



**Universidade Federal  
de Santa Catarina**

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
Programa de Pós Graduação em Engenharia Química**

**ANALISADORES VIRTUAIS BASEADOS EM MODELO  
NEURAL PARA MONITORAMENTO E CONTROLE DE  
COLUNAS DE DESTILAÇÃO COM AQUECIMENTO  
DISTRIBUÍDO**

**LEANDRO OSMAR WERLE**

**Florianópolis  
2012**

**LEANDRO OSMAR WERLE**

**ANALISADORES VIRTUAIS BASEADOS EM MODELO  
NEURAL PARA MONITORAMENTO E CONTROLE DE  
COLUNAS DE DESTILAÇÃO COM AQUECIMENTO  
DISTRIBUÍDO**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina como parte integrante dos requisitos exigidos para obtenção do título de Doutor em Engenharia Química.

Orientadores: Dr. Ricardo A. F. Machado  
Co-orientadora: Dr<sup>a</sup>. Claudia Sayer

**Florianópolis  
2012**



Catálogo na fonte pela Biblioteca Universitária  
da  
Universidade Federal de Santa Catarina

W489a Werle, Leandro Osmar

Analisadores virtuais baseados em modelo neural para monitoramento e controle de colunas de destilação com aquecimento distribuído [tese] / Leandro Osmar Werle ; orientador, Ricardo A. F. Machado. - Florianópolis, SC, 2012. 262 p.: il., grafs., tabs.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Química.

Inclui referências

1. Engenharia química. 2. Destilação. 3. Detectores. 4. Redes neurais (Computação). 5. Aquecimento. I. Machado, Ricardo Antonio Francisco. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Química. III. Título.

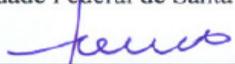
CDU 66.0

**ANALISADORES VIRTUAIS BASEADOS EM MODELO  
NEURAL PARA MONITORAMENTO E CONTROLE DE  
COLUNAS DE DESTILAÇÃO COM AQUECIMENTO  
DISTRIBUÍDO**

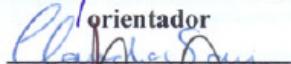
Por

**Leandro Osmar Werle**

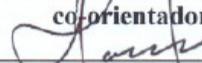
Tese julgada para obtenção do título de **Doutor em Engenharia Química**, área de Concentração de **Desenvolvimento de Processos Químicos e Biotecnológicos**, e aprovada em sua forma final pelo Programa de Pós-graduação em Engenharia Química da Universidade Federal de Santa Catarina.

  
\_\_\_\_\_  
Prof. Dr. Ricardo Antonio Francisco Machado

**orientador**

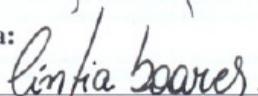
  
\_\_\_\_\_  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Claudia Sayer

**co-orientadora**

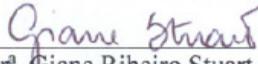
  
\_\_\_\_\_  
Prof. Dr. Leonel Teixeira Pinto

**coordenador**

**Banca Examinadora:**

  
\_\_\_\_\_  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Cintia Soares

  
\_\_\_\_\_  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Cláudia Ossanaí Ourique

  
\_\_\_\_\_  
Dr.<sup>a</sup>. Giane Ribeiro Stuart

  
\_\_\_\_\_  
Prof. Dr. Marinho Bastos Quadri

  
\_\_\_\_\_  
Prof. Dr. Pedro Henrique Hermes de Araújo

**Florianópolis, 12 de abril de 2012.**

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA****ATA DE DEFESA DE TESE DE DOUTORADO**

Ata da 109ª Sessão de Tese de Doutorado do Programa de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina, sendo candidato o Engenheiro Químico **Leandro Osmar Werle**. O trabalho tem como título "**Analísadores virtuais baseados em modelos neural para monitoramento e controle de colunas de destilação com aquecimento distribuído**", tendo como orientador o Prof. Dr. Ricardo Antonio Francisco Machado e co-orientadora Profª. Drª. Cláudia Sayer. A Banca Examinadora instituída pela Portaria nº 017/POSENQ/2012 é composta pelos seguintes membros:

**PRESIDENTE:**Prof. Dr. Ricardo Antonio Francisco Machado, **EQA****MEMBROS TITULARES:**Profª. Drª. Cintia Soares, membro interno, **EQA**Prof. Dr. Marinho Bastos Quadri, membro interno, **EQA**Prof. Dr. Pedro Henrique Hermes de Araújo, membro interno, **EQA**Profª. Drª. Cláudia Ossanaí Ourique, membro externo, **UFF**Drª. Giane Ribeiro Stuart, membro externo, **CENPES/PETROBRAS**

A sessão pública de defesa, realizada no dia **12/04/2012**, às 09:00 horas, no Auditório da Engenharia Química, foi aberta pelo Presidente da sessão, que após cumprimentar os presentes, apresentou a Banca Examinadora. Após uma breve explanação sobre as normas da UFSC e do POSENQ que regulam as defesas públicas, passou a palavra ao candidato para que este, em 50 minutos, apresentasse seu trabalho. Após a apresentação do candidato, ele foi argüido pelos integrantes da Banca Examinadora. Finalmente foi dada a oportunidade aos presentes para apresentarem perguntas ao candidato.

A Banca Examinadora, reunida em caráter sigiloso, emitiu o seguinte parecer:

- Aprovado  
 Aprovado com alterações, desde que a tese seja corrigida e entregue no prazo de até sessenta dias, nos termos sugeridos pela banca examinadora e registrados em ata  
 Reprovado

Nada mais havendo a tratar, a sessão foi encerrada às..... horas, dela sendo lavrada a presente ata, que uma vez aprovada, é assinada por todos os membros da Banca Examinadora e pela candidata.

Florianópolis, 12 de abril de 2012.

Assinaturas:

	(presidente da banca)
	(membro interno)
	(membro interno)
	(membro interno)
	(membro externo)
	(membro externo)
	(candidato)

Este trabalho foi desenvolvido no Laboratório de Controle de Processos do Departamento de Engenharia Química e Engenharia de Alimentos, Centro Tecnológico, Universidade Federal de Santa Catarina. Contou com o apoio financeiro do programa: **PRH-ANP/MCT N<sup>o</sup> 34**: Formação de Engenheiros nas Áreas de Automação, Controle e Instrumentação para a Indústria do Petróleo e Gás.





## AGRADECIMENTOS

Ao Programa de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina, em especial ao Laboratório de Controle de Processos, que ofereceu totais condições através de seu suporte logístico, para a realização do trabalho.

À Agência Nacional do Petróleo – ANP, pelo apoio financeiro prestado por intermédio de seu Programa de Recursos Humanos para Formação de Engenheiros no Setor de Petróleo e Gás – PRH34.

Ao professor Dr. Ricardo Antonio Francisco Machado, pelo apoio e idéias que contribuíram para o desenvolvimento do trabalho e, em especial, à sua confiança. A minha co-orientadora Dr<sup>a</sup>. Claudia Sayer pela orientação, pelos conhecimentos transmitidos, pela amizade e, principalmente, pela paciência.

Aos colegas de trabalho, Ronny Caytano Terán e Carlos Alberto Claumann pelas dicas preciosas que contribuíram para o desenvolvimento do projeto. Aos amigos Diogo e Iaçanã os quais estiveram sempre presentes e contribuíram com inúmeras sugestões em relação às simulações.

Gostaria de expressar meu agradecimento especial a minha amada noiva Raquel Manozzo Galante, pelo amor, apoio incondicional e conforto nos momentos difíceis, que foram imprescindíveis para prosseguimento do trabalho.

À minha família, e em especial a minha mãe Sônia Werle, pelo carinho e força, sempre estando ao meu lado.

À memória de meu pai Osmar Werle, que foi a pessoa a quem eu devo meu caráter, minha persistência e determinação.



## RESUMO

---

Colunas de destilação necessitam de sistemas de controle apropriados para obter a especificação desejada do produto e uma operação estável no processo. A medição confiável e precisa da composição, ou de alguma propriedade, dos produtos é uma das principais dificuldades em torres de destilação. Para a determinação destas, normalmente são empregadas análises de laboratório ou analisadores de processo. Contudo, existem problemas decorrentes desses métodos, quando aplicados no controle em malha fechada. Entre esses problemas estão: demora entre a coleta da amostra e a disponibilização do resultado da análise; e pouca confiabilidade dos analisadores de processo. Para tentar suprir esta dificuldade lançam-se mão de modelos inferenciais. A inferência é basicamente uma correlação de uma propriedade em função de variáveis escolhidas, que são continuamente medidas no processo de produção. O resultado da inferência é usado diretamente pelo controlador, ocasionando a atuação rápida sobre uma ou mais válvulas de controle, ou para manipular o *set point* de uma ou mais temperaturas controladas de forma a corrigir uma propriedade fora da especificação. Neste trabalho será desenvolvido um sensor por software baseado em redes neurais artificiais para inferir o valor de composição do produto de topo de uma coluna de destilação piloto, sendo esta medida disponibilizada a um sistema de controle que possui ações distribuídas de calor ao longo dos pratos. Busca-se obter um sensor robusto, porém com respostas rápidas com potencial para o desenvolvimento de estratégias de controle inferenciais avançadas. O modelo empírico será baseado em rede neural e poderá ser aplicado em uma unidade piloto de destilação de 13 pratos, processando uma mistura de etanol/água. O grande diferencial deste projeto é o fato dessa estratégia ser desenvolvida em uma torre de destilação baseada em ação distribuída de controle com aquecimentos em pontos intermediários ao longo da coluna e o desafio é que o modelo consiga representar o processo com esta característica especial. Os dados necessários para a construção, treinamento e validação da rede neural empregada no sensor, serão adquiridos através de simulações com o software comercial HYSYS® 7.0, em modo dinâmico. A validação das simulações foi realizada através da comparação com dados experimentais da unidade piloto. Foram testadas inúmeras configurações e arquiteturas de dois diferentes tipos de redes neurais: *feedforward* e redes recorrentes de Elman. Algoritmos de treinamentos distintos também foram avaliados. Adicionalmente ao projeto foi desenvolvido um segundo sensor em

malha fechada, para ser usado como um analisador de processo em tempo real, para monitoramento da planta, mantendo o sistema de controle atual. Os resultados mostraram que o modelo neural desenvolvido reproduz de forma adequada e com boa acurácia o comportamento dinâmico do processo, sendo considerado apropriado para predição de composição do produto de topo da unidade. As predições do sensor virtual foram muito similares às saídas simuladas pelo HYSYS<sup>®</sup> com erros aceitáveis. Desta forma, poderá ser utilizado para fins de monitoramento, otimização e controle no processo de destilação.

**Palavras chave:** colunas de destilação, sensor por software, redes neurais, aquecimento distribuído, controle inferencial.

## ABSTRACT

---

Distillation columns require appropriate control systems in order to obtain the desired product specification and a steady operation in the process. The reliable and precise measurement of the composition or of any other property of the product is one of the main difficulties in distillation towers. In order to determine those, laboratory analyses or process analyzers are employed. However, several problems rise from those methods, when applied in closed control loops. Among those problems might be mentioned: the long time between the sample collection and the availability of the results of the analyses; and the low reliability of the process analyzers. In order overcome this difficulty inferential models may be used. Inference is basically a correlation of a property in function of variables which are continually measured in the production process. The inference result is used either directly by the controller, resulting in fast action on one or more control valves, or to manipulate the set point of one or more controlled temperatures in order to correct an out of specification property. In this work, will be develop a soft sensor based on artificial neural networks to infer the value of the top product composition of a distillation column pilot. This measure will be available to a control system that distributed with action of heating along the column. The aim is to obtain a robust sensor, but with quick responses with potential for development of advanced inferential control strategies. The empirical model will be based on neural network and will be applied in a pilot distillation unit composed of 13 trays, processing an ethanol/water mixture. The great differential of this project is that this strategy be developed a distillation column based on distributed control action with heating on at intermediate points along the column. The challenge is that the model can represent the process with this special feature. The data needed for the construction, training and validation of the neural network employed on the sensor, were obtained through simulations with commercial software HYSYS<sup>®</sup> 7.0. The validation of the simulations was carried out through comparing to experimental data of the pilot unit. Various designs and architectures from two different types of neural networks: feedforward and Elman recurrent nets, were evaluated. Different training algorithms were also evaluated. Additionally to the project, has developed a second sensor in closed loop, to be used as an analyzer process real-time, monitoring of the plant, keeping the current control system. The results showed that the developed neural model reproduces with high accuracy the dynamic behavior of the process and considered adequate to estimate the

composition of top product of the unit. The predictions of the developed neural network model were very similar to the product composition simulated by HYSYS<sup>®</sup> with acceptable errors. Therefore, the soft sensor developed can be used for monitoring, optimization and control in the process of distillation.

