20 15 10]	<i>logsig-logsig*</i>	53	9,4473E-04	6,9784E-04	M7
	[20 20 20]	<i>tansig-tansig**</i>	58	8,9040E-04	1,1887E-03	M8
	[10 10 10]	<i>logsig-logsig*</i>	51	1,4160E-03	1,3539E-03	M9
	[20 15 10]	<i>logsig-logsig*</i>	55	9,7913E-04	8,5172E-04	M10

Legenda das Tabelas: *Normalizada no intervalo [0.1 0.9]; ** Normalizada no intervalo [-0.9 0.9].

As Tabelas 6.4 a) e b) apresentam as melhores RNA de cada modelo que obtiveram menor MQE, com uma, duas e três camadas ocultas. Verifica-se que os resultados aqui produzidos confirmam os da etapa de seleção de variáveis, ou seja, os modelos que incluem as temperaturas (Tabela 6.4 b)) são superiores aos modelos que contêm as composições do CCl₄ nas alimentações ((Tabela 6.4 a)), independentemente do número de camada ocultas. As Figuras 6.2 a) e b) descrevem, respectivamente, as 12 (doze) melhores RNA pré-selecionadas das etapas de treinamento e validação na estimativa de CCl₄.

Figura 6.2. a) Doze melhores RNA (MQE), Etapa Treinamento das RNA para a saída XCCl₄.

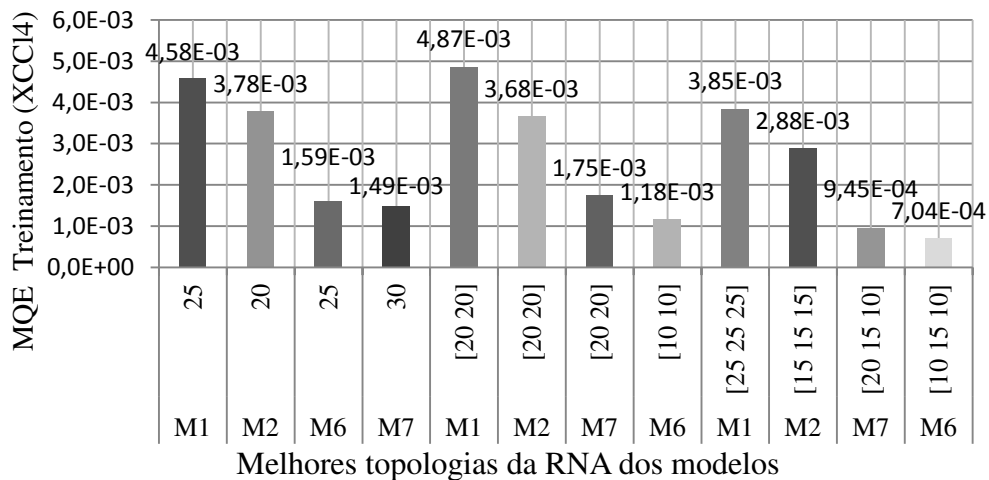
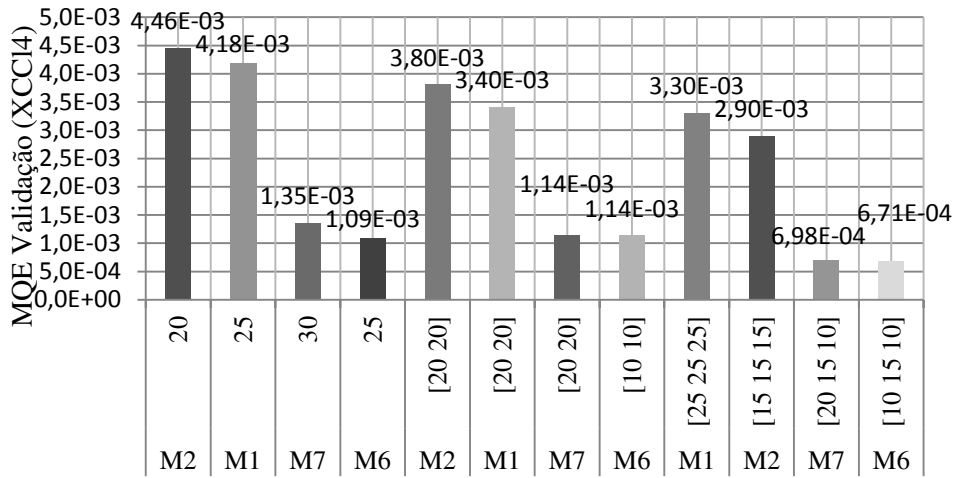


Figura 6.2. b) Doze melhores RNA (MQE), Etapa Validação das RNA para a saída XCCl_4 .



Melhores topologias da RNA dos modelos

Analisando as Figuras 6.2 a) e b), percebe-se que os modelos M6 e M7 são os que produzem melhores resultados, em qualquer número de camadas ocultas. Os valores de MQE nesses modelos são bem próximos. De fato o modelo M6 apresentou melhor desempenho e exigiu um menor esforço computacional ao utilizar um menor número de neurônios nas camadas ocultas. Logo, constata-se que o uso das variáveis u_5 e u_6 contidas no modelo M₆, proporcionam um desempenho inferior ao SV em comparação com o uso das variáveis u_9 e u_{10} , contidas em M7. Portanto, o modelo M6 (com erro médio quadrático em torno de 10^{-4}) de RNA com 3 (três) camadas ocultas e que utiliza nessa camada e na de saída a função de ativação logarítmica-sigmoide (no *software Matlab*[®] *logsig*), configuração [10 15 10], foi aquele que forneceu a melhor arquitetura à construção do sensor virtual de XCCl_4 .

De maneira análoga aos resultados produzidos à saída XCCl_4 , são feitas as mesmas observações à saída XCHCl_3 . Como definido antes, os melhores modelos M11, M12, M13, M14 e M15 (Dados 3) correlacionados com a saída XCHCl_3 , são aqueles que possuem combinações entre variáveis, nos quais estão inclusas a concentração do CHCl_3 nas alimentações. Os modelos nos quais as concentrações das alimentações foram substituídas pelas temperaturas são: M16, M17, M18, M19 e M20 (Dados 4).

Diante do exposto, as Tabelas 6.5 a) e b) apresentam os valores das melhores RNA para cada um dos dez melhores modelos correlacionados com a saída XCHCl_3 .

Tabela 6.5 Melhores RNA do tipo MLP (*Levenberg-Marquardt*) para inferência da composição do XCHCl₃: a) M11, M12, M13, M14 e M15; e b) M16, M17, M18, M19 e M20.

Número de Camadas Ocultas	Número de Neurônios da(s) Camada(s) Oculta(s)	Funções de Ativação	Número de Épocas	MQE (TREIN)	MQE (VAL)	Modelo
1	30	<i>logsig-logsig*</i>	35	7,0002E-03	7,4509E-03	M11
	25	<i>logsig-logsig*</i>	35	7,7795E-03	6,6934E-03	M12
	20	<i>logsig-logsig*</i>	48	8,1179E-03	7,1897E-03	M13
	20	<i>logsig-logsig*</i>	59	7,7467E-03	7,5334E-03	M14
	30	<i>logsig-logsig*</i>	62	8,1421E-03	7,9411E-03	M15
2	[20 20]	<i>logsig-logsig*</i>	41	6,0162E-03	6,3055E-03	M11
	[10 5]	<i>logsig-logsig*</i>	48	6,1160E-03	6,3095E-03	M12
	[15 15]	<i>logsig-logsig*</i>	38	6,9691E-03	7,3457E-03	M13
	[20 15]	<i>logsig-logsig*</i>	21	7,3332E-03	7,2408E-03	M14
	[20 20]	<i>logsig-logsig*</i>	20	7,8411E-03	6,2347E-03	M15
3	[10 5 5]	<i>logsig-logsig*</i>	38	5,7625E-03	5,9845E-03	M11
	[10 15 10]	<i>logsig-logsig*</i>	41	6,4085E-03	6,1286E-03	M12
	[15 15 15]	<i>logsig-logsig*</i>	32	6,7625E-03	6,7275E-03	M13
	[10 10 10]	<i>logsig-logsig*</i>	28	6,9045E-03	7,3883E-03	M14
	[20 15 10]	<i>logsig-logsig*</i>	32	6,8325E-03	6,3471E-03	M15

b)

1	15	<i>logsig-logsig*</i>	70	1,9878E-03	2,1306E-03	M16
	10	<i>logsig-logsig*</i>	114	1,9626E-03	1,7513E-03	M17
	20	<i>logsig-logsig*</i>	57	2,5743E-03	2,0929E-03	M18
	10	<i>logsig-logsig*</i>	83	2,9423E-03	2,6584E-03	M19
	15	<i>logsig-logsig*</i>	50	2,5919E-03	2,0447E-03	M20
2	[20 20]	<i>logsig-logsig*</i>	50	1,2083E-03	1,2444E-03	M16
	[15 10]	<i>logsig-logsig*</i>	52	1,1615E-03	1,2188E-03	M17
	[15 10]	<i>logsig-logsig*</i>	82	1,2683E-03	1,2523E-03	M18
	[20 15]	<i>logsig-logsig*</i>	53	1,4635E-03	1,9842E-03	M19
	[10 10]	<i>logsig-logsig*</i>	42	1,6785E-03	1,6728E-03	M20
3	[10 15 10]	<i>logsig-logsig*</i>	76	1,5527E-03	8,4606E-04	M16
	[15 10 5]	<i>logsig-logsig*</i>	59	1,1498E-03	8,1564E-04	M17
	[20 15 10]	<i>logsig-logsig*</i>	56	1,2154E-03	1,2041E-03	M18
	[10 15 10]	<i>logsig-logsig*</i>	84	1,3354E-03	1,1683E-03	M19
	[10 15 10]	<i>logsig-logsig*</i>	56	1,2288E-03	1,3098E-03	M20

Legenda das Tabelas: *Normalizada no intervalo [0.1 0.9]; ** Normalizada no intervalo [-0.9 0.9].

Os resultados produzidos e apresentados nas Tabelas 6.5 b), também confirmam que quaisquer modelos que incluam as temperaturas (variáveis u_9 e u_{10}), apresentam desempenho superior aos modelos que contêm as composições do CHCl_3 nas alimentações (variáveis u_7 e u_9), independentemente do número de camada ocultas. Por isso, quando se deseja maior acurácia na estimativa de XCHCl_3 , deve-se recorrer aos resultados produzidos na Tabela 6.5 b).

As Figuras 6.3 a) e b) descrevem, respectivamente, as 12 (doze) melhores RNA pré-selecionadas das etapas de treinamento e validação na estimativa de XCHCl_3 .