**Keywords:** distillation columns, soft-sensor, neural network, distributed heating, inferential control.

## LISTA DE FIGURAS

---

### Capítulo II

<b>Figura 2.1</b> – Ilustração do caminho percorrido pelo óleo cru até o refino (Fonte: Song, 2002).....	53
<b>Figura 2.2</b> – Esquema simplificado de um sensor por software (Fonte: Meneguelo, 2007).....	56
<b>Figura 2.3</b> – Soft sensor conectado a uma planta, tornando o processo assistido por modelo inferencial. (Fonte: Lotufo e Garcia, 2008).....	58
<b>Figura 2.4</b> – Topologias das principais tipos de redes neurais artificiais. Fonte: Melo, 2004).....	65
<b>Figura 2.5</b> – Rede de 2 camadas tipo <i>feedforward</i> (Fonte: Adaptado de Menezes Junior, 2006).....	67
<b>Figura 2.6</b> – Arquitetura da rede neural artificial recorrente de Elman (Fonte: Menezes Junior, 2006).....	70
<b>Figura 2.7</b> – Arquitetura da rede neural artificial recorrente de Jordan (Fonte: Menezes Junior, 2006).....	71
<b>Figura 2.8</b> – Principais funções de transferências aplicáveis às redes neurais (Fonte: Zanata, 2005).....	73

### Capítulo III

<b>Figura 3.1</b> – Representação esquemática geral em forma de diagrama de blocos, com organização da metodologia empregada.....	87
<b>Figura 3.2</b> – Unidade piloto utilizada para obtenção de dados experimentais e parâmetros para simulações.....	89
<b>Figura 3.3</b> – Esquema geral da unidade de destilação (Fonte: Pasetti, 2005).....	91
<b>Figura 3.4</b> – Vistas dos módulos com detalhamento do prato (a) Convencional, (b) Com resistência elétrica, para o aquecimento distribuído.....	92
<b>Figura 3.5</b> – Representação geral do projeto (Fonte: Marangoni, 2005).....	92
<b>Figura 3.6</b> – Seqüência de operações para aplicação do controle distribuído.....	93
<b>Figura 3.7</b> – Estrutura das malhas de controle da coluna de destilação (Fonte: Pasetti, 2005).....	95
<b>Figura 3.8</b> – Diferença entre o controle convencional e o distribuído.....	103

**Figura 3.9** – Tela do programa MATLAB 7.6 R2008a<sup>®</sup> utilizado para realizar o treinamento da rede.....114

## **Capítulo IV**

**Figura 4.1** – Diagrama esquemático para ilustração da sequencia da apresentação dos resultados com as etapas do desenvolvimento dos analisadores virtuais para controle e monitoramento.....126

**Figura 4.2** – Perfis de fração molar de etanol (vermelho) e água (azul) para a fase líquida.....128

**Figura 4.3** – Equilíbrio líquido-vapor para o etanol e para a água, na fase vapor e na fase líquida.....128

**Figura 4.4** – Efeitos da alteração da composição da alimentação da coluna: (a) no perfil de fração de etanol e (b) no perfil de temperatura.....129

**Figura 4.5** – Necessidade energética decorrente de alterações da condição de alimentação da coluna.....131

**Figura 4.6** – Tela do programa HYSYS<sup>®</sup> em modo dinâmico com (a) esquema da unidade de destilação e seus controladores (b) malha interna da coluna de destilação, com o controle de composição do topo.....133

**Figura 4.7** – Tela do programa HYSYS<sup>®</sup> com a interface gráfica desenvolvida destacando os períodos transientes após aplicações de perturbações na temperatura e composição de alimentação.....135

**Figura 4.8** – Comparação entre os perfis de temperatura da coluna: modelo versus experimental, ambos para o controle distribuído, e ilustração do perfil de composição para o modelo.....136

**Figura 4.9** – Perturbações aplicadas na temperatura da alimentação, comparação entre experimental e simulação dinâmica, para os casos de controle convencional e distribuído.....138

**Figura 4.10** – Comportamento da malha de controle da temperatura da base, antes e depois de aplicada a perturbação na temperatura da alimentação, comparação entre experimental e simulação dinâmica.....139

**Figura 4.11** – Comportamento das malhas de controle da temperatura (experimental) e da fração de etanol (simulação) do topo da coluna, antes e depois de aplicada a perturbação na temperatura da alimentação, comparação entre experimental e simulação dinâmica.....141

**Figura 4.12** – Comportamento da malha de controle do nível do condensador, antes e depois de aplicada a perturbação na temperatura da alimentação para a simulação dinâmica distribuída.....142

<b>Figura 4.13</b> – Comportamento da malha de controle do nível da base, antes e depois de aplicada a perturbação na temperatura da alimentação, para a simulação dinâmica distribuída e convencional.....	143
<b>Figura 4.14</b> – Comportamento da malha de controle da temperatura do prato 2 com controle distribuído aplicado no prato 2 e sem ação de controle no caso convencional, para a perturbação na temperatura da alimentação, comparação entre experimental e simulação dinâmica.....	144
<b>Figura 4.15</b> – Perfis de temperatura dos pratos 2, 4 e 7 e suas respectivas derivadas em relação ao tempo para a simulação distribuída, quando aplicada a perturbação na temperatura de alimentação.....	145
<b>Figura 4.16</b> – Comportamento das frações de etanol em todos os pratos da coluna, antes e depois da aplicada a perturbação na temperatura da alimentação para simulação com controle distribuído.....	147
<b>Figura 4.17</b> – Derivada das frações volumétricas de etanol em relação ao tempo, antes e depois de aplicada a perturbação na temperatura da alimentação, para simulação com controle distribuído.....	148
<b>Figura 4.18</b> – Perturbações aplicadas na vazão da alimentação, comparação entre experimental e simulação dinâmica, para os casos de controle convencional e distribuído.....	149
<b>Figura 4.19</b> – Comportamento da malha de controle da temperatura da base, antes e depois de aplicada a perturbação na vazão da alimentação, comparação entre experimental e simulação dinâmica.....	150
<b>Figura 4.20</b> – Comportamento das malhas de controle da temperatura (experimental) e da fração de etanol (simulação) do topo da coluna, antes e depois de aplicada a perturbação na vazão da alimentação, comparação entre experimental e simulação dinâmica.....	151
<b>Figura 4.21</b> – (a) Comportamento da malha de controle do nível do condensador e (b) perfil da vazão de destilado, antes e depois de aplicada a perturbação na vazão da alimentação, para a simulação dinâmica distribuída.....	153
<b>Figura 4.22</b> – Comportamento da malha de controle do nível da base da coluna, antes e depois de aplicada a perturbação na vazão da alimentação para a simulação dinâmica distribuída e convencional.....	154
<b>Figura 4.23</b> – Comportamento da malha de controle da temperatura do prato 2 com controle distribuído aplicado no prato 2 e sem ação de controle no caso convencional, para a perturbação na vazão da alimentação, comparação entre experimental e simulação dinâmica.....	155

<b>Figura 4.24</b> – Perfis de temperatura dos pratos 2, 4 e 7 e suas respectivas derivadas em relação ao tempo para a simulação distribuída, quando aplicada a perturbação na vazão de alimentação.....	157
<b>Figura 4.25</b> – Comportamento das temperaturas em todos os pratos da coluna, com aplicação do degrau negativo (-50%) na composição volumétrica da alimentação.....	161
<b>Figura 4.26</b> – Comportamento das frações de etanol em todos os pratos da coluna, com aplicação do degrau negativo (-50%) na composição volumétrica da alimentação.....	162
<b>Figura 4.27</b> – Comparação entre os perfis de temperatura e composição dos pratos da coluna, <i>antes</i> e <i>depois</i> da aplicação da perturbação degrau negativo (-50%) na fração de etanol na alimentação.....	162
<b>Figura 4.28</b> – Comportamento das temperaturas em todos os pratos da coluna, com aplicação do degrau positivo (+50%) na composição volumétrica da alimentação.....	164
<b>Figura 4.29</b> – Comportamento das frações de etanol em todos os pratos da coluna, com aplicação do degrau positivo (+50%) na composição volumétrica da alimentação.....	164
<b>Figura 4.30</b> – Comparação entre os perfis de temperatura e composição dos pratos da coluna, <i>antes</i> e <i>depois</i> da aplicação da perturbação degrau positivo (+50%) na fração de etanol na alimentação.....	165
<b>Figura 4.31</b> – Comportamento do perfil de temperatura em todos os pratos da coluna, após aplicação do degrau positivo (+100%) com incremento no fornecimento do calor cedido ao prato 2.....	166
<b>Figura 4.32</b> – Perfil de fração volumétrica de etanol no topo da coluna com suas respectivas derivadas (a) fração máxima atingida; (b) fração mínima atingida.....	169
<b>Figura 4.33</b> – Resposta da temperatura do quarto prato da coluna, após aplicação de perturbações de graus na composição, vazão e temperatura da alimentação.....	171
<b>Figura 4.34</b> – Histogramas da composição de etanol no topo da coluna. Dados gerados pelo simulador, etapa do treinamento e validação.....	175
<b>Figura 4.35</b> – Comparação do desempenho das funções de transferência para redes com uma camada intermediária e diferentes números de neurônios, na etapa do treinamento.....	176
<b>Figura 4.36</b> – Comparação do desempenho das funções de transferência para redes com duas camadas intermediárias e diferentes números de neurônios, na etapa do treinamento.....	177

<b>Figura 4.37</b> – Comparação do desempenho das funções de transferência para redes com três camadas intermediárias e diferentes números de neurônios, na etapa do treinamento.....	177
<b>Figura 4.38</b> – Comparação entre a saída simulada e a predição das melhores redes testadas, com o algoritmo <i>Levenberg-Marquardt</i> (LM), na etapa do treinamento.....	180
<b>Figura 4.39</b> – Ampliação de uma seção da predição das 3 melhores redes testadas comparadas com saída simulada, com o algoritmo <i>Levenberg-Marquardt</i> (LM), na etapa do treinamento.....	180
<b>Figura 4.40</b> – Perfil Comparação do desempenho das melhores redes, com o algoritmo <i>Levenberg-Marquardt</i> , na etapa validação cruzada.....	182
<b>Figura 4.41</b> – (a) Comparação entre a saída simulada e a predição das 3 melhores e piores redes testadas, com o algoritmo <i>Levenberg-Marquardt</i> (LM), na etapa da validação; (b) ampliação de uma seção da etapa de validação, com descrição detalhada das perturbações aplicadas.....	184
<b>Figura 4.42</b> – Comparativo do comportamento do erro médio quadrático (MSE) dos algoritmos de treinamento analisados individualmente, com o incremento do número de neurônios para redes com 1 e 2 camadas intermediárias.....	187
<b>Figura 4.43</b> – Comparativo do comportamento do erro médio quadrático (MSE) dos algoritmos de treinamento analisados individualmente, com o incremento do número de neurônios para redes com 3 camadas intermediárias.....	188
<b>Figura 4.44</b> – Dispersão dos dados estimados pela rede e esperados, melhores resultados para cada algoritmo avaliado, na etapa de treinamento.....	192
<b>Figura 4.45</b> – Comparação do desempenho das melhores redes com os algoritmos avaliados, na etapa validação cruzada.....	193
<b>Figura 4.46</b> – (a) Comparação entre a saída simulada e a predição das melhores redes com seus respectivos algoritmos de treinamento, na etapa da validação; (b) ampliação de uma perturbação.....	194
<b>Figura 4.47</b> – Comparativo do comportamento do erro médio quadrático (MSE) das redes de Elman, com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias, na etapa do treinamento.....	196
<b>Figura 4.48</b> – Comparativo do desempenho das melhores redes de Elman avaliadas, na etapa de validação cruzada.....	196
<b>Figura 4.49</b> – Comparação entre a saída simulada e a predição das 4 melhores redes de Elman, na etapa da validação.....	199

<b>Figura 4.50</b> – (a) Comparação entre a saída simulada e a predição das 2 melhores redes gerais MPL x ELMAN, na etapa da validação; (b) ampliação de uma perturbação.....	201
<b>Figura 4.51</b> – Ilustração das arquiteturas finais das 2 redes com melhor desempenho nesta etapa, (a) rede <i>feedforward</i> –MPL com 2 camadas intermediárias; (b) rede recorrente de Elman com 3 camadas intermediárias.....	203
<b>Figura 4.52</b> – (a) Diagrama de blocos e (b) fluxograma simplificado das etapas, com a proposta do sensor por software para predição da composição de etanol no topo para ser aplicada no controle da unidade.....	204
<b>Figura 4.53</b> – Histogramas da composição de etanol no topo da coluna para rede neural em malha fechada, dados gerados pelo simulador nas etapas do treinamento e validação.....	210
<b>Figura 4.54</b> – Comparativo do comportamento do erro médio quadrático (MSE) entre redes MPL e recorrentes tipo Elman, com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias, na etapa do treinamento.....	211
<b>Figura 4.55</b> – Comparativo do comportamento do erro médio quadrático (MSE) entre redes de MPL e Recorrentes com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias, na etapa da validação.....	212
<b>Figura 4.56</b> – (a) Comparação entre a saída do simulador e a predição da rede recorrente selecionada ( <i>Elman-BR-tansig-purelin</i> [15 15]) na etapa do treinamento; (b) ampliação de uma seção.....	215
<b>Figura 4.57</b> – (a) Comparação entre a saída do simulador e a predição da rede recorrente selecionada ( <i>Elman-BR-tansig-purelin</i> [15 15]), na etapa da validação; (b) ampliação de uma seção.....	216
<b>Figura 4.58</b> – Ilustração da arquitetura final da rede recorrente de Elman selecionada para ser empregada como sensor virtual no monitoramento da composição de topo na unidade piloto de destilação.....	217
<b>Figura 4.59</b> – Diagrama esquemático da coluna de destilação com a proposta do analisador virtual para monitoramento da composição de etanol no topo.....	218

## LISTA DE TABELAS

---

### Capítulo II

<b>Tabela 2.1</b> – Relação dos principais algoritmos de treinamentos com seus respectivos métodos de otimização.....	75
-----------------------------------------------------------------------------------------------------------------------	----

### Capítulo III

<b>Tabela 3.1</b> – Lista das <i>tag's</i> das malhas de controle da unidade.....	96
<b>Tabela 3.2</b> – Principais softwares comerciais de simulação dinâmica.....	100
<b>Tabela 3.3</b> – Condições de operação empregadas na simulação estática e dinâmica.....	101
<b>Tabela 3.4</b> – Intervalos adotados para as combinações usadas nas simulações em malha fechada.....	108
<b>Tabela 3.5</b> – Intervalos adotados para as combinações usadas nas simulações em malha aberta.....	108
<b>Tabela 3.6</b> – Arquiteturas das redes neurais testadas em relação ao número de camadas e neurônios.....	106
<b>Tabela 3.7</b> – Banco de dados e dimensões das matrizes de entrada/saída nas etapas de treinamento e validação, com simulador operando em malha aberta e fechada.....	120

### Capítulo IV

<b>Tabela 4.1</b> – Seleção das 12 redes com melhor desempenho na etapa do treinamento empregando o algoritmo de treinamento <i>Levenberg-Marquard</i> , com distintas funções de transferência.....	179
<b>Tabela 4.2</b> Seleção das 3 redes com melhor e pior desempenho na etapa da validação.....	183
<b>Tabela 4.3</b> – Seleção das 18 redes com melhor desempenho na etapa do treinamento, com os algoritmos avaliados e seus respectivos erros.....	191
<b>Tabela 4.4</b> – Seleção das 28 redes de Elman com melhor desempenho na etapa do treinamento, com os algoritmos avaliados e seus respectivos erros.....	197
<b>Tabela 4.5</b> – Seleção das 15 redes com alta capacidade de generalização, com seus respectivos erros na etapa do treinamento e validação.....	214



## APÊNDICE

---

### Apêndice A

<b>Código A.1</b> – Código base do treinamento da rede neural feedforward MPL (LM_TRE).....	255
<b>Código A.2</b> – Código base do treinamento da rede neural recorrente Elman (BR_TRE).....	257
<b>Código A.3</b> – Código para cálculo da saída da rede neural para dados de treinamento (Y00) e da validação cruzada (Y01F).....	259
<b>Código A.4</b> – Código para cálculo do desempenho da rede neural (MSE_tre; ME_val; MAE_val; MSE_val; MPE_val; MAPE_val).....	261



## LISTA DE ABREVIATURAS E SIGLAS

---

AG	Algoritmos Genéticos
ASTM	<i>American Society for Testing Materials</i>
ANFIS	<i>Adaptive Neuro-Fuzzy Inference Systems</i>
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
BAM	<i>Binary Associative Memory</i>
BE	Método de otimização através de Busca Exaustiva
BFGS	Algoritmo de otimização <i>Broyden-Fletcher-Goldfarb-Shanno-BFGS</i>
BR	Algoritmo de otimização <i>Regularização de Bayesian</i>
CGP	Algoritmo de otimização <i>Gradiente conjugado de Polak-Ribière</i>
CGF	Algoritmo de otimização <i>Gradiente conjugado de Fletcher-Reeves</i>
D	Vazão de destilado
DDR	<i>Dynamic Data Reconciliation</i>
$d_E$	Dimensão de imersão
DFI	<i>Fieldbus Universal Bridge</i> , Controlador de rede <i>fieldbus</i>
DWC	Divided Wall Column
EE	Estado Estacionário
ELV	Equilíbrio líquido-vapor
F	Vazão de produto de fundo
$F_t$	Saída inferida pela rede neural para os pontos do grupo de dados
FC	<i>Flux controller</i> , Controlador de vazão
FF	FOUNDATION <i>Fieldbus</i>
FIC-100	Controlador da vazão de refluxo nas simulações com HYSYS <sup>®</sup>
FIC-102	Controlador da vazão da alimentação nas simulações com HYSYS <sup>®</sup>
FT	<i>Flow Transmitter</i> , Transmissor de vazão
FK	Filtro de Kalman
FKE	Filtro de Kalman Estendido
FV306	Válvula na corrente de alimentação
GB	Algoritmo de otimização <i>Gradiente conjugado de Powell-Beale</i>
GD	Algoritmo de otimização <i>Gradiente descendente básico</i>

GDA	Algoritmo de otimização <i>Gradiente descendente com taxa aprendizado adaptativa</i>
GDM	Algoritmo de otimização <i>Gradiente descendente com momento</i>
GDX	Algoritmo de otimização <i>Gradiente desc. com momento e taxa de aprendizado adaptativa</i>
GLP	Gás Liquefeito dePetróleo
<i>hardlim</i>	Função de transferência linear (limite de 0 a +1)
<i>hardlims</i>	Função de transferência <i>Hard-Limit</i> (limite de -1 a +1)
HIDiC	<i>Heat Integrated Distillation Column</i>
<i>initnw</i>	<i>Comando do MATLAB<sup>®</sup> com função de inicialização dos pesos da rede (Nguyen Widrow)</i>
k	Tempo discreto
LC	<i>Level controller</i> , Controlador de nível
LCP	Laboratório de Controle de Processos
LC700	Controlador Lógico Programável
<i>learnqdm</i>	Algoritmo de aprendizado gradiente descendente com momento
LIC-101	Controlador do nível do refeedor nas simulações com HYSYS <sup>®</sup>
LIC-100	Controlador do nível acumulador nas simulações com HYSYS <sup>®</sup>
<i>logsig</i>	Função de transferência <i>logarítmica sigmóide</i> (limite de 0 a +1)
LM	Algoritmo de otimização <i>Levenberg-Marquardt</i>
LNN	Rede Neural do tipo <i>linear perceptron</i>
LT	<i>Level transmitter</i> , Transmissor de nível
LV	Posicionador da válvula
MLR	<i>Multiple Linear Regression</i>
MSE	<i>Mean Squared Error</i> , Erro Médio Quadrático
<i>MSE_train</i>	Média do Erro Médio Quadrático na etapa do treinamento
<i>MSE_val</i>	Média do Erro Médio Quadrático na etapa da validação cruzada
MPE	<i>Mean Percentage Error</i> , Percentual de Erro Médio
MAPE	<i>Mean Absolute Percentage Error</i> , Percentual de Erro Médio Absoluto
<i>MAPE_val</i>	Média do Percentual de Erro Médio Absoluto na etapa da validação cruzada
MODBUS	Protocolo de comunicação serial
MPL	<i>Multi-Layer Perceptron</i>

MPC	<i>Model Predictive Control</i>
NNMPC	<i>Neural Network Model Predictive Control</i>
n	Número de pontos do conjunto de dados utilizados para treinamento da rede
<i>n</i>	Índice indicativo de tempo discreto
net	Arquitetura da rede neural
<i>newff</i>	Comando do MATLAB <sup>®</sup> para treinamento de redes <i>feedforward</i>
<i>newelm</i>	Comando do MATLAB <sup>®</sup> para treinamento de redes recorrente do tipo Elman
OSS	Algoritmo de otimização <i>One Step Secant</i>
PE	Erro Percentual
<i>poslin</i>	Função de transferência <i>positiva linear</i> (limite de 0 a +1)
PC	<i>Pressure controller</i> , Controlador de pressão
PCA	<i>Principal Component Analysis</i>
PCR	Regressão nos Componentes Principais
PI	Controlador Proporcional-Integral
PID	Controlador Proporcional-Integral-Derivativo
PLC	<i>Programmable Logic Controller</i> , Controlador Lógico Programável
PLS	<i>Partial Least Squared</i>
Pt100	Sensor de temperatura
PT	<i>Pressure Transmitter</i> , Transmissor de pressão
PRH	Programa de Recursos Humanos
PV	<i>Process Variable</i> , Variável de processo
<i>purelin</i>	Função de transferência <i>linear</i> (limite de -1 a +1)
Qa	Vazão da mistura na corrente de alimentação
q1	Número de neurônios da primeira camada
q2	Número de neurônios da segunda camada
UPGN	Unidade de Processamento de Gás Natural
R	Refluxo
R/D	Razão de refluxo
RBF	Rede Neural do tipo <i>radial basis function</i>
RNA	Rede Neural Artificial
RMSE	<i>Root Mean Square Error</i> , raiz quadrada média do erro
R2	Resistência elétrica instalada no prato 2
RP	Algoritmo de otimização <i>Resilient backpropagation</i>
RPV	Pressão de vapor da nafta leve
SA	Métodos de otimização <i>Simulated Annealing</i>
<i>satlin</i>	Função de transferência <i>satlin</i> (limite de 0 a +1)
<i>satlins</i>	Função de transferência <i>satlins</i> (limite de -1 a +1)

SDBP	Algoritmo de otimização do tipo <i>Steepest Descendent Back Propagation</i>
SHRT	<i>Self-Heat Recuperation Technology</i>
SMAR	<i>Empresa de Automação Industrial</i>
SOM	<i>Self-Organizing Maps</i>
SCG	Algoritmo de otimização <i>Gradiente conjugado escalonado</i>
<i>tansig</i>	Função de transferência <i>tangente sigmóide</i> (limites de -1 a +1)
<i>t</i>	Índice indicativo de tempo contínuo
T	Temperatura dos pratos ao longo da coluna
T <sub>a</sub>	Temperatura da mistura na corrente de alimentação
T <sub>b</sub>	Temperatura mistura na base (corrente de <i>boilup</i> )
T <sub>t</sub>	Temperatura da mistura na corrente de topo da coluna
T <sub>2</sub>	Temperatura da mistura no prato 2da coluna
T <sub>5</sub>	Temperatura da mistura no prato 5 da coluna
T <sub>7</sub>	Temperatura da mistura no prato 7 da coluna
TC	<i>Temperature controller</i> , Controlador de temperatura
TIC-101	Controlador da temperatura do <i>boillup</i> nas simulações com HYSYS <sup>®</sup>
LIC-R2	Controlador da temperatura do prato 2 nas simulações com HYSYS <sup>®</sup>
TT	<i>Temperature Transmitter</i> , Transmissor de temperatura
tr	Parâmetros da rede neural
TV	Válvula
V	Vazão de vapor no topo da coluna
VCR	Recompressão Direta de Vapor
<i>x</i>	Variável de entrada da rede
<i>x</i>	Fração volumétrica de etanol ao longo dos pratos da coluna
X	Matriz de dados de entrada do simulador HYSYS <sup>®</sup>
X <sub>a</sub>	Fração volumétrica do componente etanol na corrente de alimentação da coluna
X <sub>b</sub>	Fração volumétrica do componente etanol na corrente de base da coluna
X <sub>1</sub> a X <sub>13</sub>	Frações volumétricas do componente etanol na mistura ao longo dos respectivos pratos
X <sub>t</sub>	Fração volumétrica do componente etanol na corrente de topo da coluna
X <sub>etanol</sub>	Fração molar de etanol ao longo da coluna para simulações estáticas

$x_{H_2O}$	Fração molar de água ao longo da coluna para simulações estáticas
$X_{tre}$	Matriz com dados do HYSYS <sup>®</sup> usados como entrada da rede na etapa de treinamento
$X_{val}$	Matriz com dados do HYSYS <sup>®</sup> usados como entrada da rede na etapa de validação
$y$	Saída de ativação da rede neural
$Y$	Matriz de dados de saída do simulador HYSYS <sup>®</sup>
$Y_i$	Saída gerada pelo simulador HYSYS <sup>®</sup> para os pontos do grupo de dados
$Y_{tre}$	Matriz com dados do HYSYS <sup>®</sup> usados como saída da rede na etapa do treinamento
$Y_{val}$	Matriz com dados do HYSYS <sup>®</sup> usados como saída da rede na etapa de validação
$Z^1$	Atraso unitário



## LISTA DE SÍMBOLOS

---

$\tau$	Constante de tempo
$\alpha$	Passo de ajuste a cada iteração da rede neural
$\emptyset_{ref}$	Calor do refervedor
$\Delta T$	Diferença de temperatura



## SUMÁRIO

---

<b>Capítulo I – INTRODUÇÃO.....</b>	<b>39</b>
1.1 IMPORTÂNCIA DO ESTUDO.....	42
1.2 ORGANIZAÇÃO DO TRABALHO.....	44
1.3 OBJETIVOS.....	45
1.3.1 OBJETIVOS ESPECÍFICOS.....	46
<b>Capítulo II – REVISÃO BIBLIOGRÁFICA.....</b>	<b>47</b>
2.1 INTRODUÇÃO.....	47
2.2 PROCESSO DE DESTILAÇÃO.....	48
2.2.1 Controle de colunas.....	49
2.2.2 Transientes de operação: principais causas.....	50
2.2.3 Dificuldades do controle de composição e alternativas.....	52
2.3 CONTROLE DO PROCESSO AUXILIADO POR MODELO INFERENCIAL.....	57
2.3.1 Técnicas de geração de modelos empíricos.....	61
2.3.1.1 Métodos estatísticos.....	61
2.3.1.2 Métodos baseados em filtro de Kalman.....	62
2.3.1.3 Métodos baseado em inteligência artificial - redes neurais artificiais.....	63
2.3.1.3.1 Redes feedforward.....	65
2.3.1.3.1.1 Multi-Layer Perceptron (MPL).....	66
2.3.1.3.2 Redes Recorrentes.....	68
2.3.1.3.2.1 Rede de Elman.....	69
2.3.1.3.2.2 Rede de Jordan.....	71
2.3.1.3.3 Funções de ativação.....	72
2.3.1.3.4 Algoritmos de treinamento.....	74
2.4 REDES NEURAIAS COMO MODELO DE ANALISADORES VIRTUAIS EMPREGADOS EM COLUNAS DE DESTILA- ÇÃO.....	80
2.5 CONCLUSÕES.....	86
<b>Capítulo III – MATERIAL E MÉTODOS.....</b>	<b>87</b>
3.1 INTRODUÇÃO.....	88
3.2 DESCRIÇÃO DA UNIDADE EXPERIMENTAL E TESTES.....	88
3.2.1 Instrumentação.....	93
3.2.2 Sistema de controle.....	94

3.2.3	Condições experimentais e operacionais.....	97
3.3	METODOLOGIA PARA IDENTIFICAÇÃO DE SISTEMAS.....	97
3.3.1	Obtenção do grupo de dados de treinamento.....	98
3.3.1.1	Software comercial usado para construção do modelo.....	98
3.3.1.2	Parâmetros usados nas simulações.....	100
3.3.1.3	Sistema de controle da unidade simulada.....	102
3.3.1.4	Validação do modelo.....	105
3.3.1.5	Faixa de operação e perturbações realizadas.....	106
3.3.1.6	Intervalo entre as aplicações das perturbações.....	109
3.3.1.7	Dimensão do conjunto de dados entrada/saída.....	109
3.3.1.8	Tratamento dos dados.....	110
3.3.2	– Determinação da Estrutura da rede neural.....	110
3.3.2.1	Seleção das variáveis de entrada/saída.....	110
3.3.2.1.1	Variáveis de entrada.....	111
3.3.2.1.2	Variáveis de saída.....	113
3.3.2.2	Ambiente de desenvolvimento da rede neural.....	114
3.3.2.3	Configuração da rede neural.....	114
3.3.2.3.1	Tipos de redes neurais avaliadas.....	115
3.3.2.3.2	Métodos e algoritmos de treinamento avaliados.....	116
3.3.2.3.3	Funções de transferência avaliadas.....	117
3.3.2.3.4	Número de camadas e neurônios.....	117
3.3.2.3.5	Procedimento de treinamento, inicialização e parada.....	119
3.3.3	Validação da rede – Medidas de desempenho.....	122
3.4	CONCLUSÕES.....	123

## **Capítulo IV – RESULTADOS E DISCUSSÃO.....125**

4.1	SIMULADOR EM MODO ESTÁTICO.....	126
4.1.1	Análise de uma simulação padrão.....	127
4.1.2	Avaliação do efeito da alteração nas condições de alimentação da coluna.....	129
4.2	SIMULADOR EM MODO DINÂMICO.....	132
4.2.1	Validação das simulações dinâmicas e análise do comportamento do processo em malha fechada.....	134
4.2.1.1	Perturbação na temperatura de alimentação.....	137
4.2.1.2	Perturbação na vazão de alimentação.....	148