Figura 6.3. a) Doze melhores RNA (MQE) da etapa de treinamento das RNA para a saída XCHCl_3 .

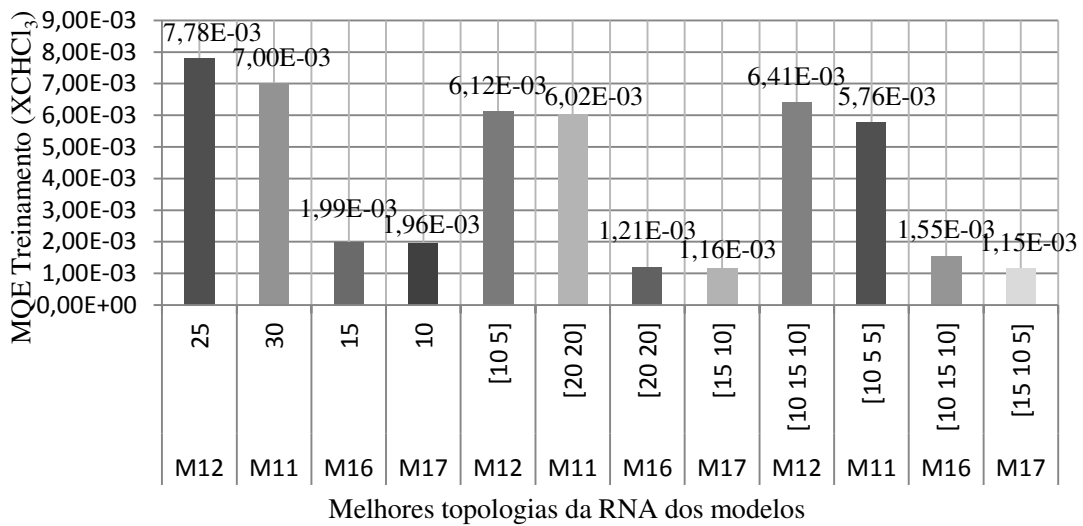
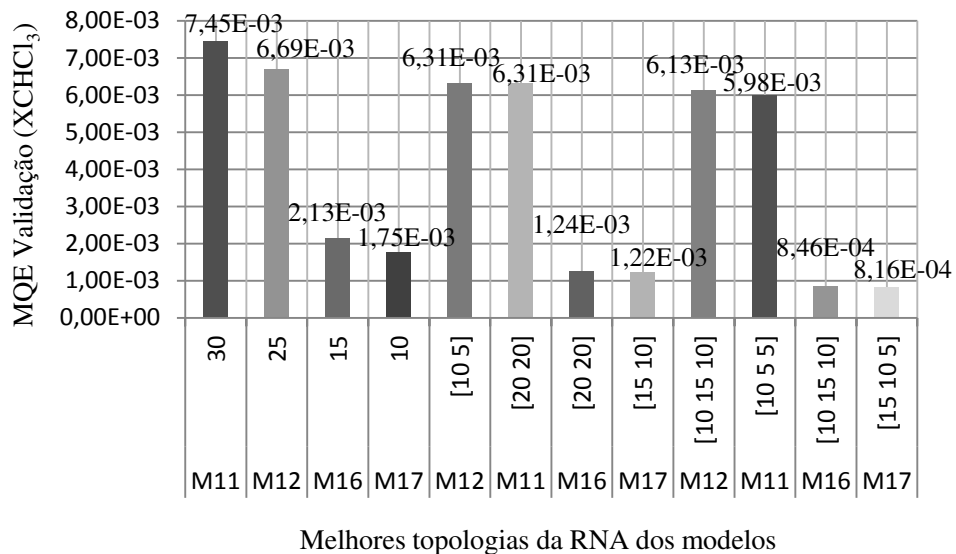


Figura 6.3. b) Doze melhores RNA (MQE) da etapa de validação das RNA para a saída XCHCl_3 .



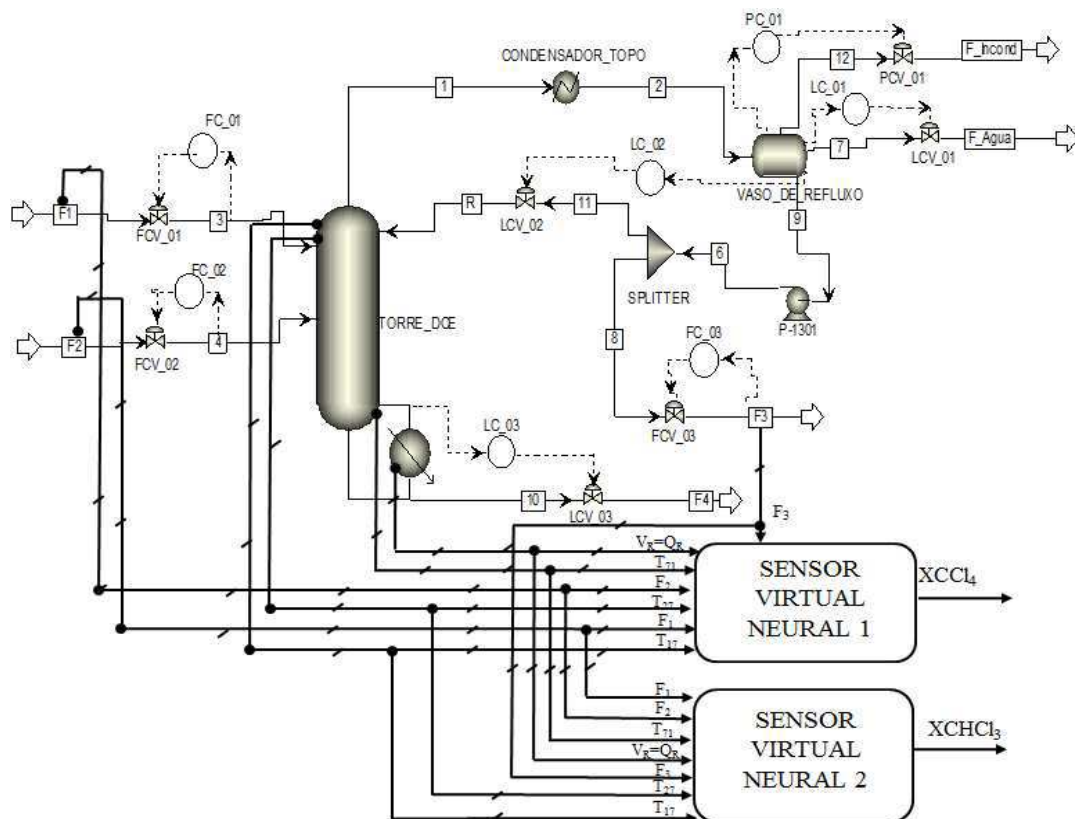
De acordo com as Figuras 6.3 a) e b), os modelos M16 e M17 produziram os melhores resultados, com quaisquer número de camadas ocultas das etapas de treinamento e validação, apresentando valores de MQE similares. De fato, o modelo M17 apresentou melhor desempenho, pois apresenta três das doze melhores RNA apresentadas. Além disso, M17 requer um menor esforço computacional ao utilizar um menor número de neurônios na camada oculta.

Por fim, o modelo M17 (com erro médio quadrático em torno de 10^{-4}) da RNA com 3 (três) camadas ocultas e que utiliza nessa camada e na de saída, a função de ativação logarítmica-sigmoide (no *software* Matlab[®] *logsig*), configuração [15 10 5], foi aquele que forneceu a melhor arquitetura à construção do sensor virtual de XCHCl₃.

6.3.2. Análise Gráfica dos Sensores Virtual de XCCl₄ e XCHCl₃

Nesta seção, são apresentadas às análises gráficas dos sensores virtuais desenvolvidos. Os modelos finais, fornecidos na seção anterior, possibilitaram a confiabilidade nas estimativas, haja vista os diversos critérios utilizados no desenvolvimento do trabalho. Conforme é apresentado na Figura 6.4, verifica-se a estrutura dos sensores virtuais em malha aberta, visando o objetivo de estimar as concentrações do CCl₄ e do CHCl₃, ambas na corrente de base da coluna.

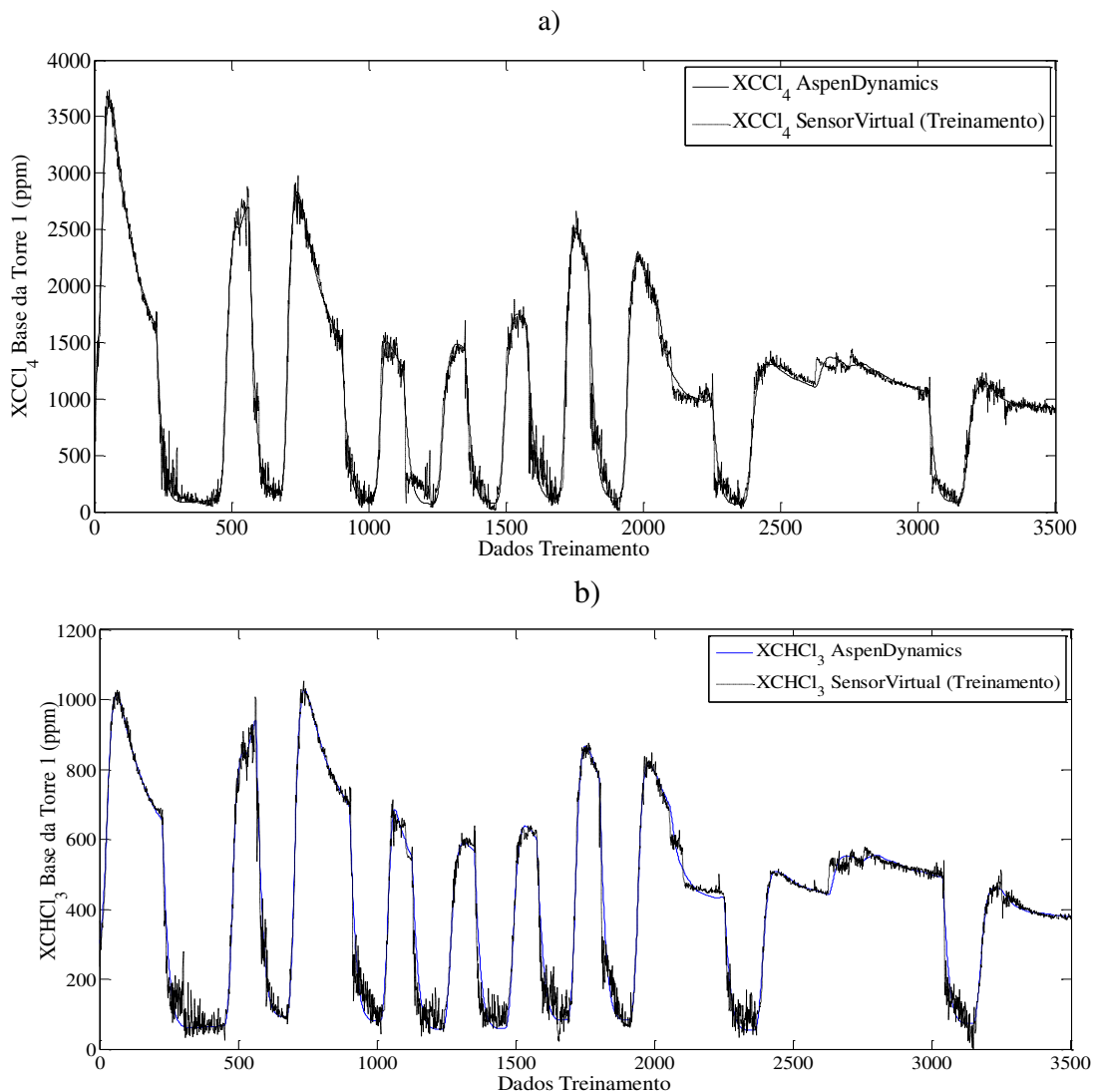
Figura 6.4. Configuração dos SVN na corrente de base da torre de destilação do 1,2-DCE.