4.2.1.3	Avaliação final em relação a validação das simulações dinâmicas.....	157
4.2.2	– Comportamento do processo em malha aberta.....	158
4.2.2.1	Perturbação degrau negativo (-50%) na composição de etanol na corrente de alimentação.....	160
4.2.2.2	Perturbação degrau positivo (+50%) na composição de etanol na corrente de alimentação.....	163
4.2.2.3	Perturbação degrau positivo (+100%) no calor cedido à resistência do prato 2.....	165
4.2.2.4	Avaliação da dinâmica em malha aberta após aplicadas perturbações.....	167
4.2.2.5	Determinação do intervalo máximo entre as perturbações no simulador na etapa de obtenção dos dados de treinamento.....	168
4.2.2.6	Determinação do tempo de amostragem.....	170
4.2.2.7	Análise de sensibilidade.....	172
4.3	– MODELO EMPÍRICO DINÂMICO BASEADO EM REDE NEURAL MALHA ABERTA.....	174
4.3.1	Determinação da função de transferência mais adaptada ao conjunto de dados para redes neurais <i>Multi-Layer-Perceptron</i> (MPL).....	174
4.3.1.1	Medidas de desempenho das redes MPL avaliadas – Etapa de treinamento.....	176
4.3.1.2	Medidas de desempenho das redes MPL avaliadas – Etapa da validação cruzada.....	181
4.3.2	Determinação do algoritmo de treinamento mais adaptado ao conjunto de dados para redes neurais <i>Multi-Layer-Perceptron</i> (MPL).....	185
4.3.2.1	Medidas de desempenho das redes MPL avaliadas – Etapa de treinamento.....	186
4.3.2.2	Medidas de desempenho das redes MPL avaliadas – Etapa da validação cruzada.....	193
4.3.3	Redes recorrentes de Elman.....	195
4.3.3.1	Medidas de desempenho das redes recorrentes de Elman avaliadas – Etapa de treinamento e validação cruzada.....	196
4.3.4	Comparativo entre Redes <i>Multi-Layer-Perceptron</i> (MPL) e Redes Recorrentes de Elman .....	200
4.3.5	Avaliação final em relação ao modelo empírico obtido para ser empregado como analisador virtual em malha aberta.....	206

4.4	MODELO EMPÍRICO DINÂMICO BASEADO EM REDE NEURAL – MALHA FECHADA.....	207
4.4.1	Redes neurais <i>Multi-Layer-Perceptron</i> (MPL) e redes recorrentes de Elman.....	209
4.4.1.1	Medidas de desempenho das redes avaliadas – Etapa de treinamento e validação cruzada.....	210
4.4.2	Avaliação final em relação ao modelo empírico obtido para ser empregado como analisador virtual em malha fechada.....	219
	<b>Capítulo V – CONCLUSÕES.....</b>	<b>221</b>
5.1	CONCLUSÕES.....	221
	<b>Capítulo VI – SUGESTÕES PARA TRABALHOS FUTUROS...229</b>	
6.2	SUGESTÕES PARA TRABALHOS FUTUROS.....	227
	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>229</b>
	<b>APÊNDICE .....</b>	<b>255</b>
	<b>Apêndice A – Códigos da rede neural.....</b>	<b>255</b>

# Capítulo I

## INTRODUÇÃO

Em processos de destilação, a qualidade dos produtos depende de um sistema de controle bem ajustado, o qual por sua vez, depende de medições de algumas variáveis do processo, uma delas a composição de saída das colunas. Infelizmente a medição precisa e rápida da composição dos produtos em uma torre de destilação é um problema antigo e típico na indústria, conforme afirmado por vários pesquisadores (QUINTERO-MARMOL e LUYBEN, 1992; BARATTI *et al.*, 1995; KANO *et al.*, 2000; BAHAR e ÖZGEN, 2008). Para avaliação da qualidade de um produto geralmente são utilizadas análises de laboratório e analisadores de processo. Porém, as deficiências desses métodos são o elevado tempo para obtenção das respostas e a baixa confiabilidade, respectivamente. Além de que as análises de laboratório não podem ser utilizadas para controle *on-line* e são empregadas apenas para ajustes e correções. A maioria das colunas industriais não dispõe de analisadores em linha para medir a composição da carga ou dos produtos, em função da complexidade e do elevado custo, tanto de investimento quanto de manutenção. Segundo Meneguelo (2007) uma situação comum no meio industrial é a existência de analisadores de processo apenas nas unidades de maior motivação econômica.

A utilização de sensores por software substituem a medição de variáveis que necessitariam de um analisador em linha. Desta forma, um grande desafio é obter um modelo preciso do processo, uma vez que as medições experimentais serão substituídas por dados gerados por esse modelo. Os sensores possibilitam o controle direto e constante da qualidade do produto ou corrente desejada, além de deslocar parte dos custos de implantação e manutenção de analisadores para engenharia (MENEGUELO, 2007).

Existem na literatura vários trabalhos (MEJDELL e ANDERSSON, 1993; SZIGETI, *et al.*, 2000; ZAMPROGNA *et al.*, 2004; ZANATA, 2005; SINGH *et al.*, 2005; PATIL e NIGAM, 2009) que tratam este problema, propondo o uso de modelos inferenciais para estimar a composição. Alguns fazem a inferência direta da composição através de medidas de temperatura (sistemas binários), outros empregam modelos matemáticos fenomenológicos ou empíricos para a predição. Porém, na grande maioria trata-se apenas de estudos simulados e em colunas convencionais (com ação de controle na base e topo).

Diferentes estratégias têm sido propostas para desenvolver sensores virtuais. De acordo com FORTUNA *et al.* (2007), são três as principais abordagens utilizadas para o desenvolvimento destes sensores: modelagem fenomenológica, estatística multivariada e modelagem por inteligência artificial (redes neurais, lógica *fuzzy* e métodos híbridos).

As estratégias de controle baseadas em modelos fenomenológicos não-lineares de processo, também conhecida como modelagem caixa branca, requerem um vasto conhecimento das leis físicas. Como as plantas usualmente envolvem fenômenos que se identificam melhor com modelos não lineares e de maior complexidade, o desenvolvimento do modelo pode requerer muito tempo e esforço de programação e implementação e, em muitos casos, o tempo de processamento para obtenção da informação de interesse pode comprometer a eficiência do sistema de controle. Estes motivos, são os maiores obstáculos à aplicação destes métodos em ambientes industriais.

Além disso, com os atuais avanços nos sistemas de automação industrial, um grande número de variáveis de processo podem ser simultaneamente medidas e armazenadas, podendo ser facilmente acessadas, o que encoraja a construção de modelos a partir de dados históricos, já que o registro histórico acumulado representa uma fonte muito útil de informação, que permite identificar características relevantes do processo (FORTUNA *et al.*, 2007). Todavia, é importante enfatizar que estas abordagens são usadas para executar uma mesma função: transformar medidas disponíveis na planta em outras medidas não disponíveis, com auxílio de uma transformação matemática. Desta forma, modelos baseados em redes neurais apresentam vantagem de necessitarem apenas da aquisição de um conjunto de dados suficientemente grande para representar o processo, que serão usados para treinamento da rede.

Este trabalho propõe um estudo no qual se pretende desenvolver um sensor por software baseado em redes neurais artificiais para inferir o valor de composição do produto de topo de uma coluna de destilação piloto, com ações distribuídas de calor ao longo dos pratos.

O grande diferencial deste projeto é o fato do modelo empírico a ser desenvolvido ter a capacidade de modelar um processo com configuração diferenciada baseada em ação distribuída de controle com aquecimentos em pontos intermediários. Esta configuração já se encontra implementada em uma unidade piloto experimental (MARANGONI, 2005) e foi estudada por Werle (2007), o qual realizou a análise energética, avaliando também a minimização de transientes.

Porém ainda não foi obtido um modelo representativo com tais características. Este modelo poderia auxiliar novas pesquisas em projetos de controladores e estratégias avançadas de controle na busca da otimização do processo. O objetivo principal desta tese vem a esse encontro, buscando sanar esta dificuldade e obter um modelo empírico da unidade agregando conhecimentos à nova abordagem e, através deste, fornecer ferramentas para futuramente aprimorar a forma de controle da unidade.

O fato da unidade possuir ações distribuídas de calor ao longo dos pratos lhe confere um controle diferenciado das colunas convencionais. Com a obtenção do modelo deste processo, pretende-se empregá-lo como um sensor por software para estimar a composição do topo da coluna. A proposta é utilizar a composição como variável controlada *on-line*. Busca-se obter um sensor robusto, porém com respostas rápidas, com potencial para o desenvolvimento de estratégias de controle inferenciais avançadas.

Desta forma, agregando estas duas concepções: ação distribuída e uso de *soft-sensor*, pretende-se desenvolver um ambiente inovador de controle aplicado nesse tipo de processo. Uma vez obtido o modelo inferencial, este pode ser utilizado no controle em malha fechada, visto que eles fornecem uma maneira para aumentar a robustez da malha de controle. Neste caso, o resultado da inferência do produto de topo pode ser usado diretamente pelo controlador, ocasionando a atuação rápida sobre uma ou mais válvulas de controle, ou para manipular o *set point* de uma ou mais temperaturas controladas, com um controle em cascata paralelo de forma a corrigir uma propriedade fora da especificação.

Neste projeto será apresentada, inicialmente, a elaboração de um modelo da coluna em um simulador comercial, com o intuito de gerar o conjunto de dados necessários para treinamento e geração de padrões para as redes neurais. Esta etapa foi necessária em virtude da quantidade insuficiente de dados experimentais dinâmicos dos perfis de temperatura e composição. Em seguida será contemplada a elaboração do modelo empírico do processo, onde foram testadas inúmeras configurações e arquiteturas de dois diferentes tipos de redes neurais: *feedforward* e redes recorrentes. Estes testes visaram encontrar uma rede com boa capacidade de generalização e conseqüente boa predição da propriedade alvo, para ser empregada como sensor por software, foco do estudo. Não será contemplada a implementação final do sensor na unidade, apenas será apresentado o modelo empírico com potencial para tal.

Neste capítulo será apresentada uma introdução ao tema, expondo o problema do controle da composição dos produtos em tempo real no processo de destilação e as dificuldades e alternativas para minimizá-lo. A organização do trabalho também será detalhada.

## 1.1 IMPORTÂNCIA DO ESTUDO

A destilação é um dos processos de separação mais empregados na indústria química e petroquímica. A especificação dos produtos segue a demanda do mercado e o objetivo econômico destas unidades consiste em manter os produtos mais próximos possível das condições desejadas.

Essa não é uma tarefa fácil, pois o processo apresenta várias características indesejáveis, porém intrínsecas ao mesmo, tais como: elevado tempo morto e elevadas constantes de tempo, além de inúmeras restrições de operação. O comportamento dinâmico é extremamente não linear e existe um elevado grau de acoplamento entre as variáveis. Trata-se de um processo com respostas lentas e presença de atraso em virtude de sua geometria em estágios (HENLEY e SEADER, 1981). Estas características impõem transientes de operação quando o processo é perturbado. Ainda, em alguns casos específicos, existem fatores associados ao processo que dificultam o controle da unidade.

Um sistema de controle bem projetado e ajustado não é suficiente para eliminar estes transientes de operação. Um aspecto que contribui para tal situação consiste na centralização do controle nas variáveis da base e do topo da coluna. Com isso existe a propagação da ação corretiva por toda a unidade, gerando um período com produção fora da especificação desejada. Mesmo com toda a tecnologia envolvida e com sistemas de controle bem ajustados, ainda torna-se necessário implementar uma abordagem adequada à dinâmica do processo.

Logo, baseado na distribuição do calor na destilação aplicada em colunas diabáticas a qual se mostra como uma opção importante para reduzir as perdas de energia (SCHALLER *et al.*, 2001; BJÖRN *et al.*, 2002; KOEIJER e RIVERO, 2003; KOEIJER *et al.*, 2004; SALAMON *et al.*, 2004; ROSJORDE e KJELSTRUP, 2005; SHU *et al.*, 2007) e buscando-se controladores adequados à dinâmica do processo, propôs-se o emprego da abordagem de controle com aquecimento nos pratos de uma coluna (MARANGONI, 2005). Em conjunto com esta abordagem inovadora, a aplicação de controle inferencial, para estimar a composição aparece como uma alternativa promissora. Com isso, acredita-se que o resultado seja a redução de transientes de operação e

de consumo energético da unidade, decorrentes da ação distribuída, além da redução de produtos fora da especificação, devido a ação mais rápida e precisa do controle inferencial.

Entretanto, para minimizar estes transientes é necessário primeiramente uma otimização da produção. Para tal, é necessário o acompanhamento e o controle de algumas propriedades do processo. A medição confiável e precisa da composição, ou de algumas dessas propriedades dos produtos é uma das principais dificuldades em torres de destilação. Para a determinação destas, normalmente são empregadas análises de laboratório ou analisadores de processo. Contudo, conforme citado anteriormente, existem diversos problemas decorrentes desses métodos, que os tornam inadequados para a aplicação no controle em malha fechada. Entre esses problemas estão: a demora entre a coleta da amostra e a disponibilização do resultado da análise; e a baixa confiabilidade dos analisadores de processo. Normalmente, esses valores são obtidos por analisadores *of-line*, sendo que os resultados destas análises podem levar cerca de 15 a 30 minutos (MENEGUELO, 2007).

Uma alternativa mais eficiente é através da estimação da propriedade do produto, ou seja, os modelos inferenciais. Uma vez obtidos os modelos inferenciais, estes podem ser utilizados no controle em malha fechada. Neste caso, o resultado da inferência é usado diretamente pelo controlador, ocasionando a atuação rápida sobre uma ou mais válvulas de controle, ou para manipular o *set point* de uma temperatura controlada, com um controle em cascata paralelo de forma a corrigir uma propriedade fora da especificação. Este é o conceito que rege o controle inferencial desenvolvido por Joseph e Brosilow (1978).

Esta abordagem pode trazer uma série de vantagens em relação às análises de laboratório e aos analisadores de processo, dentre elas: resultados mais freqüentes que as análises de laboratório ou de alguns tipos de analisadores de processo; o controle baseado em modelo inferencial é mais rígido, produzindo um menor desvio padrão da propriedade inferida (PINTO e LIMA, 1993).

A vinculação dessas medidas virtuais com estratégias de controle avançado consiste em nobre benefício para a operação de processos não-lineares, tais como a destilação.

A inferência de propriedades em função de variáveis medidas no processo é uma alternativa que traz vantagens diretas, tais como: boa precisão, baixo tempo de resposta e baixo custo de manutenção, quando comparados com analisadores de processo. O enfoque principal do estudo efetuado é a obtenção de modelo empírico do processo de destilação, com diferencial de ações distribuídas no fornecimento de

calor à coluna. Este modelo será baseado em redes neurais artificiais o qual será empregado em um sensor por software. Alguns métodos tradicionalmente utilizados para obtenção de modelos inferenciais serão avaliados brevemente no capítulo da revisão bibliográfica.

Devido a sua importância, o cenário de aplicações dos sensores virtuais é cada vez mais amplo, sendo aplicados em distintos segmentos da indústria, como por exemplo, polímeros onde a qualidade do produto é especificada principalmente em termos de propriedades difíceis de medir e geralmente indisponíveis em tempo real (LIU, 2007; FACCO *et al.* 2008), na área de biotecnologia (KERMANI *et al.*, 2005; DESAI *et al.*, 2006; DAI *et al.*, 2006) e na indústria de celulose (DUFOUR *et al.*, 2005). Finalmente a indústria química e petroquímica, uma das pioneiras na implementação de estratégias baseadas em sensores virtuais (ZAMPROGNA *et al.*, 2001a; FORTUNA *et al.*, 2007; BOLF *et al.*, 2008; PATIL e NIGAM, 2009). Neste segmento as Redes Neurais Artificiais (RNA's) têm sido amplamente aplicadas como modelo empírico para o sensor, em diversas áreas. Na área de inferência de propriedades, as RNA's permitem a obtenção de modelos de inferência abrangentes e precisos, podendo ser utilizadas no controle em tempo real, nos arranjos de controle em malha simples ou em arquiteturas de controle multivariável.

## 1.2 ORGANIZAÇÃO DO TRABALHO

Para facilitar a apresentação, este trabalho é dividido em 6 (seis) capítulos.

Iniciou-se com o Capítulo I, apresentando uma introdução em relação ao tema da proposta e a importância de seu desenvolvimento no cenário atual. Também são contemplados os objetivos e o desafio do projeto.

No Capítulo II é apresentada a revisão bibliográfica tratando dos principais conceitos relacionados ao tema, com ênfase para modelagem inferencial empregando redes neurais artificiais e sua utilização como sensor por *software*.

No Capítulo III é relatada a metodologia aplicada. Neste é feito uma descrição e caracterização da unidade piloto indicando as rotinas analíticas e condições operacionais que foram utilizadas nos experimentos. Em seguida é descrita a metodologia adotada para identificação de sistemas para emprego de redes neurais. A metodologia envolve desde o desenvolvimento de um modelo dinâmico da coluna de destilação, através do software de simulação comercial HYSYS®

empregado para gerar os dados utilizados como entradas/saída que serão utilizados para treinamento das redes avaliadas, até as ferramentas utilizadas para determinação da configuração das redes neurais empregadas neste trabalho.

No Capítulo IV, inicialmente são apresentados os resultados obtidos através das simulações estáticas e dinâmicas e a comparação dos resultados simulados com os dados experimentais de perfis de temperatura, com o intuito de validar as simulações com a unidade simulada operando em malha fechada, além de um breve estudo avaliando conceitualmente as respostas das simulações, que também servirá de validação das mesmas. Posteriormente, com o emprego do simulador dinâmico em malha aberta, é realizada a caracterização da dinâmica, através de aplicações de degraus na carga processada. Esta etapa fornecerá dados para seleção das variáveis de entrada/saída do sensor.

Na sequência o foco se voltará para a concepção de redes neurais capazes de realizar a inferência da composição de topo da coluna com maior precisão, tanto em malha aberta, quanto em malha fechada. Foram avaliados dois tipos de redes neurais dinâmicas: Redes *feedforward*, mais especificamente do tipo MLP, e redes recorrentes de Elman, para verificar qual obtêm melhor performance na inferência da composição de produto de topo e que tenha potencial para ser empregada como sensor por software.

E para finalizar, no Capítulo V apresentam-se as conclusões da tese e no Capítulo VI as sugestões para trabalhos futuros.

### 1.3 OBJETIVOS

- Desenvolver um modelo empírico dinâmico não-linear baseado em redes neurais artificiais para ser empregado como sensor por *software*, com capacidade para realizar a estimação da composição de etanol no topo de uma coluna de destilação. O modelo desenvolvido deve ser representativo da característica principal da unidade que é o aquecimento distribuído nos pratos. A informação da predição da rede neural deverá ser empregada posteriormente para projeto de controladores com técnicas avançadas, buscando a otimização na unidade piloto de destilação.

### 1.3.1 Objetivos Específicos

Além de desenvolver o sensor por software com modelagem empírica do processo em malha aberta, que consistiu no objetivo principal do projeto, também se optou por fazer uma modelagem do processo em malha fechada, com o intuito de obter um sensor virtual para fins de monitoramento, com as atuais malhas de controle. Para satisfazer ambas condições, são destacados os objetivos específicos a seguir:

Em relação ao simulador para geração do conjunto de dados:

- Definir o software a ser empregado como simulador;
- Desenvolver modelo da coluna de destilação piloto;
- Validar experimentalmente o modelo simulado com dados experimentais da unidade piloto.

Em relação à rede neural:

- Definir o tipo de rede neural a ser empregada para desenvolvimento do modelo;
- Definir a função de transferência, o método de otimização e o algoritmo de treinamento adequados para o conjunto de dados e treinamento da rede;
- Avaliar o desempenho de distintos algoritmos de treinamento frente ao mesmo conjunto de dados e analisar seu comportamento, de forma a contribuir com as pesquisas nesta área;
- Selecionar as redes que possuam boa capacidade de generalização e arquitetura suficientemente simples para servirem de modelo ao sensor;
- Obter uma rede neural artificial capaz de prever adequadamente a composição de topo da coluna para uma ampla faixa de operação, com potencial para ser empregada como sensor por software.

Em relação ao sensor por *software*:

- Desenvolver um sensor virtual, baseado em rede neural, representativo do processo com configuração distribuída de controle, para estimativa da composição de produto de topo, capaz de ser implementado *on-line*;

Em relação à unidade:

- Verificar mudanças necessárias na planta para implementação futura do sensor virtual.

# Capítulo II

## REVISÃO BIBLIOGRÁFICA

Neste capítulo a proposta é descrever brevemente características do processo de destilação. O controle destas unidades, fatores que geram os transientes de operação e as dificuldades do controle de composição. Enfatizou-se o emprego de modelos inferenciais para o controle do processo. Serão apresentadas as principais técnicas usadas para obtenção de modelos empíricos, com destaque para as redes neurais artificiais, que serão utilizadas no desenvolvimento do *soft-sensor*. Será exposta também uma revisão apresentando trabalhos que envolvem a aplicação de sensores por software para estimativas de propriedades em alguns processos.

### 2.1 INTRODUÇÃO

Um dos principais objetivos das indústrias petroquímicas em geral é operar a planta de forma que o produto final atenda aos padrões de qualidade especificados sob o menor custo possível. Para tal, se torna necessário a implementação de estratégias de monitoramento e controle capazes de realizar diversas operações, tais como otimização da produção, monitoramento, previsão e detecção de falhas. Porém, a dificuldade no desenvolvimento destas estratégias é diretamente proporcional à complexidade da operação em questão.

Outro fator de extrema relevância, é que geralmente os índices de qualidade dos produtos não são medidos em tempo real, devido aos custos expressivos dos equipamentos de medição *on-line* (LONGUI *et al.*, 2009), cujas técnicas de medição costumam ser lentas e complicadas e com baixa confiabilidade (MENEGUELO, 2007).

Buscando contornar este problema, várias pesquisas são realizadas em torno do desenvolvimento de técnicas alternativas capazes de reduzir o custo de sensores *on-line*. Geralmente são baseadas em algoritmos que simulam o comportamento dos sensores físicos, disponibilizando em tempo real estimativas das variáveis primárias a partir de variáveis secundárias e são conhecidos como modelos inferenciais ou sensores por software.

Desta forma, a modelagem inferencial, aplicada às torres de destilação, pode ser utilizada juntamente com os sistemas digitais de controle e os controladores multivariáveis buscando amenizar estes problemas e encontrar soluções viáveis para sua melhoria e consequente obtenção de produtos com maior qualidade.

## 2.2 PROCESSO DE DESTILAÇÃO

Destilação é um dos processos mais importante de separação física, sendo utilizada em 95% das separações de fluidos na indústria química (PONCE, 2011), sendo considerado um dos processos básicos de produção em uma refinaria. Segundo Jesus *et al.*(2003), na maioria das indústrias petroquímicas, 80% do custo operacional energético é devido a esta operação unitária. Além disso, a eficiência global termodinâmica de uma destilação convencional é por volta de 5-20% (HUMPHREY *et al.*, 1991; DE KOEIJER e KJELSTRUP, 2002).

Mesmo com a importância da destilação como processo de separação, observa-se nela uma grande degradação da energia, que ocorre devido às grandes diferenças de temperatura no condensador e no refeedor da coluna (PONCE, 2011). O grande consumo energético concentrado numa única operação industrial está atraindo pesquisadores e empresas da área e vêm impulsionando diversos estudos, resultando em novas configurações para o processo de destilação. Dentre as diversas configurações propostas, podem ser citadas com maior importância a Recompressão Direta de Vapor - VRC (OLIVEIRA *et al.*, 2001; ENWEREMADU *et al.*, 2009; SUJIT e DAOUTIDIS, 2009), Destilação Diabática (KOEIJER *et al.*,2004; SALAMON *et al.*, 2004; ROSJORDE e KJELSTRUP, 2005; SHU *et al.*, 2007; JANA, 2010), Destilação Múltiplo-Efeito (SOPHOS *et al.*, 1978; PINTO, 1987; DE RIJKE, 2007; WEI-ZHONG e XI-GANG, 2009; TSUN-JEN HO, 2010), Colunas de Parede Dividida - DWC (PREMKUMAR e RANGAIAH, 2009; RONG, 2011), e mais recente, a tecnologia de Auto-Recuperação de Calor (*Self-Heat Recuperation Technology*- SHRT) (KANSHA *et al.*, 2010; MATSUDA *et al.*, 2011), e a Coluna de Destilação com Integração Interna de Calor (CDIIC) sendo mais conhecido na literatura como Coluna HIDiC (*Heat Integrated Distillation Column*) (HO, *et al.*, 2009; SUPLANIT, 2010; SUPLANIT, 2011; PULIDO, 2011) sendo esta a de maior relevância atualmente. Todas estas tecnologias tem como objetivo principal o aproveitamento de calor na coluna.

No processo de destilação convencional, além das perdas com energia, existem transientes de operação elevados em virtude das características intrínsecas do mesmo, tais como o comportamento dinâmico não linear e um elevado grau de acoplamento entre as variáveis. Um sistema de controle bem projetado e ajustado não é suficiente para eliminar estes transientes de operação. Outro aspecto que contribui para tal situação consiste na centralização do controle nas variáveis da base e do topo da coluna. Com isso existe a propagação da ação corretiva por toda a unidade gerando um período com produção fora da especificação desejada. Algumas das principais causas que geram estes transientes serão abordadas na *seção 2.2.2* deste capítulo. Logo, a destilação necessita de uma estratégia de controle adequada, de forma a melhorar a qualidade dos produtos e garantir maior economicidade.

O presente trabalho aborda estas duas características referenciadas do processo, baixa eficiência termodinâmica com elevados transientes de operação. Nele é proposto o uso de sistema inferencial de controle baseado em tecnologia neural para estimar a composição de topo da coluna, aliada a uma nova configuração do processo, com controle com aquecimentos individuais nos pratos da coluna (MARANGONI, 2005), proposta esta baseada na distribuição do calor na destilação aplicada em colunas diabáticas. Estas ações se mostram com grande potencial para reduzir as perdas de energia, além da redução de produtos fora da especificação, devido à ação mais rápida e precisa do controle inferencial.

### 2.2.1 Controle de colunas de destilação

O controle de colunas de destilação é amplamente estudado (LUYBEN, 1989; SKOGESTAD, 2007a). Este processo de separação permite a implementação de técnicas convencionais à aplicação de algoritmos avançados.

A destilação fracionada é um problema essencialmente multivariável e com restrições de operação. Trata-se de um processo com respostas lentas e presença de atraso (HENLEY e SEADER, 1981). Pode ser representado por um processo acoplado em série em virtude da separação ser realizada em estágios, e cada estágio representar um ponto de equilíbrio; portanto, o tempo morto é incrementado a cada estágio da coluna gerando o atraso hidráulico. Quando o valor desta variável cresce em relação à constante de tempo, existe a possibilidade do controlador

não oferecer um bom desempenho. Estas características impõem transientes de operação quando o processo é perturbado. Ainda, em alguns casos específicos, existem fatores associados ao processo que dificultam o controle da unidade.

Colunas de destilação necessitam de sistemas de controle bem ajustados para se obter a especificação desejada do produto e uma operação estável no processo. Segundo Abdulah *et al.* (2007) o controle efetivo das colunas de destilação pode melhorar a qualidade de seus produtos, aumentar a capacidade de produção e reduzir o consumo de energia, além de auxiliar na segurança do processo e dos operadores. Definir uma estratégia de controle a ser implementada é fundamental. Nesta etapa são especificadas as variáveis a serem controladas, as variáveis a serem manipuladas e toda a instrumentação necessária. O sucesso de qualquer estratégia de controle por realimentação depende da confiabilidade e do nível de exatidão das medidas das variáveis controladas do processo (LINHARES, 2010). Em seguida, geralmente é feita a otimização da planta de produção. Para tal lança-se mão de técnicas cada vez mais avançadas referentes à otimização, supervisão e controle de processos. De acordo com Longhi *et al.* (2009), a otimização da operação deste equipamento é uma meta constante nas indústrias petroquímicas, pois ela possibilita retornos financeiros bastante atrativos. Nestas indústrias as técnicas de controle avançado têm sido amplamente empregadas para esta finalidade e, por consequência, na obtenção dos produtos finais dentro de padrões de qualidade especificados por normas (WANG *et al.*, 2006).

Elevadas capacidades de produção demandam equipamentos de grande porte. O processo de destilação possui características, como o processamento de altas vazões, presença de elevados gradientes de temperatura e tempo morto, fortes interações e um alto grau de acoplamento entre as variáveis manipuladas e controladas (MOURA, 2003). O comportamento dinâmico desta operação é bastante lento, o que aumenta a importância de um sistema de controle apropriado e que minimize os estados transientes oriundos de alterações entre condições de operação, seja por uma mudança desejada dessas condições, seja pela incidência de perturbações imprevistas.

### 2.2.2 Transientes de operação: principais causas

A formação de transientes em uma coluna de destilação ocorre quando o processo é perturbado e as características deste limitam a

eficiência do sistema de controle ou quando um fator externo induz a modificação do ponto de operação da unidade.