Os dois modelos finais de inferência selecionados M6 (XCCl_4) e M17 (XCHCl_3) são definidos como modelos de múltiplas entradas e única saída (MISO). Para ambos os modelos são utilizadas as medições das seguintes variáveis: F_1 , F_2 , F_3 , $V_R=Q_R$, T_{17} , T_{27} e T_{71} . De acordo com a Figura 6.4, verifica-se a estrutura final proposta para a implementação dos sensores virtuais. É importante mencionar, que na planta real não existem medições físicas das temperaturas dos estágios 17 e 27. Nesse caso, os sensores físicos são baratos e práticos para instalação. Essa possibilidade é viável, haja vista que em casos de implantação dos sensores no processo, pois permite confiabilidade maior nas estimativas.

Os gráficos das Figuras 6.5 a) e b) apresentam, respectivamente, os resultados do conjunto de dados de treinamento dos sensores virtuais neurais para estimativas de XCCl_4 e XCHCl_3 , diante dos resultados produzidos pelo simulador *Aspen Dynamics*TM.

Figura 6.5. Dados de Treinamento: a) Sensor virtual neural do XCCl_4 e XCCl_4 fornecida pelo simulador
b) Sensor virtual neural do XCHCl_3 e XCHCl_3 fornecida pelo simulador.



Conforme se vê nas Figuras 6.5 a) e b), os dois sensores virtuais foram submetidos ao treinamento em um ambiente extremamente ruidoso, mesmo assim ocorre sobreposição dos valores fornecidos pelo simulador *Aspen Dynamics*TM e aqueles do treinamento dos SVN, estimador das composições do CCl_4 e do CHCl_3 na corrente de base da torre. Em uma sequência de dez treinamentos consecutivos, para as arquiteturas escolhidas, foi fornecido um MQE em torno de 10^{-4} , similares aos obtidos na etapa de validação. Em observância aos mesmos gráficos, durante a etapa de treinamento, as RNA do tipo *Perceptron* multicamadas, com topologias [10 15 10] e [15 10 5] e utilizam o algoritmo de treinamento *Levenberg-Marquardt*, foram capazes de criar padrões de aprendizado e fornecer de maneira eficaz as concentrações, XCCl_4 e XCHCl_3 .

A validação cruzada foi utilizada nesse trabalho como critério de parada das RNA, onde o treinamento é interrompido a cada C iterações (épocas) e a partir do momento em que o erro medido no conjunto de teste apresentar crescimento, o treinamento é encerrado. Vale ressaltar que essa etapa foi alimentada com um conjunto de dados distintos dos quais elas foram treinadas (15% dos dados totais), entretanto sem extrapolar os limites de treinamento.

Os resultados da etapa de validação cruzada versus os produzidos pelo simulador *Aspen Dynamics*TM para XCCl_4 e XCHCl_3 , são apresentados conforme os gráficos das Figuras 6.6 a) e b), respectivamente.

Figura 6.6. a) Sensor virtual neural do XCCl_4 e saída do simulador (Validação Cruzada);

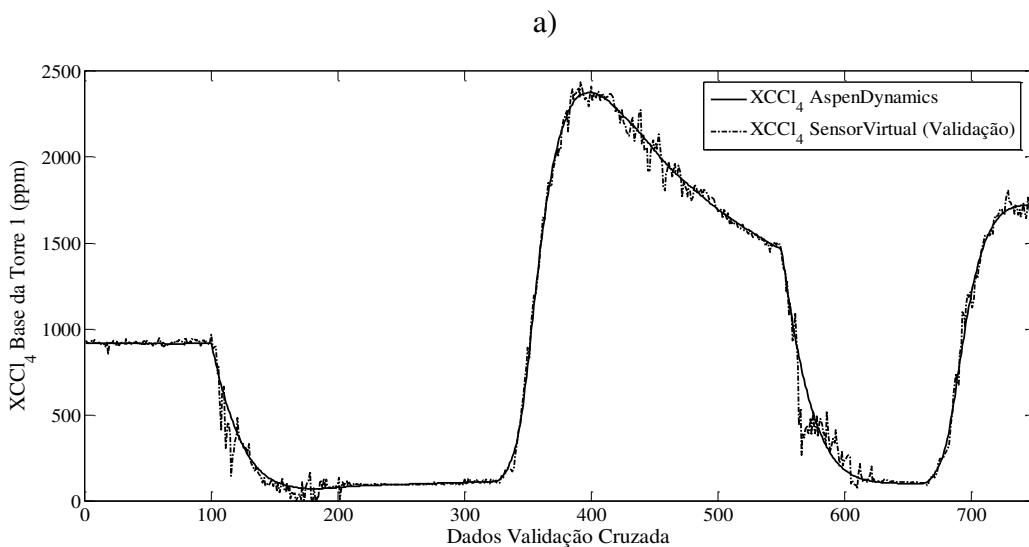
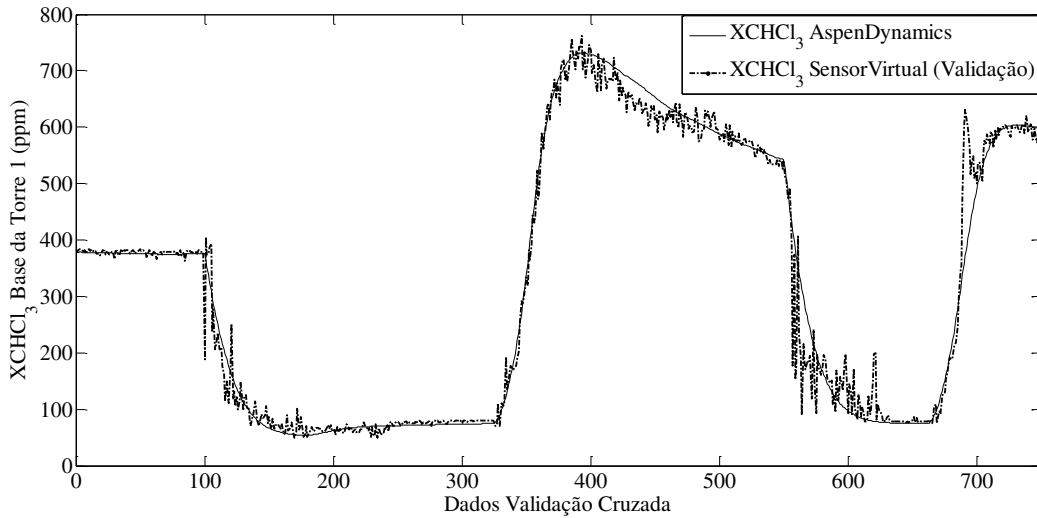


Figura 6.6. b) Sensor virtual neural do XCHCl_3 e saída do simulador (Validação Cruzada).



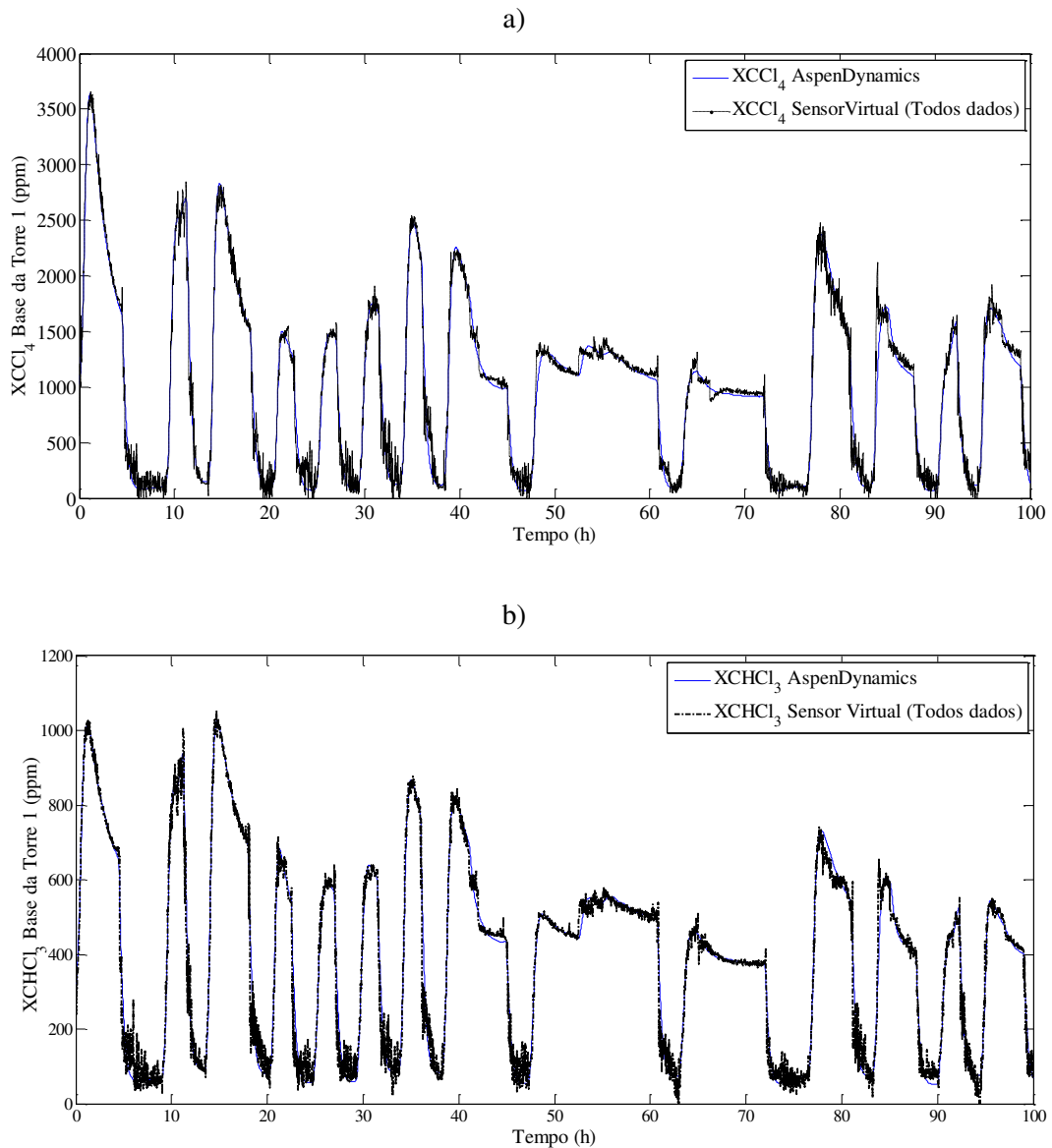
Conforme é verificado através das Figuras 6.6 a) e b), os dados de validação cruzada forneceram bons resultados, o que aumenta a capacidade de generalização do sensor virtual neural desenvolvido.