Colunas de destilação possuem algumas características inerentes ao processo, um tanto quanto “indesejáveis”, como *elevadas constantes de tempo* que geram atrasos na resposta, *as não-linearidades, restrições e acoplamento das variáveis*. Tais características específicas interferem diretamente nos transientes gerados quando a unidade é perturbada. Diante destas perturbações, a eficiência dos sistemas de controle destas unidades é prejudicada.

A formação de longos transientes pode ser acarretada em virtude destas características indesejadas ou também em função do *sistema a ser destilado, das mudanças na composição da alimentação e/ou das transições de operação* necessárias em virtude de alterações no mercado (MARANGONI, 2005).

A dinâmica de processos de destilação é lenta. Embora o atraso e a constante de tempo sejam pequenos para cada prato (na ordem de segundos), o comportamento dinâmico da coluna toda é dominado por grandes constantes de tempo (VENTIN, 2010), conforme mostrado nos estudos de Wah e Harriot (1970) e Kim e Friedly (1974), na ordem de horas para algumas unidades. Pode-se observar que existe o atraso hidráulico encontrado na transferência de líquido prato a prato, além do atraso na resposta de uma variável a perturbações. Este fato pode ser verificado quando ocorrem mudanças na composição da alimentação. Esta perturbação provoca mudanças primárias nas composições dos pratos. Por conseqüência, as correntes líquidas e vapor no interior da coluna são alteradas e mudanças secundárias nas composições são observadas. Quando ocorre uma perturbação este equilíbrio é afetado e com isso gera um efeito em série em cada estágio. Este comportamento é importante e deve ser considerado na decisão do sistema de controle.

Enquanto não ocorre a estabilização das composições nos pratos, o novo estado estacionário não será alcançado. O que se observa ao final é o incremento do tempo morto em virtude da separação em estágios. Além disso, ações na base devem superar o tempo morto de toda a coluna para refletirem mudanças no topo, verificando-se o mesmo na situação contrária. Este fator pode ser considerado um forte responsável pela geração de transientes em colunas de destilação.

O comportamento não linear das colunas de destilação muitas vezes está associado ao acoplamento das variáveis controladas e manipuladas, ao equilíbrio líquido-vapor e as características dos equipamentos (ANSARI e TADÉ, 2000). Observam-se ainda características não lineares decorrentes do tipo de mistura a ser

processada, por exemplo, mistura azeotrópica; no funcionamento dos sensores e atuadores que compõe a instrumentação; na perda de carga gerada pela construção das tubulações, entre outros.

As restrições referentes às limitações hidráulicas, de separação e de transferência de calor também atuam como não linearidades no processo de destilação. Temperatura e pressão não podem se aproximar das condições críticas, pois o fluxo hidráulico depende da diferença entre a massa específica da fase líquida e vapor.

Colunas de destilação apresentam forte grau de acoplamento entre suas variáveis, característica do próprio processo, ou seja, uma ação tomada em certo ponto da coluna irá ou poderá afetar o controle de outra variável em outra malha. Interações entre dois controladores podem ser minimizadas parcialmente através de um bom ajuste do controlador. A escolha do sistema de controle destas unidades selecionando as variáveis manipuladas e/ou pares de variáveis controladas/variáveis manipuladas é fundamental para obtenção de um maior ou menor grau de acoplamento (MORAES, 2004).

Assim, os aspectos da dinâmica do processo interferem na eficiência dos sistemas de controle. Isto resulta facilmente em produtos fora da especificação provenientes do transiente gerado até que se obtenha uma resposta satisfatória quando o processo é perturbado.

### 2.2.3 Dificuldades do controle de composição e alternativas

Mudanças na composição de carga representam o mais significativo distúrbio que o sistema de controle da coluna deve rejeitar durante uma operação contínua (MORAES, 2004). Este tipo de perturbação irá deslocar o perfil de composição da coluna, resultando em uma grande perturbação na composição dos produtos. A maioria das colunas industriais não dispõe de analisadores em linha para medir a composição na carga. Portanto, esta perturbação usualmente aparece como uma perturbação não medida.

Um problema muito comum enfrentado em colunas de destilação de refinarias de petróleo consiste em mudanças no tipo de petróleo processado. Uma vez que cada óleo possui características dependentes de sua origem, este fato propaga o transiente gerado para toda a refinaria. O sistema de controle deve ser capaz de ajustar novos valores de referência de acordo com tais características. Como alternativa ao impacto da composição do óleo, as refinarias fazem uso

de tanques de estocagem para a mistura de diferentes óleos, conforme citado por Song (2002) e ilustrado na Figura 2.1.

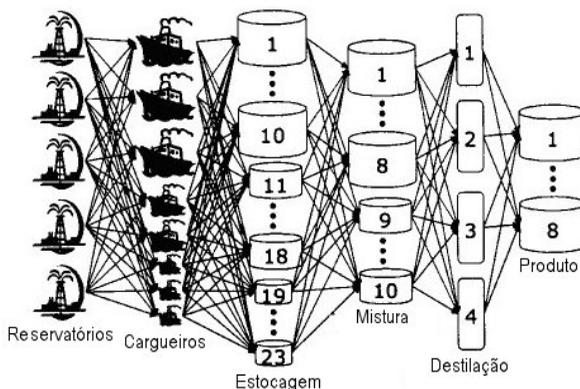


Figura 2.1 Ilustração do caminho percorrido pelo óleo cru até o refino

Fonte: Song, (2002).

Esta alternativa muito empregada industrialmente de misturar óleos de procedência diferentes, busca minimizar o impacto da diferença de composição na alimentação do refino. No entanto, a adaptação do controle das unidades atmosféricas às propriedades químicas e físicas do petróleo a ser refinado ainda gera transientes consideráveis. O ajuste de novas condições de operação geram em torno de 4 (quatro) horas até a estabilização em novo ponto de operação de uma refinaria de porte médio (NASCIMENTO, 1994). Em muitos casos a refinaria opta pela mudança de campanha. Esta consiste no direcionamento da produção para alguns tipos de derivados com vantagens econômicas quando aplicados como fontes de combustíveis ou lubrificantes.

Em todos os casos, o que se busca atingir é a especificação dos produtos. Para tal é necessário manter sob controle determinadas propriedades ou características dos mesmos. Para avaliação da qualidade de um produto podem ser utilizadas análises de laboratório e analisadores de processo. Existem diversas técnicas de análise de laboratório (EWING, 1972) como, por exemplo: a cromatografia, a espectrometria molecular, espectrometria atômica e a destilação segundo a norma ASTM (*American Society for Testing Materials*). As análises de laboratório podem ser feitas rotineiramente (diariamente, semanalmente, etc.) ou por demanda.

Uma deficiência desse método é que o operador só efetua uma correção após o resultado da análise estar disponível, que pode levar de 15 a 30 minutos, segundo Meneguelo (2007). Nesse ínterim, muito produto já terá sido produzido fora da especificação. Além disso, para avaliar as conseqüências de uma ação do operador, é necessária uma nova análise.

Já os analisadores de processo são equipamentos instalados junto às tubulações onde escoam os produtos. Geralmente são complexos e caros, tanto para aquisição quanto para manutenção (KANO *et al.*, 2000; OISIOVICI e CRUZ, 2000; VENKATESWARLU e AVANTIKA, 2001; FORTUNA *et al.*, 2005; LONGUI *et al.*, 2009), normalmente compostos de um módulo processador de sinal e um módulo analítico. Podem ser utilizados no sistema de controle da torre, pois suas análises levam de alguns segundos a dezenas de minutos. Porém, seus resultados podem não ser satisfatórios. São utilizados, principalmente, onde for necessário um resultado de análise mais imediato e se não for essencial um valor muito preciso. Existem vários tipos de analisadores de processo, tais como: cromatógrafos, viscosímetros, espectrômetros, densímetros, entre outros. Em Passos *et al.* (1988), Santos (2004) e Borges (2004) há uma descrição detalhada de diversos tipos de analisadores de processo. Fieg (2002) faz uso de cromatógrafos em linha para a análise dos produtos de topo de uma coluna de destilação. O autor descreve o alto tempo morto desta técnica e propõem um sistema de controle distribuído (DCS) com o emprego destas medidas.

Diversas técnicas são estudadas para controlar a composição de uma coluna de destilação através de modelos rigorosos e não lineares (BAROLO e BERTO, 1998; FIKAR, 1998; KIM e KIM, 1998; SINHA e RIGGS, 1989; CHATTERJEE e SARAF, 2004). Muitos modelos não lineares formulam o algoritmo como um problema geral de otimização com restrições e o resolvem por programação não linear.

Adetola e Guay (2010) e Hosen *et al.* (2011) destacam que a otimização em tempo real (RTO), que avalia métodos de como melhor operar sistemas sujeitos a distúrbios freqüentes como alterações na composição e vazão de alimentação ou quantidade de energia disponível, pode produzir benefícios econômicos significativos. Ambos empregaram modelos não-lineares preditivos, sendo que o segundo autor empregou uma rede neural artificial baseada em modelo preditivo de controle (NNMPC). Porfírio e Odloak (2011) propuseram a implementação de um modelo preditivo de controle (MPC) em uma coluna de destilação industrial que integrasse o controle e a otimização

econômica. Seus resultados mostram que a abordagem é eficiente e pode ser usada em vários casos práticos.

A abordagem de lógica difusa (*fuzzy*) consiste em uma alternativa interessante para o controle, uma vez que o objetivo desta técnica de controle é simular um operador, sendo capaz de controlar um processo através da tradução de regras (YAMAZAKI, 1996; RAVI *et al.*, 1997, ARAÚJO JUNIOR, 2007). Alguns estudos utilizam lógica difusa com controladores preditivos (MAHFOUF *et al.*, 2002), com compensação de tempo morto (ZHAO *et al.*, 2000) e outros em conjunto com redes neurais artificiais (SALAHSHOOR e HAMZEHNEJAD, 2010) denominado de sistemas *neuro-fuzzy*. Somsung e Pratisthamanda (2005) utilizaram esta técnica para controlar a composição do produto de topo e de fundo, quando distúrbios no fluxo da corrente de alimentação foram aplicados na coluna. Em Khazraee e Jahanmiri (2010) um esquema inferencial para estimativa da composição de uma coluna de destilação reativa baseado em um sistema inferencial *neuro-fuzzy* (ANFIS) foi empregado para estimar a composição usando medições de temperatura da coluna em uma destilação multicomponente.

Uma técnica amplamente usada para controle da composição em torres de destilação é o desenvolvimento e emprego de analisadores virtuais, em substituição aos analisadores de processo e como alternativa a outras técnicas conforme visto anteriormente.

Segundo James *et al.* (2000), sensores por software ou sensores virtuais são algoritmos matemáticos (modelos inferenciais) capazes de estimar, em tempo real, variáveis de interesse não medidas, como por exemplo a composição de colunas, a partir de outras variáveis disponíveis medidas instantaneamente.

A denominação “sensor virtual” é um termo recente cuja origem do nome está relacionada a uma das principais aplicações desta técnica, que é a substituição de sensores físicos via “software”. Porém, são também encontrados sob diversos outros nomes, sendo que os mais comuns são: analisador virtual, termo bastante difundido principalmente no ambiente industrial, *soft sensor*, estimador *on-line*, ou ainda medidor inferencial (ZANATA, 2005). No transcorrer do texto algumas destas denominações são utilizadas.

Os sensores são os olhos pelos quais o comportamento e o desempenho da planta podem ser observados. Entretanto, existem diversas situações, já comentadas anteriormente, nas quais o sensor não está disponível, devido a falhas, retirada para manutenção ou mesmo inexistência de um sensor adequado. A utilização de analisadores

virtuais poderia suprir a não operação destes analisadores, contribuindo assim para que as unidades operem mais próximas quanto possível do ponto ótimo.

Um sensor por software pode ser descrito de forma simplificada como uma associação entre um sensor (*hardware*) e um estimador, conforme se observa na Figura 2.2. O estimador é parte de um software que produz estimações em linha da variável a partir de medidas obtidas pelos sensores.

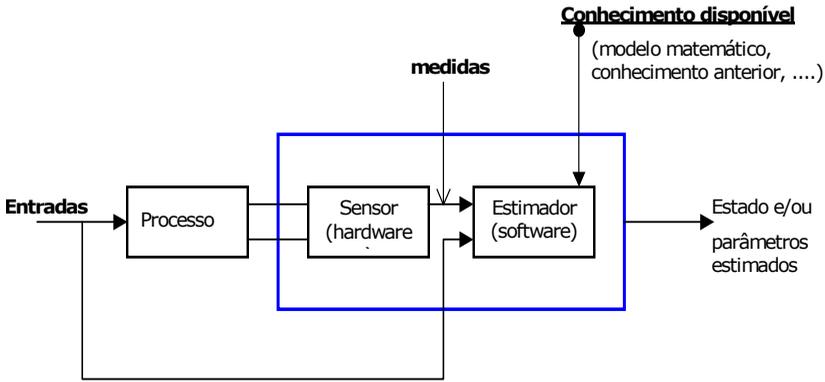


Figura 2.2 Esquema simplificado de um sensor por software  
Fonte: Meneguelo (2007).

A utilização de um sensor por software está sujeita a algumas complicações. Dentre elas pode-se citar duas: a utilização de um modelo inadequado do processo, ou mesmo erro na modelagem ou o sensor responsável pela medida da variável mensurável estar com defeito (MENEGUELO, 2007).

Um das principais motivações para utilização de um sensor por software na predição da composição de uma coluna de destilação é sem dúvida econômica. A rapidez com que a resposta é obtida, comparada com métodos *off-line*, pode trazer benefícios para o controle do processo, fazendo com que a planta opere menos tempo em regime transiente, sem reprocessamento de produtos fora de especificação e, conseqüentemente, diminuição do gasto de energia.

Jianxul e Hiuhe (2002) preferem os *soft-sensors* pela maior confiabilidade e precisão nas estimativas, principalmente em faixas extremas de operação, como afirmam no trabalho de desenvolvimento de um *soft-sensor* para uma unidade de craqueamento catalítico de combustíveis para produção de diesel.

Sobre as tecnologias empregadas para desenvolvimento de *soft-sensors*, encontram-se na literatura três classes distintas de sensores de acordo com os modelos empregados, os quais podem ser: modelagem caixa branca ou princípios fundamentais; modelagem caixa preta ou empírica e os modelos híbridos.