Convém ressaltar que a validação do sensor XCHCl_3 , Figura 6.6 b), apresentou um maior pico de erro, MQE em torno de $5,9515\text{E}-04$, isso se deve ao fato das baixíssimas concentrações do composto envolvido em torno de 203 ppm no valor estacionário. No entanto, isso não compromete a robustez da estimativa do sensor de XCHCl_3 , haja vista que um erro MQE de $3,4095\text{E}-4$ gerado por todo o conjunto de dados é bastante aceitável do ponto de vista de engenharia.

Já os resultados obtidos no valor inferido da composição de base do XCCl_4 e do XCHCl_3 feita pelo sensor virtual com todos os dados e durante 100 horas de operação da coluna do 1,2-DCE, comparado com os valores desejados (Aspen DynamicsTM), são, respectivamente, apresentados nos gráficos das Figuras 6.7 a) e b).

O modelo desenvolvido mostrou-se capaz de desempenhar a função de estimador do XCCl_4 , conforme resultado apresentado no gráfico da Figura 6.7 a), e isso ocorreu após 59 iterações realizadas pelo sistema (RNA MLP [10 15 10]). O que de fato também foi apresentado para a estimativa de XCHCl_3 , Figura 6.6 b), onde nesse caso a RNA tipo MLP possui arquitetura [15 10 5], necessitando esta de 48 iterações para convergir o sistema.

Figura 6.7. a) Sensor Virtual Neural dinâmico (XCCl_4) versus a saída desejada (Simulador); e b) Sensor Virtual Neural dinâmico (XCHCl_3) versus a saída desejada (Simulador).



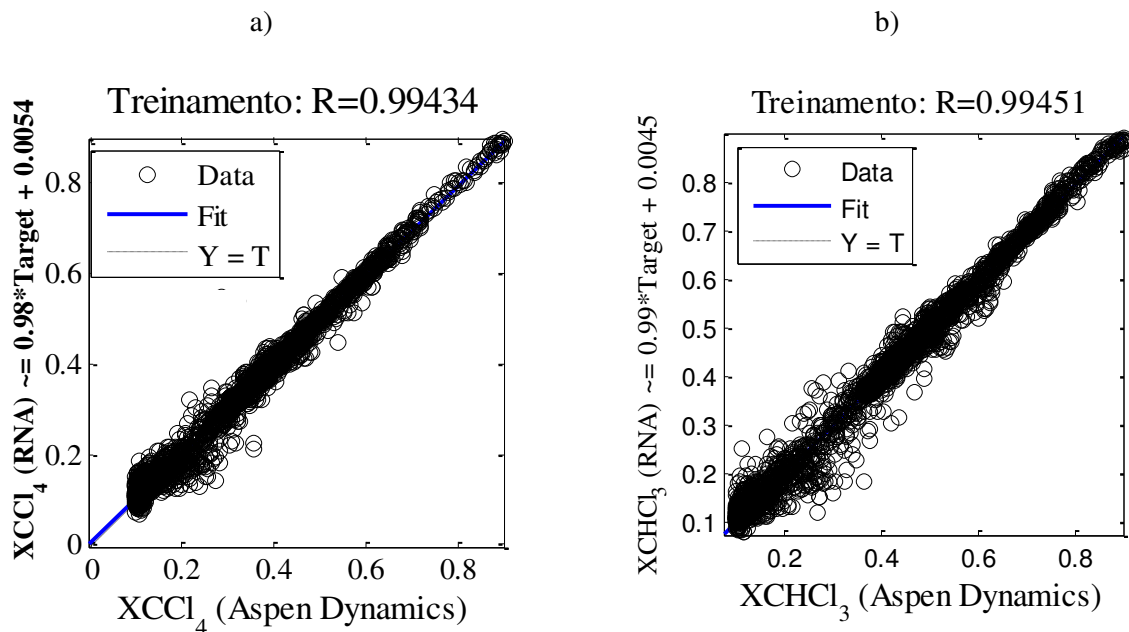
Logo, os sensores virtuais apresentaram bons resultados para inferência dos compostos críticos na base da torre do 1,2-DCE. É importante notar sobre a ocorrência de diversas perturbações nas variáveis de entrada e ruídos inseridos no sistema de medição, o que aumenta a robustez dos sensores e garante uma boa confiabilidade de estimativa. Haja vista, que em ambiente industrial a coleta dos dados é feita sobre diversas fontes de erros.

Geralmente, utiliza-se a análise de regressão dos dados como técnica para modelagem do relacionamento entre duas ou mais variáveis. Ela é um método que permite explorar e inferir a

relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias) (MONTGOMERY e RUNGE, 2014; HAYKIN, 2001).

Sendo assim, a análise de regressão pode ser vista como o estudo de correlação entre as variáveis geradas por sensor virtual em comparação com as saídas desejadas (geradas pelo simulador). os gráficos das Figuras 6.8 a) e b) mostram, respetivamente, a regressão dos dados para $XCCl_4$ e $XCHCl_3$ referente ao treinamento dos sensores virtuais.

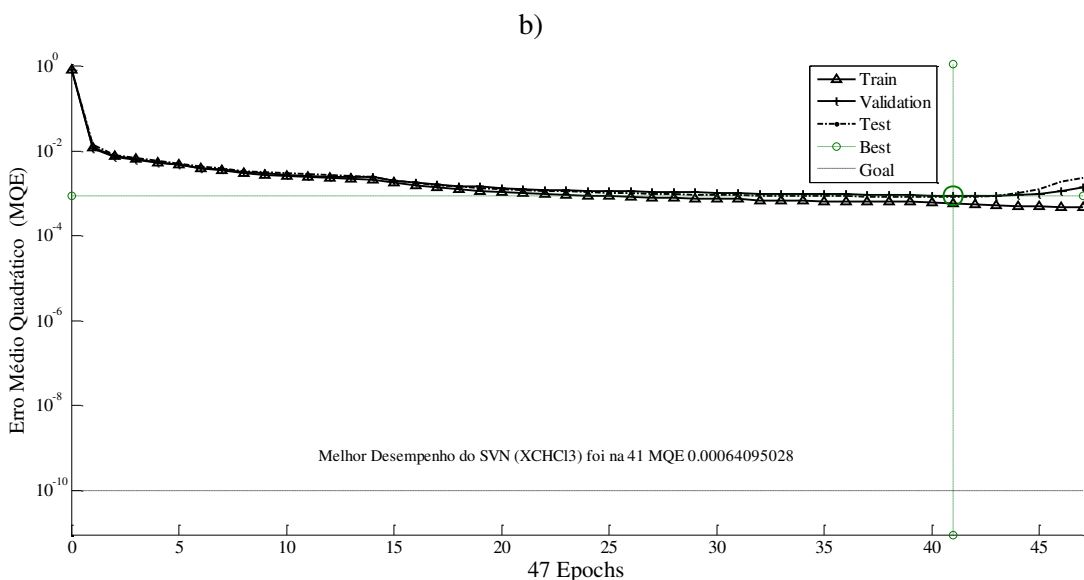
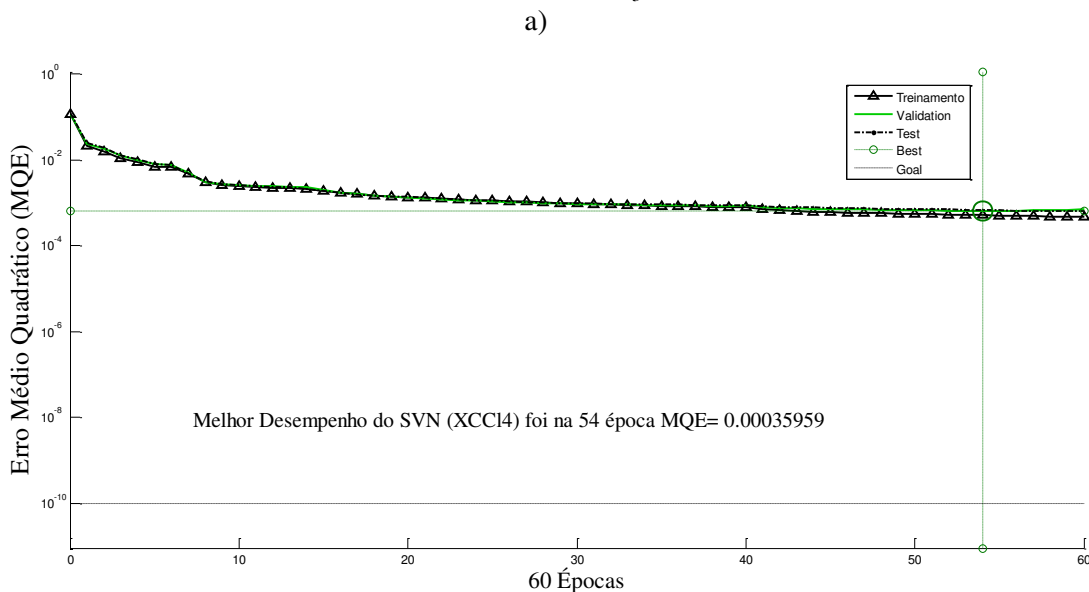
Figura 6.8. Dispersões dos dados estimados pelos SVN: a) regressão dos dados para $XCCl_4$; e b) regressão dos dados para $XCHCl_3$.



A análise da regressão é utilizada apenas com um método descritivo da análise de dados sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Os resultados apresentados pelos gráficos de treinamento das Figuras 6.8 a) e b) demonstraram ser bastante coerentes.

Ainda, para uma melhor avaliação do sensor virtual construído a partir do modelo de rede neural, foi empregado como índice de desempenho o erro médio quadrático médio; outros critérios podem ser utilizados como o da média do erro absoluto percentual (MAPE). O gráfico da Figura 6.9 a) e b) apresentam, respectivamente, os valores de MQE durante as etapas dos sensores virtuais de $XCCl_4$ e $XCHCl_3$.

Figura 6.9.a) Erro Médio Quadrático do sensor virtual de XCCl₄; b) Erro Médio Quadrático do sensor virtual de XCHCl₃.



Os gráficos de erros das Figuras 6.9 a) e b) explanam como a simulação se mostrou próxima da real, gerada pelo simulador, gerando um erro da ordem de 10^{-4} para XCCl₄ e XCHCl₃. Para a composição do CCl₄, dez treinamentos consecutivos, o menor erro alcançado foi na iteração 54 das 60 realizadas e, ainda, o MQE para todo o conjunto de dados calculado foi de $3,5959 \times 10^{-4}$.

Para a composição do CHCl₃ o desempenho foi bastante similar, exceto na etapa de validação, onde o menor erro alcançado foi na iteração 41 das 47 realizadas e, ainda, o MQE para todo o conjunto de dados calculado foi de $6,4095 \times 10^{-4}$. Todos os valores dos erros gerados

são considerados aceitáveis diante de uma análise residual, do ponto de vista de engenharia de controle podem ser insignificantes.

6.3.3. Inferência do $XCCl_4$ $XCHCl_3$ Com Remoção de Erros Grosseiros

A identificação de erros grosseiros é uma abordagem importante na construção de sensores virtuais, principalmente quando se tem conjuntos de dados de alta dimensão. Através dos *outliers* é possível encontrar informações que podem levar a erros graves. De acordo com Zanata (2005), os *outliers* levam o sensor ao cometimento de dois tipos de erros: erros espúrios e erros sistêmicos. O primeiro está relacionado a uma variação brusca da saída estimada durante um curto intervalo de tempo. Enquanto o segundo relaciona-se à diferença considerável entre o valor real e o estimado pelos SV.

Poucos trabalhos pesquisados até a presente data trazem uma abordagem quanto à significação ou melhorias alcançadas após a eliminação *outliers*. Uma exceção é o trabalho de Galicia *et al.* (2012), onde os autores utilizam o erro residual e uma abordagem bayesiana para remoção dos *outliers*, comparado o treinamento do sensor sem e com a remoção dos erros grosseiros. Especificamente, a remoção de *outliers* em sensores virtuais de colunas de alta pureza não foi encontrado nada na literatura abordando o tema. Vale ressaltar que ao trabalhar com dados fornecidos pelo modelo matemático do processo, o surgimento dos *outliers* está ligado diretamente às aproximações e imprecisões dos cálculos efetuados durante as simulações. No entanto, remover pontos sem um critério previamente definido pode levar ao comprometimento de informações preciosas sobre a dinâmica do sistema.

Nesse trabalho foram gerados dados alternativos, que consistem em dados com remoção de *outliers* utilizando as técnicas da PCA e do erro residual padronizado. Verificou-se que os conjuntos de dados pré-processados forneciam os mesmos melhores modelos de inferência que os dados não tratados, para ambas as saídas ($XCCl_4$ e $XCHCl_3$). Dessa forma, os melhores modelos avaliados são os mesmos da sessão anterior, respectivamente, M6 para saída $XCCl_4$ e M17 para saída $XCHCl_3$. Esses correspondem aos conjuntos de dados pré-processados “Dados 6 e 8”.

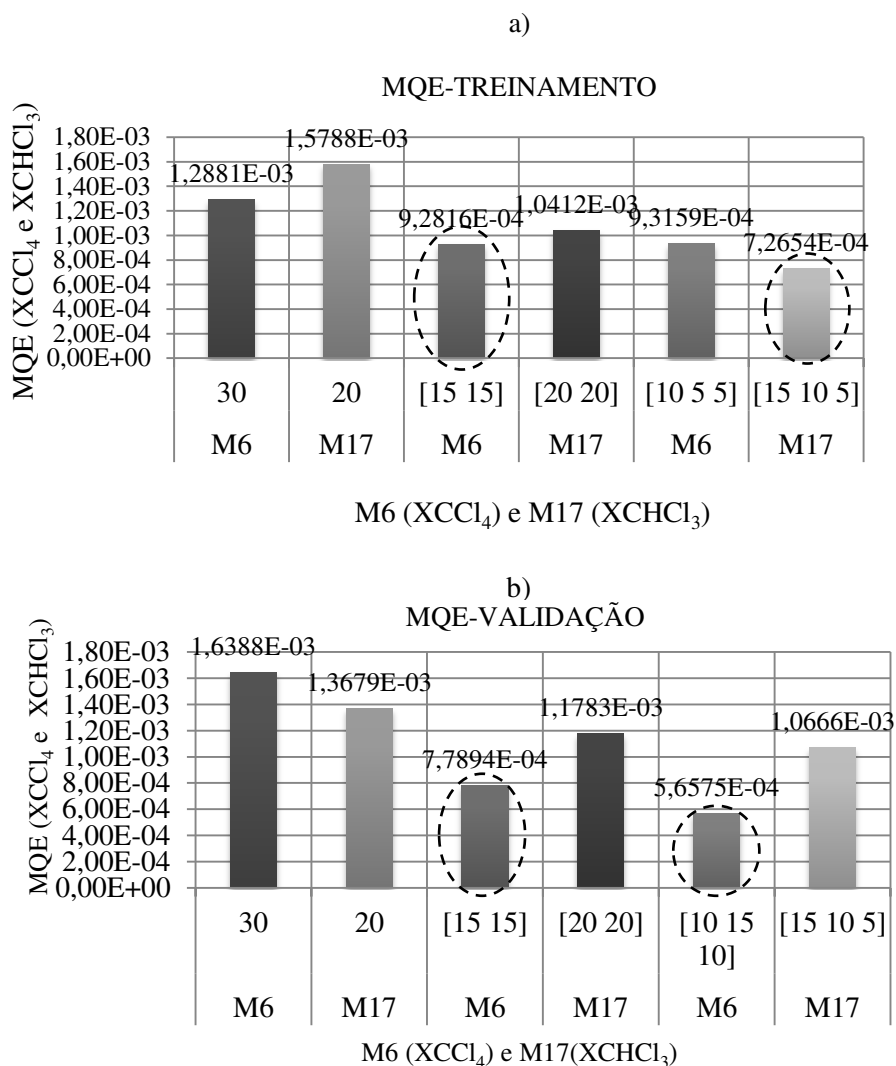
A Tabelas 6.6 a) e b) apresentam, respectivamente, os valores das melhores RNA para os dois melhores modelos correlacionados com a saída $XCCl_4$ e $XCHCl_3$. Nessas tabela, são apresentados os menores MQE (média de três simulações) relativos ao treinamento e a validação das RNA com uma, duas e três camadas ocultas, variando os números de neurônios nas mesmas.

Tabela 6.6 Melhores RNA do tipo MLP (*Levenberg-Marquardt*) para inferência da composição do XCCl_4 (M6) e do XCHCl_3 (M17).

Inferência	Modelo	Topologia	MQE Treinamento	MQE Validação	Épocas
XCCl_4	M6	30	1,2881E-03	1,6388E-03	40
XCHCl_3	M17	20	1,5788E-03	1,3679E-03	56
XCCl_4	M6	[15 15]	9,2816E-04	7,7894E-04	75
XCHCl_3	M17	[20 20]	1,0412E-03	1,1783E-03	40
XCCl_4	M6	[10 5 5]	9,3159E-04	9,2689E-04	82
XCHCl_3	M17	[15 10 5]	7,2654E-04	1,0666E-03	42

Conforme apresentado na Tabela 6.6, os melhores modelos produziram às saídas MQE similares, tanto no treinamento quanto na validação. Para uma melhor apresentação dos resultados, são gerados gráficos mostrando as RNA e sua arquitetura que produziram melhores erros (MQE), conforme apresentados nas Figuras 6.10 a) e b).

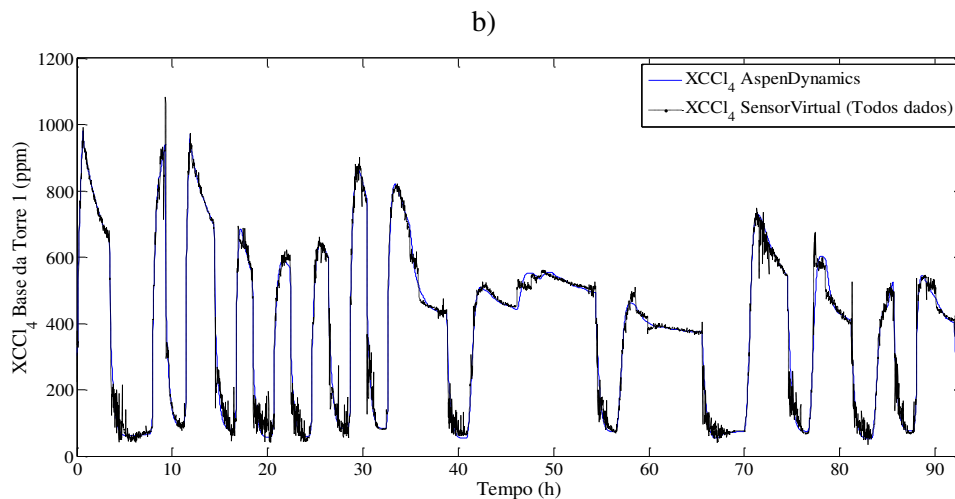
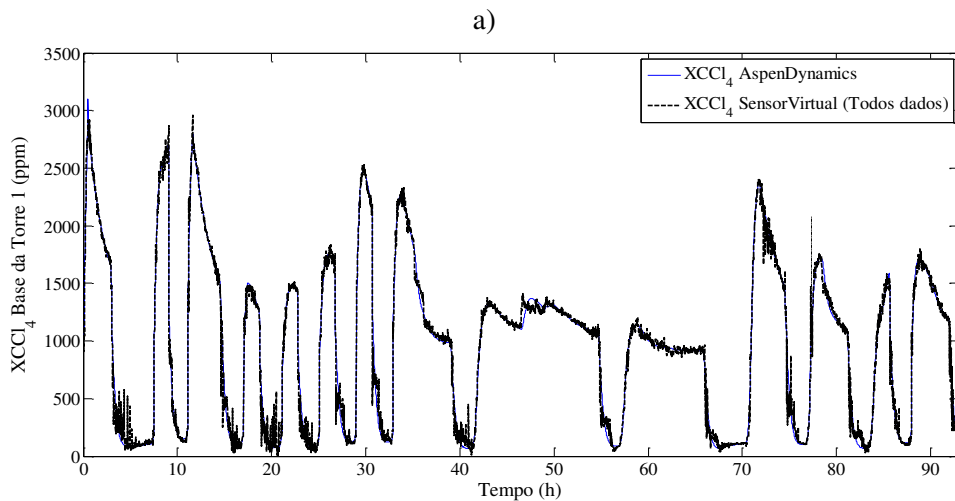
Figura 6.10. Sensores virtuais com remoção de *outliers*: a) MQE do XCCl_4 e XCHCl_3 dados de treinamento; b) MQE do XCCl_4 e XCHCl_3 dados validação.