Independentemente de qual enfoque se utilize, uma preocupação que deve haver ao desenvolver-se um *soft-sensor* é determinar quais são as melhores variáveis a serem consideradas como entradas para o modelo inferencial, sempre seguindo a metodologia: 1) conhecimento preliminar do processo; 2) qualidade dos dados experimentais e; 3) determinação e validação do modelo.

Os métodos empregados para viabilizar o controle de qualidade dos produtos através de modelos inferenciais apresentam vantagens como baixo custo e rapidez de resposta. A seção seguinte descreve o método inferencial usado para a avaliação de propriedades do produto.

### **2.3 CONTROLE DO PROCESSO AUXILIADO POR MODELO INFERENCIAL**

Os sistemas de inferência possuem ampla aplicação na indústria de processos. Atuam auxiliando no monitoramento, controle e otimização de plantas industriais. Fornecem medições mais exatas, rápidas e confiáveis a um custo reduzido, quando comparado a métodos tradicionais de análises (LINHARES, 2010). De acordo com Fortuna *et al.* (2007), os métodos de caracterização compostos por sistemas inferenciais, são ferramentas valiosas para serem empregadas em diversos setores da indústria, tais como em refinarias, plantas químicas e de processamento de alimentos, usinas nucleares, dentre tantas outras.

No controle auxiliado por inferência, o sistema recebe os sinais dos sensores e também o valor da variável estimada por um modelo inferencial. O enfoque, porém, passa a ser controlar a propriedade estimada. Isto é, os controladores das malhas se encarregam de controlar as variáveis medidas para manter a especificação do produto no valor correto (KRESTA *et al.*, 1994). Para isso, o resultado da inferência é usado diretamente pelo sistema de controle.

Um exemplo de controle do processo pode ser visualizado na Figura 2.3, sendo que, neste caso, temos a opção de utilizar a medida de um sensor físico (analisador de processo) ou substituí-lo pelo analisador virtual, tornando-o um controle por modelo inferencial. A figura ilustra

o diagrama de uma planta na qual a variável ‘ $y$ ’, medida usando um sensor, está para ser substituída pelo sinal ‘ $y_m$ ’, fornecida pelo *soft-sensor*. Na grande maioria dos casos o *soft-sensor* é o modelo de uma parte da planta. As entradas para este modelo são outras variáveis do processo de fácil aquisição e distúrbios medidos. O modelo usado pode ser do tipo entrada-saída ou na forma de equações de estado e de saída. Este último caso, comumente conduz ao uso de um Filtro de Kalman Estendido. Em certos casos pode ser apropriado usar modelos nebulosos (*fuzzy*) ou modelos baseados em redes neurais, principalmente naqueles casos onde não se podem utilizar informações a respeito do processo, tendo-se como informações apenas os dados medidos da entrada e saída do processo (LOTUFO e GARCIA, 2008).

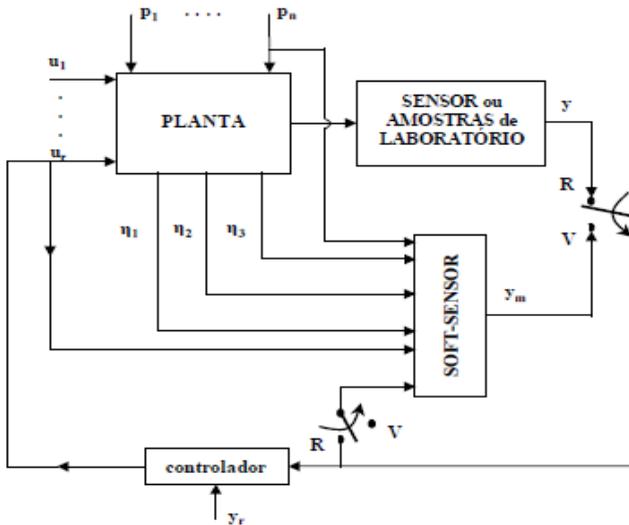


Figura 2.3. *Soft sensor* conectado a uma planta, tornando o processo assistido por modelo inferencial. Fonte: Lotufo e Garcia (2008).

Com referência ainda à Figura 2.3, quando o sensor real está ativo, a chave está na posição ‘ $R$ ’. Então o modelo do *soft-sensor* é identificado, enquanto a malha de controle usa o sinal ‘ $y$ ’ dado pelo sensor real. Quando o sinal do sensor não está mais disponível, a chave deve ser virada para a posição ‘ $V$ ’, em cujo caso a malha de controle usa o sinal ‘ $y_m$ ’ fornecido pelo *soft-sensor*, e a identificação do modelo ou a estimação paramétrica é interrompida. O sensor virtual pode ser empregado em situações que o sensor real passe por manutenção ou

pode-se optar pela substituição definitiva, em casos que o alto custo do analisador de processo for impeditivo.

Existem várias possibilidades de aplicações de modelos para colunas de destilação. Em geral as técnicas mais usadas atualmente são empíricas, mas podem ser utilizados também modelos fenomenológicos, conforme citado anteriormente. Apesar de apresentar diversas vantagens, o modelo inferencial requer um acompanhamento de suas estimativas, o que pode ser feito através de análises de laboratório.

As categorias de modelagem podem ser divididas em três grupos: modelos fenomenológicos (*white-box ou caixa branca*); modelos empíricos (*black-box ou caixa preta*) e modelos semi-empíricos (*gray-box ou caixa cinza*).

Os modelos fenomenológicos são os modelos com maior capacidade de extrapolação. Sua obtenção, no entanto, geralmente demanda um elevado esforço e conseqüentemente muito tempo para o seu desenvolvimento, especialmente em sistemas complexos (FORTUNA *et al.*, 2007, VENTIM, 2010). Este modelo, sendo mais completo na sua descrição, é chamado modelo rigoroso. Um modelo rigoroso pode ser preciso nos resultados, por outro lado, é extremamente complexo (CONZ, 2005). Segundo Ramchandran e Rhinehart (1995) e Kresta *et al.* (1994), normalmente os modelos rigorosos não são apropriados para aplicações em tempo real, pois em muitos casos o tempo de processamento para obtenção da informação de interesse não é aceitável e/ou pode comprometer a eficiência do sistema de controle. Estes motivos são os maiores obstáculos à aplicação destes métodos em ambientes industriais.

Modelos empíricos, por sua vez, são caracterizados por não possuírem nenhuma base fenomenológica do sistema a ser modelado. Eles são baseados na aplicação de métodos matemáticos, como técnicas de regressão, capazes de representar a relação entre as variáveis de entrada e saída de um processo. Sua capacidade de extrapolação é reduzida quando comparada a dos modelos fenomenológicos. Outro fator que deve ser levado em conta, quando se utiliza um modelo empírico baseado em técnicas de regressão, é que eles são capazes somente de modelar os dados e não de modelar o processo como um todo (CONZ, 2005).

Quintero-Marmol *et al.* (1991) aplicaram a técnica de estimativa do estado de sistemas denominada Observadores Estendidos de *Luenberger* para aplicação de perfis de composição instantâneos de colunas de destilação a partir de valores de temperatura. O observador de *Luenberger* é um método determinístico, ou seja, pressupõe que o

modelo do processo é perfeito e que há ausência de ruídos e perturbações. Quintero-Marmol *et al.* (1991) relataram problemas de convergência quando a presença de ruídos nas medidas de temperatura foi considerada. Barolo e Berto (1998) utilizaram um observador estendido de *Luenberger*. Segundo os autores, à medida que o número de pratos da coluna aumenta, torna-se mais difícil de sintonizar esse tipo de observador. Eles recomendam o uso de um estimador estocástico (como o filtro de Kalman) para inferência de composições.

Existem vários trabalhos sobre inferência utilizando modelos empíricos em torres de destilação. Normalmente, é feita uma correlação entre a composição e o perfil das várias temperaturas ao longo da torre e mais algumas variáveis como pressão de topo e de fundo e vazões de refluxo (MEJDELL e SKOGESTAD, 1991; MEJDELL e ANDERSSON, 1993).

Para compensar a falta de tecnologia na época para aquisição de medidas *on-line* para o controle da qualidade do produto, Weber e Brosilow (1972) usaram medidas secundárias de temperatura, vazão de refluxo e de vapor em conjunto com um modelo matemático do processo para estimar a qualidade dos produtos. O estimador foi desenvolvido para controle multicomponente de colunas de destilação.

Na inferência em torres de destilação, apesar de vários tipos de modelos terem sido desenvolvidos, alguns métodos genéricos têm sido propostos. Na destilação, uma prática comum é selecionar a temperatura de um dos pratos. Esta metodologia fornece um modelo chamado de estimador a uma temperatura (*one temperature estimator*) (MEJDELL e SKOGESTAD; 1991, GUPTA *et al.*, 2008). A escolha do prato, ou da temperatura mais adequada, baseia-se na temperatura de maior sensibilidade às mudanças de composição da carga e das vazões de refluxo (BAROLO *et al.*, 2000). Este método possui a desvantagem de limitar o modelo a apenas uma temperatura de prato. Yildiz *et al.*, (2005) desenvolveram um estimador de composição para coluna de destilação baseado em modelo empírico. Os autores concluíram que as posições de medição de temperatura devem ser distribuídos de forma homogênea em toda a coluna, desde o refeedor até o prato superior. Segundo Kresta *et al.* (1994), para desenvolver um modelo com múltiplas variáveis medidas, ou multivariável, deve-se utilizar uma técnica de modelagem linear ou não-linear para ajustar o modelo.

Gupta *et al.* (2008) propuseram o uso de um modelo inferencial usando um observador de posição não-linear que utiliza uma medida de temperatura de um prato para controle de uma coluna de destilação debutanizadora simulada. Os resultados das simulações mostraram que a

abordagem tem o potencial para satisfazer a exigência do controle de processos industriais.

Um modelo inferencial proporciona uma estimativa da variável, que poderá ser utilizada no projeto de um controlador num arranjo *feedback*, para permitir um controle aproximado. A metodologia utilizada no projeto de um estimador, ou modelo de inferência deverá fornecer, preferencialmente, um modelo capaz de trabalhar com qualquer tecnologia de controle, desde o tipo PID até o MPC (*Model Predictive Control*), em arranjos de controle monovariável e multivariável.

Modelos inferenciais podem ser utilizados de forma isolada, como uma alternativa de baixo custo para sensores usuais, ou ainda podem ser empregados em conjunto com sensores físicos gerando, assim, técnicas baseadas em modelos, para o desenvolvimento de esquemas de detecção de falhas e inspeção do funcionamento desses sensores físicos enquanto estes são reparados, com o objetivo de manter as malhas de controle em funcionamento de forma adequada (RIZZO e XILIBIA, 2002).

Os sistemas de inferência possuem grande aplicação, pois atuam auxiliando na monitoração, controle e otimização de processos em geral, fornecem medições mais exatas, rápidas e confiáveis a um custo mais baixo (ARAÚJO JUNIOR, 2007).

### 2.3.1 Técnicas de geração de modelos empíricos

Apesar dos métodos por inferência serem amplamente difundidos nas indústrias, uma quantidade limitada de técnicas são utilizadas para seu desenvolvimento. Os principais tipos de abordagens empíricas empregadas são: os métodos estatísticos, métodos baseados em observadores de estado e a modelagem a partir de alguma técnica de inteligência artificial.

Nesta seção serão abordadas apenas as técnicas de geração de modelos empíricos, mais usadas atualmente, em virtude de algumas limitações do emprego de modelos fenomenológicos, conforme já mencionado. Outro motivo da ênfase para as técnicas empíricas é o fato de esta ser a escolhida neste trabalho para modelagem do processo, sendo que a técnica selecionada foi a baseada em inteligência artificial.

#### 2.3.1.1 Métodos estatísticos

Métodos estatísticos multivariáveis, baseados em projeção linear, têm atraído grande interesse dos pesquisadores (GELADI e

KOWALSKI, 1986; LU e FISHER, 1989; MEJDELL e SKOGESTAD, 1991; LU e FISHER, 1992; KRESTA *et al.*, 1994; ZHANG, *et al.*, 1997; HARTNETT *et al.*, 1998; ZHANG, 2001; EVANGELISTA *et al.*, 2001; KANO *et al.*, 2003; LU *et al.*, 2004; ABDUL-WAHAD *et al.*, 2006) para construção de modelos, especialmente quando os dados do processo apresentam alta dimensionalidade e colinearidades. Dentre estes métodos podemos citar a Regressão Linear Múltipla (MLR), a Regressão nos Componentes Principais (PCR) e os Quadrados Mínimos Parciais (PLS), tanto pela sua praticidade quanto pelos bons resultados apresentados.

De acordo com Nagai (2006), embora estes métodos de projeção linear possam manipular alta dimensionalidade e colinearidade, eles são capazes de extrair somente informações lineares a partir dos dados. Como grande parte dos processos apresentam não-linearidades, métodos mais robustos com capacidade de modelar tais não-linearidades seriam interessantes. Nas próximas seções estes métodos serão abordados. No entanto o método estatístico PLS possui variantes não-lineares (WOLD *et al.*, 1989; QIN e McAVOY, 1992; KRESTA *et al.*, 1994) e são bastante empregados como modelos para estimação de composição em colunas de destilação.

### 2.3.1.2 Métodos baseados em filtros de Kalman

O filtro de Kalman (FK) é um método recursivo para a estimativa do estado de sistemas dinâmicos, chamado de estimador dinâmico estocástico. É uma das técnicas bastante empregadas para estimação de estados no meio industrial. Por volta de 1960, Kalman desenvolveu essa técnica para sistemas lineares. Um detalhamento do algoritmo empregado nesta técnica pode ser encontrado em Kalman (1960) e Welch e Bishop (2004), outras descrições do procedimento em Brown e Hwang (1992). Posteriormente, o filtro de Kalman foi estendido (FKE) para sistemas não-lineares. Outra opção é o uso do Filtro de Kalman Estendido com Restrições que, conforme demonstrado em Marcon e Trierweiler (2002) apresenta um desempenho mais robusto quando implementado.

O filtro de Kalman é considerado uma ferramenta matemática com capacidade de filtrar sinais de medição sob o regime dinâmico. É usado na estimação e reconciliação dinâmica de variáveis de processos. Um filtro também usado para sistemas dinâmicos é o *Dynamic Data Reconciliation* (DDR) que possui o mesmo princípio matemático do FK, que é baseado no uso da informação de medição e um modelo

estocástico do processo (MARTINS *et al.*, 2009). Uma aplicação do uso destes filtros dinâmicos é discutida em Bai *et al.* (2006).