De acordo com os gráficos das Figuras 6.10 a) e b), as melhores arquiteturas de RNA para as saídas XCCl_4 e XCHCl_3 são, respectivamente, [10 15 10] e [15 10 5]. Essas RNA foram aquelas que produziram um menor valor de MQE tanto para o treinamento quanto à validação, utilizando a função *logsig* nas camadas ocultas e de saída. Os conjuntos de dados pré-processados “Dados 6 e 8” não trouxeram melhoras significativas ao modelo quando comparados aos dados sem pré-processamento, isso se compararmos os seus respectivos MQE.

No entanto, deseja-se saber se as estimativas dos sensores virtuais com os dados pré-processados trouxeram melhorias quanto à visualização de erros espúrios. Sendo assim, os gráficos das Figuras 6.11 a) e b) apresentam, respectivamente, os treinamentos das RNA para as saídas XCCl_4 ([10 15 10]) e XCHCl_3 ([15 10 5]).

A Figura 6.11. Dados de Treinamento: a) Sensor virtual neural do XCCl_4 remoção de *outliers* b) Sensor virtual neural do XCHCl_3 remoção de *outliers*.



De acordo com as Figuras 6.11 a) e b), verifica-se que comparadas às Figuras 6.5 a) e b), aquelas dos SV com dados sem pré-processamento, não ocorre melhoria significativa nos resultados produzidos. Essas observações também foram verificadas na análise de regressão e nos dados de treinamento e validação dos sensores.

Essas conclusões devem-se a dois principais motivos: o primeiro é o fato de que os *outliers* removidos não são suficientemente quantitativos, enquanto o segundo trata mais uma vez das baixíssimas concentrações envolvidas. Ou seja, considerando que a remoção de *outliers* não é uma técnica automatizada, pois sua remoção é extremamente criteriosa. Recomenda-se nesse caso o treinamento do sensor virtual sem a etapa do pré-processamento. Entretanto, vale ressaltar que o procedimento de remoção de *outliers* trouxe melhores desempenho nos resultados em sensores onde não se tratava de processos de alta pureza, como nos trabalhos de Galicia et al. (2012) e Fortuna et al. (2005).

6.4. CONCLUSÕES DO CAPÍTULO

Nesse capítulo foram apresentados os resultados referentes à metodologia proposta para construção de dois sensores virtuais neurais, um para estimativa da concentração do composto tetracloreto de carbono e outro na estimativa do composto clorofórmio, ambos na corrente de base da torre de destilação do 1,2-Dicloroetano.

A metodologia aplicada demonstrou ser eficiente para tais fins, haja vista que os resultados produzidos demonstraram ser eficientes na busca do objetivo. Na etapa de seleção de variáveis regressoras foram selecionados os dez melhores modelos regressores, que aqui foram submetidos à uma análise de erros criteriosa e que permitiu escolher o melhor modelo preditivo. Inicialmente, os conjuntos de dados iniciais foram pré-processados, dando origem a outros conjuntos de dados com remoção de erros grosseiros (*outliers*). No entanto, verificou-se que esses novos conjuntos de dados não trouxeram melhorias significativas aos modelos de inferência.

Com relação às redes neurais avaliadas nesse trabalho, optou-se pela utilização das RNA *perceptron* multicamadas (MLP), onde foram avaliadas um total de 30 RNA para cada modelo pré-selecionado. Foram também utilizados aqui 4 (quatro) tipos de algoritmos de treinamento das RNA, chegando à conclusão que o algoritmo de treinamento *Levenberg-Marquardt* (LVM), que emprega o método de otimização de *Quasi-Newton*, resultou em erros médios quadrático muito menores que os demais avaliados. Ainda sobre as RNA, o número de camadas ocultas variou de

1 a 3, cujo número de neurônios nessas camadas foram combinações de 3-30 neurônios (a partir de 5 neurônios). Entretanto, foram selecionadas 30 melhores topologias para avaliação dos modelos finais.

Bons resultados foram alcançados para estimativa de XCCL_4 e XCHCl_3 , as melhores inferências eram seguidas pelo uso da função de ativação do tipo logarítmica-sigmoide, tanto na camada oculta quanto na camada de saída. Essa função faz com que a busca do problema de otimização não linear assuma uma única direção, logo proporciona um menor número de iteração e menores erros médios quadráticos na convergência do problema de otimização não linear.

As melhores configurações de RNA do tipo MLP obtidas para estimativa de XCCL_4 e XCHCl_3 , foram, respectivamente, [10 15 10] e [15 10 5]. A justificativa pela qual foram utilizadas três camadas ocultas, deve-se ao ponto que os dados fornecidos ao treinamento das RNA apresentavam muitas descontinuidades. Ambas as RNA apresentaram erros em torno de 10^{-4} , boa análise de regressão e um erro MQE satisfatório para o conjunto de dados totais, sendo que o sensor virtual de composição do XCCL_4 apresentou um MQE um pouco menor.

CAPÍTULO 7:

CONCLUSÕES E SUGESTÕES

7.1. CONCLUSÕES DO TRABALHO

Para o processo estudado, o método proposto demonstrou ser capaz de alcançar o objetivo desejado, ou seja, a inferência das composições das impurezas de base na torre de destilação. Para isso, a pesquisa foi dividida em etapas, as quais abordaram: a modelagem em regime estacionário e dinâmico, representação do sistema típico de controle da unidade, análise de sensibilidade estacionária e análise dinâmica, indicação das variáveis candidatas, coleta dos dados, a remoção de *outliers*, a seleção de variáveis candidatas e, por fim, o treinamento das RNA.

Durante as simulações estacionárias, observaram-se algumas dificuldades na reprodução adequada das concentrações dos compostos de interesse (CHCl_3 , CCl_4 e $\text{C}_2\text{H}_4\text{Cl}_2$) na corrente de topo e base da torre do 1,2-DCE. De fato, esse problema teve que ser corrigido, pois poderia comprometer os resultados das estimativas dos sensores virtuais. Sendo assim, foram utilizados diversos modelos termodinâmicos disponíveis no simulador *Aspen*TM, com intuito de aproximar a solução do modelo quanto às concentrações daquelas obtidas em análise de laboratório. Portanto, para representar o equilíbrio entre as fases (Líquido-Vapor e Líquido-Líquido-Vapor) foi utilizada uma abordagem do tipo γ - ϕ através do modelo NRTL-HOC.

Ainda sobre a modelagem estacionária, verificou-se os efeitos das variáveis manipuladas e distúrbios do processo sobre as concentrações XCCl_4 e XCHCl_3 , através da análise de sensibilidade. As contribuições foram além do entendimento físico do processo, contribuindo também para conversão do regime dinâmico. Na análise transiente, a configuração típica do controle da unidade foi inserida, os controladores foram sintonizados e às respostas demonstraram ser inversas de concentração, tanto para XCCl_4 quanto para XCHCl_3 . Essas observações foram importantes e determinantes na metodologia de construção dos sensores virtuais, haja vista que uma má representação física do sistema real poderia comprometer o objetivo final.

Com as observações transientes e o conhecimento fenomenológico da coluna de destilação em estudo, foram indicadas as variáveis candidatas no estudo de inferência quanto à correlação com saídas. Logo, de acordo com o método proposto foi requerido a coleta de diversas amostras de

dados, as quais foram efetuadas através da comunicação OLE do *software* VBA® e o simulador de processo *Aspen Dynamics*TM, sendo os dados armazenados em arquivos *.xls* e lidos pelo *software* Matlab®. Esses dados foram submetidos à análise de erros grosseiros, os quais foram avaliados por meio de duas técnicas distintas, a análise de erros residuais padronizados e a técnica da análise de componentes principais, dando origem assim a novos conjuntos de dados.

O trabalho trouxe relevantes contribuições quanto à seleção de variáveis para modelos modelo de inferência em colunas de alta pureza visto que foram utilizadas duas abordagens distintas para seleção de variáveis (ambas implementadas no *software* Matlab®): análise de regressão multivariada por meio da técnica de todas as regressões possíveis (TRP) e análise de componentes principais (PCA). Destaca-se que no primeiro algoritmo é permitido a inserção de outras análises estatísticas, qualquer um dos procedimentos sequenciais (*Stepwises*). A TRP demonstrou ser mais eficiente em comparação a PCA, porque diante das baixíssimas concentrações dos compostos a serem estimados, a primeira informou os melhores modelos regressores, enquanto a segunda forneceu apenas os regressores que apresentavam maior variância explicada dos dados. O que de fato não é um critério relevante diante de um pequeno grupo de variáveis candidatas correlacionadas.

Foram selecionados os dez melhores modelos de inferência para cada uma das saídas. Diante dessa informação, os melhores modelos produzidos não utilizavam as concentrações dos compostos a serem estimados nas correntes de alimentação. Uma importante conclusão do ponto de vista de construção de sensores virtuais, porque na maioria dos trabalhos desenvolvidos sobre *soft sensors* essas variáveis são cruciais na produção de bons resultados. Essa conclusão é coerente e de boa indicação, haja vista que medições físicas como as das temperaturas dos estágios são preferíveis àquelas provenientes de medições *off-line* (concentração), no caso de construção de sensores virtuais e controle de processos químicos.

As redes neurais artificiais são bastante difundidas na literatura e sem dúvidas o seu uso nesse trabalho foi motivado pelas suas características de não linearidade, robustez e capacidade de generalização ao aprender com exemplos. Ela se mostrou eficaz ao ponto de obter erros consideráveis nos resultados apresentados, quando a melhor configuração da RNA era alcançada. Quanto ao desempenho dos sensores virtuais desenvolvidos é importante notar que a coluna em questão é de destilação de alta pureza, considerada não convencional. Tratando-se de um problema de estimativa das impurezas (XCCl₄ e XCHl₃) da ordem de ppm, inferências as quais poderiam ser totalmente comprometidas ao serem inseridas perturbações e ruídos no sistema.

Entretanto, verificou-se que os dados foram coletados quando ocorrem variações degrau e PRBS nas variáveis: composições de alimentação, carga térmica do refeedor e vazões. Ainda assim, foram simulados ruídos inerentes a interferências de leituras de sensores e de transmissão de dados, os ruídos brancos Gaussianos. Os ruídos foram inseridos nas medições das vazões F_1 , F_2 e F_3 , com um desvio padrão (σ) de 2% dos valores de referência dessas variáveis. Conforme foi verificado, o treinamento dos sensores virtuais foi efetuado em um ambiente extremamente ruidoso. De fato, é preciso ainda investigar uma maneira para se obter erros ainda menores, mas que na prática industrial sabemos que é não corriqueira. Foi verificado ainda, que os dados pré-processados pela remoção de *outliers* não proporcionaram melhoras significativas às estimativas finais. Sendo preferível o uso dos dados originais, sem remoção de erros grosseiros, pois a estimativa com *outliers* podem inviabilizar o uso do sensor para controle por inferência.

7.2. SUGESTÕES PARA TRABALHOS FUTUROS

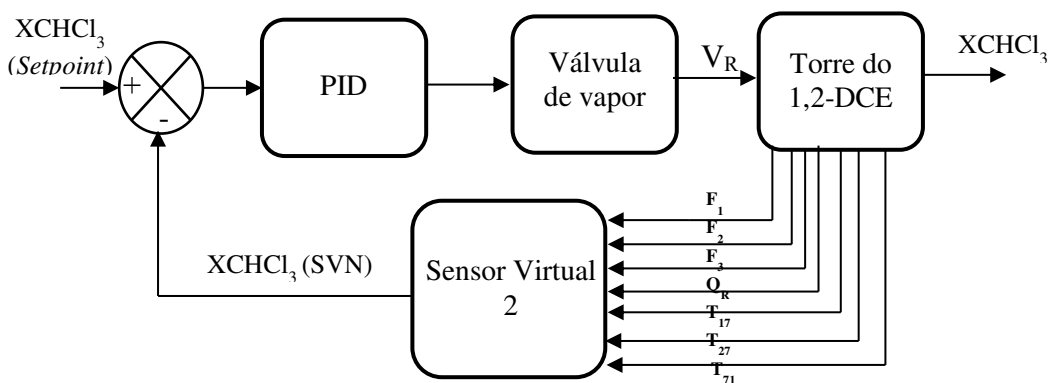
Os estudos realizados nesse trabalho sobre as estimativas por inferência se restringiram ao uso de um único tipo de rede neural artificial (*perceptron* multicamadas). Sugere-se que outras RNA possam ser avaliadas como, por exemplo, as RNA de realimentadas de Elman, visando à redução da etapa de treinamento para proposta de controle inferencial. Isso não garante melhores resultados nas estimativas, de fato o uso das RNA recorrentes possibilitam a convergência dinâmica do sistema. Sugere-se também para trabalhos futuros à comparação da estimativa do sensor com as análises analíticas de laboratório, seja pela correção de fatores ou pela proposta de implementação dos sensores virtuais em tempo real (*on-line*). Mas, vale ressaltar que o modelo termodinâmico foi ajustado às composições do CCl_4 , CHCl_3 e o $\text{C}_2\text{H}_4\text{Cl}_2$, valores esses que foram obtidos de uma média de três análises de cromatografia da linha do processo em estudo. Esse fato, torna o método proposto confiável diante da possibilidade implementação *on-line*.

7.3. PROPOSTA DE CONTROLE POR INFERÊNCIA

Uma vez atingindo o principal objetivo dessa tese, que foi o desenvolvimento de dois sensores virtuais dinâmicos para estimativa das concentrações de composto críticos no processo de purificação do 1,2-DCE, através de redes neurais artificiais, recorre-se como etapa adicional o emprego de uma RNA para aplicações de controle por inferência.

Como se trata de duas concentrações na mesma corrente do processo, recomenda-se que o controle seja feita para a variável mais crítica, a concentração do CHCl_3 na base na coluna. Nesse caso, quanto ao uso da variável manipulada ideal, os estudos através da análise SVD aqui forneceram os pares mais efetivos para tal finalidade. Sendo que, as vazões de alimentação F_1 e F_2 não podem ser utilizadas na manipulação da referida concentração, haja vista que ambos são distúrbios inerentes da planta. Uma análise feita aqui por meio matriz de ganho relativo (RGA), indicou que a vazão de vapor ($V_R=Q_R$) é a próxima da lista no sentido de interação da malha de concentração. Assim, o controle inferencial de concentração obedece ao diagrama de controle realimentação apresentado na Figura 7.1.

Figura 7.1. Diagrama de realimentação negativa, controle por inferencial do clorofórmio na corrente de base da torre do 1,2-DCE.



Para tal objetivo, talvez seja mais viável o emprego de redes recorrentes como, por exemplo, redes neurais de Elman. As RNA recorrente tipo Elman apresentam a vantagem de incorporar a dinâmica do processo dentro da própria rede, onde ocorre a realimentação interna (a cada iteração) das saídas de cada camada intermediária para as unidades de contexto e, por conseguinte, estas informações são repassadas à primeira camada oculta da rede, propagando-a para as demais durante seu treinamento. A realimentação dos valores internos em um passo seguinte à RNA, sendo capaz de armazenar as informações do passo anterior e utilizá-las a posteriore. O trabalho se concentrou na definição do melhor modelo de inferência que pudesse fornecer as estimativas das impurezas de baixíssimas concentrações. O desenvolvimento dessas minuciosas etapas possibilita a implementação de um *loop* externo, com o intuito de inferir e fazer o monitoramento das impurezas do processo em tempo real, similarmente a um analisador em linha de composição, mantendo o sistema de controle atual da planta, com a vantagem do baixo custo de aquisição, manutenção e um eficiente custo computacional.

REFERÊNCIAS BIBLIOGRÁFICAS

- AGUIRRE, L.A. **Introdução à identificação de sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais**. Belo Horizonte: Editora UFMG, Cáp 4 e 11 da 2ª Ed. 2007.
- AL-KANDARI, N.M. and JOLLIFFE, I.T. **Variable selection and interpretation of covariance principal components**. Commun. Statist.—Simul. Computat., 30, 339-354, (2001).
- ANDERSON, T.A., PRAUSNITZ, J. M., GRENS, E.A., ECKERT, C.A., HSIEH, R. e O'CONNELL, J.P. **Computer Calculations for Multicomponent Vapor Liquid and Liquid Liquid Equilibria**. Prentice Hall, 1980.
- BHAT, N. V., P. A. MINDERMAN, T. J. MCAVOY, AND N. WANG. **Modeling Chemical Process systems via neural Computation**. IEEE Control Systems Magazine, pp. 24-30, 1990.
- BISWAS, P. P.; RAY, S.; AND SAMANTA, A. N. **Nonlinear control of high purity distillation column under input saturation and parametric uncertainty**. Journal of Process Control, Pages 75–84, 2009.
- BOLF, N., IVANDIĆ, M.; and GALINEC, G. **Soft sensors for crude distillation unit product properties estimation and control**. Proc. 16th Mediterranean Conference on Control and Automation, Ajaccio, IEEE, pages 1804-1809, 2008.
- BOSTON, J. F., SULLIVAN, JR. **A new class of solution methods for multicomponent, multistage separation processes**. Canadian Journal of Chemical Engineering, v.52, p.52-63, 1974.
- BRAGA, A.P; LUDERMIR, T. B. e de LEON, A. C. P. **Redes neurais artificiais: teoria e aplicações**. Livros Técnicos e Científicos Editora, 2000.
- BRISTOL, E. H. **On a New Measure of Interaction for Multivariable Process Control**. IEEE, Transactions on Automatic Control, pages 133-134, January, 1966.
- BROSILOW, C. AND JOSEPH, B. **Techniques of model-based control**. Prentice Hall PTR, Date: April 03, 2002.
- BROSILOW, C.B.; JOSEPH, B. **Inferential control of processes**. AIChE. pp. 485, 1978.
- CAMPOS, M., GOMES, M. E PEREZ, J. **Controle Avançado e Otimização na Indústria do Petróleo**. Ed. Interciência, 2013.
- CAMPOS, M. M.; SAITO, K. **Sistemas inteligentes em controle e automação de processos**. Rio de Janeiro: Editora Ciência Moderna Ltda., 2004.
- CHONG, IL-GYO, JUN, CHI-HYUCK. **Performance of some variable selection methods when multicollinearity is present**. Chemometrics and Intelligent Laboratory Systems, 78: 103, 2005.
- COOPER, Douglas. **"Practical Process Control Using Control Station"**. Published by Control Station, Inc, Storrs, CT (2004).

CURRY, B.; and MORGAN, P. H. **Model selection in neural networks: some difficulties.** European Journal of Operational Research, 170 (2), pp. 567-577, 2006.

DIMIAN, A. C.; and SORIN, C. B. **Vinyl Chloride Monomer Process.** Chemical Process Design: Computer-Aided Case Studies. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2008.

DOYLE III, F.J. “**Nonlinear inferential control for process applications**”, J. Proc. Con. 8 (5-6), 339-353, 1998.

DUWAISH, H., GHOUTI, L., HALAWANI, T., AND MOHANDES, M. **Use of artificial neural networks process analyzers: a case study.** ESANN'2002 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium), pp. 465-470, 2002.

FINKLER, T. F. **Desenvolvimento de uma ferramenta para obtenção de modelos empíricos.** Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Química, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.

FORTUNA L., GRAZIANI S., XIBILIA M. G. **Soft sensors for product quality monitoring in debutanizer distillation columns.** CONTROL ENGINEERING PRACTICE, V.13(4), 499-508, 2005.

FORTUNA, L.; GRAZIANI, S.; RIZZO, A.; & XIBILIA, M.G. **Soft sensors for monitoring and control of industrial processes.** London, UK: Springer, 2007.

GALICIA, HECTOR J.; PETER, Q. HE; AND WANG, JIN. **Adaptive Outlier Detection and Classification for Online Soft Sensor Update.** 8th IFAC Symposium on Advanced Control of Chemical Processes Furama Riverfront, Singapore, July 10-13, 2012.

GANI, R., RUIZ, C.A. e CAMERON, I.T. **A Generalized Model for Distillation Columns – I. Model description and applications.** *Comput. Chem. Eng.*, **10, 3**, 181-198, 1986.

GARCIA, C. **Modelagem e simulação.** Editora Edusp, Cap. 2pp.32-34, 2005.

HALDER, GOPINATH. **Introduction to Chemical Engineering Thermodynamics.** Prentice-Hall of India Pvt. Ltd; 2nd Revised edition, 2014.

HAGAN, M.T., DEMUTH, H.B., BEALE, M.H. **Neural Network Design.** PWS, Boston, 1996.

HAYKIN, S. **Redes neurais: princípios e prática.** 2^a Edição, Bookman, 2001.

HIMMELBLAU, D. “**Accounts of Experiences in the Application of Artificial Neural Networks in Chemical Engineering**”, *Industrial & Engineering Chemistry Research*, 47, 5782-5796, 2008.

JANG, J. S. R.; SUN, C. T.; e MIZUTANI, E. **Neuro-fuzzy and soft computing.** Editora Prentice Hall, New Jersey, 1^a edição, 1997

JERBIĆ, I.; BOLF, N.; PAVELIĆ, H. **Development of soft sensors for debutanizer Product quality estimation and control.** Proceedings of European Congress of Chemical Engineering (ECCE-6), 2007.

JOLLIFFE I.T. **Discarding Variables in a Principal Component Analysis**. I: Artificial Data. Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 2, n. 21, pp. 160- 173, 1972.

JOLLIFFE, I. T. **Principal component analysis**. 2nd ed., Springer series in statistic, 2002.

KADLEC, P.; GABRYS, B.; STRANDT, S. **Data-driven Soft Sensors in the process industry**. Computers and Chemical Engineering, pages 795–814, 2009.

KALID, R.A. **Controle de colunas de destilação**. 2009. Disponível em: <http://www.eng.ufba.br/cecapi/ccd.html>. Último acesso em: Janeiro 2010.

KANO, M.; MIYAZAKI,K.; HASEBE,S.; e HASHIMOTO, I. **Inferential control system of distillation compositions using dynamic partial least squares regression**. Journal of Process Control, 2000.

KANO et al. **Two-stage subspace identification for softsensor design and disturbance estimation**. Journal of Process Control, pages179–186, 2009.

KIRK-OTHMER. **Encyclopedia of Chemical Technology**. John Wiley & Sons, Inc., quarta edição, 24, 851-882, 1997.

KISTER, HENRY Z. **Distillation troubleshooting**. Ed. John Wiley & Sons, 2006.

KUTNER, M. H. **Applied linear statistical models**. 5th ed., McGraw-Hill, 2004.

LIN, B. et al. **A systematic approach for soft sensor development**. **Computers & Chemical Engineering**, v. 31, n. 5-6, p. 419-425, ISSN 0098-1354. May-Jun, 2007.

LOTUFO F.A., GARCIA C. **Sensores virtuais ou soft sensors: uma introdução**. DINCON, 2008.

LUO, R.F.; SHAO, H.H.; AND ZHANG, Z.J. **Fuzzy-neural-net-based inferential control for a high-purity distillation column**. Control Eng. Practice, Vol. 3, No. 1, pp. 31-40, 1995.

LUYBEN, W. L. **Practical distillation control**. Ed. Van Nostrand Reinhold. 1st Edition, 1992.

LUYBEN, M. L.; TYREUS, B. D.; and LUYBEN, W. L. **Plantwide control desing procedure**. AIChE Journal, 43 (12), pp. 3161-3174, 1997.

LUYBEN, W.L. **Process modeling, simulation and control for chemical engineers**. Editora: McGraw-Hill International Editions, 2^a Edição, 1999.

LUYBEN, W. L. and CHIEN, I-LUNG. **Design and Control of Distillation Systems for Separating Azeotropes**. Ed. John Wiley & Sons, 2010.

LUYBEN, W. L. **Distillation design and control using Aspen simulation**. 1st ed., Copyright by John Wiley & Sons, Inc. All rights reserved 2006.

LUYBEN, W. L. **Distillation design and control using Aspen simulation**. 2nd ed., Copyright by John Wiley & Sons, Inc. All rights reserved 2013.

MCCULLOCH, W. AND PITTS, W. **A logical calculus of the ideas immanent in nervous activity**. Bulletin of Mathematical Biophysics, 5:115–133, 1943.

MING-DA MA, JING-WEIKO; SAN-WANG, MING-FENGWU, SHI-SHANGJAN, SHYAN-SHUSHIEH, DAVIDSHAN-HILLWONG. **Development of adaptive soft sensor based on statistical identification of key variables**. Control Engineering Practice (2009), Control Engineering Practice, pp. 10-19, 2009.

MONTGOMERY, D.C.; PECK, E.A.; and VINING, G.G. **Introduction to Linear Regression Analysis**. 4th ed, Wiley, 2006.

MONTGOMERY, D.C.; AND RUNGER, G. C. **Applied Statistics and Probability for Engineers**. Sixth Edition, WILEY, 2014.

MOORE, C. **Application of Singular Value Decomposition to the Design, Analysis, and Control of Industrial Processes**. Proc. Am. Control Conf., 643, 1980.

MORAIS JR, A. A. **Elaboração de um Analisador Virtual Utilizando Sistema Híbrido Neuro-Fuzzy para Inferenciar a Composição em um Processo de Destilação**. Dissertação de mestrado, Universidade Federal de Alagoas, 2011.

Neural Network Toolbox™. User's Guide, Copyright, 2010 e 2015 by The MathWorks, Inc.

PIANG, L. W. **Inferential estimation and control of chemical processes using partial least squares based model**. Thesis of Master of Engineering (Chemical).Faculty of Chemical and Natural Resources Engineering, University Technology, Malaysia, 2005.

PIAIA, J. C. Z.; MACHADO, R.A.F.; BOLZAN, A. C. **Egg Hatchability Prediction by Multiple Linear Regression and Artificial Neural Networks**. Brazilian Journal of Poultry Science. v.10, n.2, p. 97-102, apr-jun, 2008.

PLUMB, P. A; ROWE, R.C.; YORK, P.; BROWN, M. **Optimization of the predictive ability of artificial neural network (ANN) models: A comparison of three ANN programs and four classes of training algorithm**. European Journal of Pharmaceutical Sciences, v.25 p. 395–405, 2005.

POLING, B. E.; PRAUSNITZ, J. M. & O'CONNELL, J. P. **The properties of Gases and Liquids**. 5^a Ed, New York, McGraw-Hill, 2001.

RENON, H., ASSELINEU, L., COHEN, G. e RAIMBAULT, C. **Calcule sur Ordinateur des Équilibres Liquid-Vapeur et Liquide-Liquide**. Edição Technip, Paris, 1971.

RODOLFO JR., A., NUNES, L. R., ORMANJI, W. **Tecnologia do PVC**. 2^a Ed., Editora Pro-Editores, Braskem 2006.

- SAITO, J. H. **RNA modelos estacionários e dinâmicos**. UFSCAR, (Apresentação de Trabalho/Seminário). Disponível: http://www2.dc.ufscar.br/~saito/download/top_cospis/. Último acesso em: Abril 2013.
- SALTELLI, A.; TARANTOLA, S.; CAMPOLONGO, F.; RATTO, M. **Sensitivity analysis in practice**. Inglaterra: John Wiley & Sons, 2004.
- SANDLER. STANLEY L. **Chemical, Biochemical, and Engineering Thermodynamics**. 4th Edition, John Wiley & Sons, 2006.
- SCHEFFLAN, R. **Teach yourself the basics of Aspen Plus**. John Wiley & Sons, 2011.
- SEADER, J.D.; HENLEY, E.J. AND ROPER, D. K. **Separation Process Principles**. 3rd Ed. John Wiley & Sons, New York, 2011.
- SEBORG, D; EDGAR, T.; MELLICHAMP, E. D.; and DOYLE III, F. J. **Process dynamics and control**. John Wiley & Sons, Inc., 3rd Edition, Reprint (2013), 2011.
- SIEGEL, M. **Sensor modeling and simulation: can it pass the turing test vims 2001: workshop virtual and intelligent measurement systems**. Budapest, 2001. Disponível em: <<http://www.cs.cmu.edu/~mws>>. Acesso em: 15 de junho 2014.
- SILVA, A.A.R. **Aspectos Estacionário e Dinâmico de uma Destilação Azeotrópica não Convencional**. dissertação de mestrado, Universidade Federal de Campina Grande, Campina Grande, Brasil, 2002.
- SILVA, I. N.; SPATTI, D. H.; e FLAUZINOR, R. A. **Redes neurais artificiais para engenharia e Ciências aplicadas: curso prático**. São Paulo, Editora Artliver, 2010.
- SMITH, J. M.; VAN NESS; H. C.; and ABBOTT, M. M. **Introduction to Chemical Engineering Thermodynamics**. McGraw-Hill Education, 7 edition, 2007.
- SMITH, C. A. E CORRIPIO, A. **Princípios e prática de controle automático de processo**. 3^a Edição, Editora Gen e LTC, Cap. 4 e 7 pp. 226-227, 2008.
- SRINIWAS, G. R.; ARKUN, Y.; CHIEN, I-LUNG; and BUBATUNDE, A. **Ogunnaike Nonlinear identification and control of a high-purity distillation column: a case study**. J. Proc. Cont., Vol. 5, No. 3, pp 149 162, 1995.
- STEPHANOPOULOS, G. “**Chemical process control - an introduction to theory and practice**”, Prentice-Hall Inc., 1984.
- TAYLOR, B. J. **Methods and procedures for the verification and validation of artificial neural networks**. Springer, 2006.
- TREVISAN, M. G. e POPPI, R. J. **Química analítica de processos**. Quim. Nova, Vol. 29, No. 5, pp.1065-1071, 2006.

WENXIANG, L.; YING, ZHU; DEXIAN, HUANG; YONGHENG, JIANG; and YIHUI, JIN. **A New Strategy of Integrated Control and On-line Optimization on High-purity Distillation Process**. Chinese Journal of Chemical Engineering, Volume 18, Pages 66–79, 1, 2010.

WERLE, L. O. **Analísadores virtuais baseados em modelo neural para monitoramento e controle de colunas de destilação com aquecimento distribuído**. Tese de Doutorado, Florianópolis, SC, 2012.

YAN, W., SHAO, H., & WANG, X. **Soft sensing modeling based on support vector machine and Bayesian model selection**. Computers and Chemical Engineering, 28(8), 1489–1498, 2004.

ZAMPROGNA, E.; BAROLO, M.; SEBORG, D. E. **Composition estimations in a middlevessel batch distillation column using artificial neural networks**. Chemical Engineering Research and Design -Trans IChemE, v.79, part A, n.6, pp.689-696, September, 2001.

ZAMPROGNA, E., BAROLO, M. and SEBORG, D. **Optimal selection of soft sensor inputs for batch distillation columns using principal component analysis**. Journal of Process Control 15: 39–52, 2005.

ZANATA, D. R. P. **Desenvolvimento de um sensor virtual empregando redes neurais para medição da composição em uma coluna de destilação**. Dissertação de Mestrado, Universidade de São Paulo-USP, 2005.