Várias pesquisas empregaram o filtro de Kalman e sua versão estendida para prever a composição de topo de colunas de destilação seja em processo em batelada ou contínuo, binária ou multicomponente (BARATTI *et al.*, 1995; OISIOVICI e CRUZ, 2001; VENKATESWARLU e AVANTIKA, 2001; VENKATESWARLU e KUMAR, 2006), demonstrando a eficácia do esquema de estimação desenvolvido, quando aplicada esta metodologia.

### 2.3.1.3 Métodos baseados em inteligência computacional – RNA's

Redes neurais artificiais (RNA's) podem ser usadas para obter o modelo empírico para o sensor a ser desenvolvido, uma vez que esse método pode fornecer uma representação precisa do comportamento do processo e requer baixa carga computacional comparativamente com modelos fenomenológicos (ZAMPROGNA, 2001b). Este método será empregado para elaboração do sensor por software, e por isso será mais detalhado nesta revisão.

No campo da inteligência artificial, redes neurais artificiais são modelos paramétricos não-lineares que procuram realizar uma aproximação do mecanismo de processamento do cérebro humano (SANTOS *et al.*, 2005).

A capacidade de aprender através de exemplos e de generalizar a informação aprendida são, sem dúvida, os atrativos principais da solução de problemas através de redes neurais. A generalização, que está associada à capacidade da rede aprender através de um conjunto reduzido de exemplos e, posteriormente, dar respostas coerentes para dados não conhecidos, é uma demonstração de que a capacidade das redes neurais vai muito além do que simplesmente mapear relações de entrada e saída. Elas são capazes de extrair informações não apresentadas de forma explícita através dos exemplos.

As redes neurais são capazes de atuar como mapeadores universais de funções multivariáveis, com custo computacional que cresce apenas linearmente com o número de variáveis. Outras características importantes são a capacidade de auto-organização e de processamento temporal que, aliadas àquelas citadas anteriormente, fazem das redes neurais uma ferramenta computacional extremamente poderosa e atrativa para a solução de problemas complexos (ZANATA, 2005).

Esta seção tem o objetivo de contextualizar a metodologia adotada na execução deste trabalho. Serão abordadas as principais estruturas neurais aplicáveis no tipo de problema abordado e seus princípios, com o intuito de introduzir os conceitos a respeito das RNA's *feedforward* e recorrentes, além da descrição de alguns algoritmos de aprendizado, com o intuito de facilitar a compreensão do método de predição proposto e empregado nos capítulos seguintes.

De forma geral, redes neurais artificiais podem ser divididas quanto ao tipo de aprendizado em duas categorias: redes com aprendizado supervisionado e redes com aprendizado não-supervisionado. No caso supervisionado, cada entrada apresentada à rede vem acompanhada de uma saída desejada, a fim de permitir uma modificação dos parâmetros ajustáveis em função do erro entre a resposta fornecida pela rede e a saída real desejada. Ao final da etapa de ajuste dos parâmetros, chamada genericamente de treinamento, as respostas da rede para todas as entradas devem ser próximas das saídas desejadas. No caso não-supervisionado, a rede neural detecta padrões e características estatísticas do espaço de entrada, de forma a construir uma representação de dimensionalidade reduzida do mesmo no conjunto de pesos sinápticos de seus neurônios. Neste tipo de redes não existe nenhuma meta estabelecida para alcançar, e a rede deve descobrir categorias interessantes ou características nos dados de entrada. Pertencem a esta categoria redes de Hopfield (HOPFIELD, 1982) e redes ou Mapas de Auto Organização de Kohonen (*Self-Organizing Maps* – SOM) (KOHONEN, 1982; SILVA, 1998; SOUZA, 2005; GONÇALVES, 2009; CHERIF *et al.*, 2011).

A rede Hopfield emprega um princípio chamado de armazenamento de informação como atratores dinamicamente estáveis, considerando a rede um conjunto de neurônios binários comunicando-se entre si através de eficácias sinápticas simétricas. Apresenta comportamento dinâmico e fluxo de dados multidirecional devido à integração total dos neurônios, desaparecendo, assim, a idéia das camadas bem distintas. Também conhecida como Memória Associativa, as redes Hopfield armazenam padrões que são recuperados a partir de estímulos de entrada. O armazenamento de tais padrões é realizado via *Aprendizado Hebbiano*, de acordo com a teoria de Donald Hebb (HEBB, 1949).

Algumas aplicações de redes de Hopfield podem ser encontradas em Metz (2008) e Ruéla (2005). Um levantamento bibliográfico de alguns trabalhos com o emprego de redes de Kohonen estão disponíveis em Oja *et al.* (2003). Este tipo de rede utiliza o

*Aprendizado Competitivo*, onde os neurônios competem entre si para responder a um estímulo apresentado.

As topologias das principais redes são ilustradas na Figura 2.4. Algumas destas topologias relacionadas ao escopo deste projeto serão discutidas com maiores detalhes na próxima seção.

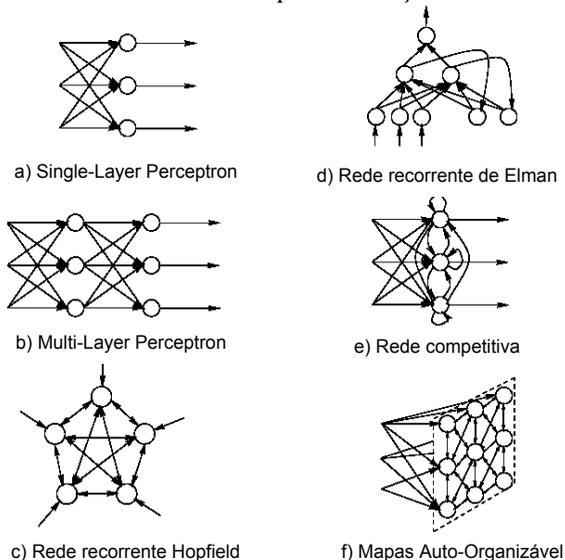


Figura 2.4. Topologias das principais tipos de redes neurais artificiais.

Fonte: Adaptado de Melo (2004).

Neste trabalho serão usadas apenas redes neurais com aprendizado supervisionado, diferenciando umas das outras apenas pelo modo que processam informação temporal, ou seja, se utilizam ou não de laços de realimentação. Desta forma, arquiteturas supervisionadas que não contenham tais laços são comumente chamadas de não-recorrentes ou *feedforward*. Já as arquiteturas contendo laços de realimentação são denominadas de redes neurais recorrentes. Nesta revisão serão contempladas as redes com treinamento supervisionado.

### 2.3.1.3.1 Redes *Feedforward*

Redes neurais *feedforward* são aproximadores universais de funções, ou seja, são capazes de aproximar mapeamentos entrada-saída complexos com grau de precisão arbitrário, sejam tais mapeamentos contínuos ou descontínuos. Esta propriedade é uma das responsáveis

pela ampla popularização do uso de redes neurais artificiais em tarefas de reconhecimento de padrões e aproximação de funções. Segundo Menezes Junior (2006), a formulação geral do problema de aprendizado de uma rede neural se aplica a arquiteturas de redes neurais envolvidas tanto em tarefas de aproximação de funções, quanto em tarefas de classificação de padrões.

#### 2.3.1.3.1.1 Multi-Layer Perceptron – MPL

Este tipo de rede se enquadra na classe com treinamento supervisionado. Tipicamente, uma rede Perceptron Multicamada (*Multilayer Perceptron* - MLP) é constituída de uma camada de entrada que recebe os sinais, uma ou mais camadas intermediárias, compostas por neurônios somadores com função de ativação não-linear e uma camada de saída, também composta por neurônios somadores, embora estes possam ter funções de ativação lineares.

As camadas intermediárias são comumente chamadas de camadas escondidas, visto que os neurônios nelas localizados não têm acesso direto aos sinais da entrada nem da saída. A existência de camadas intermediárias não-lineares confere à rede MLP o poder computacional de resolver problemas complexos, pois tais camadas têm a função de promover sucessivas alterações na representação dos dados originais até que o problema possa ser resolvido pela última camada de neurônios (camada de saída).

Outra característica da rede MLP é seu alto grau de conectividade, determinado pelas sinapses da rede, interligações entre os neurônios de diferentes camadas, em que cada uma delas está associada a um valor numérico chamado de peso sináptico. A Figura 2.5 ilustra a arquitetura geral de uma rede MLP de uma única camada intermediária. A rede possui ainda 5 entradas ( $x_1, x_2, x_3, x_4$  e  $x_5$ ) e 1 saída ( $y_1$ ). Uma vez especificado o número de camadas e a quantidade de neurônios em cada uma delas, o processo de aprendizado da rede MLP é realizado através do ajuste dos pesos sinápticos e limiares de ativação por meio do algoritmo de retropropagação do erro (*backpropagation*). Este algoritmo é baseado em uma regra de aprendizagem que “corrige” o erro durante o treinamento (HAYKIN, 1999). O MLP com aprendizado baseado no algoritmo de retropropagação do erro é o modelo mais utilizado em áreas de conhecimento, tais como reconhecimento de padrões, processamento de fala e na previsão de séries temporais (MAJI, 2007; PINO *et al.*, 2008; GIORDANO *et al.*, 2007).

Uma aplicação típica das redes tipo *feedforward* é o desenvolvimento de modelos complexos não lineares que podem ser usados por sua vez, em modelos inferenciais e classificação (SINGH *et al.*, 2005; FORTUNA, 2005; SHAHSAVAND *et al.*, 2011; LEE e HONG, 2011; ROGINA *et al.*, 2011; PINTÉR, 2012).

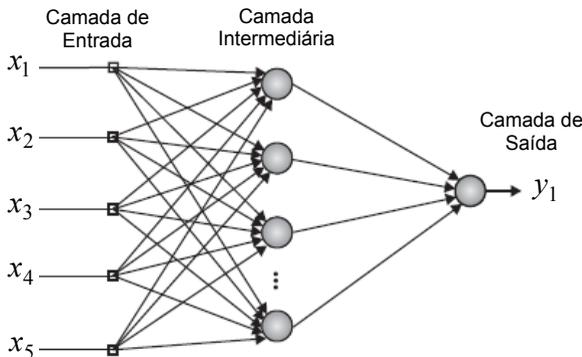


Figura 2.5. Arquitetura de rede tipo *feedforward-MPL*.

Fonte: Adaptado de Menezes Junior (2006).

Neste tipo de configuração não existem conexões ligando um neurônio de uma camada a outro de uma camada anterior, nem a um neurônio da mesma camada. As conexões são feitas somente para neurônios da camada seguinte.

Redes de uma camada escondida são capazes de aproximar com precisão arbitrária funções contínuas, enquanto redes MLP de duas ou mais camadas podem aproximar até funções descontínuas. Em relação ao número de neurônios em cada camada intermediária, um valor otimizado para o número de neurônios em cada camada escondida é geralmente encontrado por experimentação, em função da capacidade de generalização da rede. Se muitos neurônios existirem na camada intermediária, o desempenho será muito bom para os dados de treinamento, mas tende a ser ruim para os novos dados. Se existirem poucos neurônios, o desempenho será ruim também para os dados de treinamento. O valor ideal é aquele que permite atingir as especificações de desempenho adequadas tanto para os dados de treinamento, quanto para os novos dados. Este item, juntamente com o número de camadas intermediárias, define o poder computacional necessário para a rede MLP.

### 2.3.1.3.2 Redes recorrentes

Uma rede neural dinâmica recorrente, ou simplesmente rede recorrente, é aquela que contém conexões sinápticas realimentadas (ou laços de realimentação), permitindo o fluxo de sinais de ativação e saída neurais entre neurônios de camadas distintas, entre neurônios de uma mesma camada, ou ainda de um neurônio para ele mesmo, gerando interconexões bem mais complexas que as redes não recorrentes (JAIN, 2001). A recorrência é um tipo de mecanismo de memória de curta duração que permite a rede relembrar informações de um passado recente.

A realimentação significa que de alguma forma a saída de um neurônio é reprocessada como nova entrada daquele mesmo neurônio ou da camada de entrada da rede, caracterizando a recorrência. Este tipo de rede é de particular interesse na previsão de séries temporais, já que a realimentação de seus neurônios permite a elas reprocessar de forma dinâmica o erro destas saídas e, com isto, lidar melhor com certas características de padrões não lineares (PASQUOTO, 2010).

Um modo de classificar redes recorrentes consiste em verificar a extensão espacial das conexões de realimentação existentes, ou seja, se ela envolve apenas os neurônios de uma única camada ou se envolve neurônios de outras camadas. Pode-se enquadrar esta definição em três grupos:

- *Conexão Recorrente Local*: este tipo de conexão envolve apenas um neurônio. Neste caso, o termo local refere-se ao fato da saída do neurônio ser realimentada para a entrada deste mesmo neurônio.
- *Conexão Recorrente Global*: se a realimentação engloba alguma(s) camada(s) completa(s), este tipo de conexão acontece entre um neurônio de uma camada para um neurônio de uma camada anterior, ou seja, um sinal de saída de um neurônio é realimentado para a entrada de outro neurônio localizado em uma camada distinta.
- *Conexão Recorrente Não-Local*: esta é um tipo especial de conexão global, visto que envolve neurônios distintos. Porém, a conexão é estabelecida entre neurônios de uma mesma camada. Assim, a saída de um neurônio de certa camada é realimentada para a entrada de outro neurônio da mesma camada.

Redes neurais recorrentes constituem uma das mais importantes famílias de arquiteturas de redes neurais. Tsoi e Back (1997) discutem e listam como diversas arquiteturas neurais recorrentes podem ser geradas pelas mais variadas combinações de realimentações entre neurônios de uma mesma camada e entre neurônios de diferentes camadas, conforme

destacado anteriormente, de tal forma que se pode facilmente entender a razão da grande diversidade de arquiteturas recorrentes encontrada na literatura. Vale ressaltar que, como os pesos das conexões de realimentação não são ajustáveis, pode-se usar o algoritmo *backpropagation* padrão para treinar tais redes.

#### 2.3.1.3.2.1 Rede de Elman

Elman (1990) propôs uma RNA recorrente, adicionando uma camada intermediária, denominada camada de contexto, que consiste de atrasos unitários. A camada é responsável por guardar a informação temporal relativa às entradas anteriores da rede. Observa-se na arquitetura da rede recorrente de Elman, ilustrada na Figura 2.6, que ocorre uma redefinição da camada de entrada da rede, que passa a ser dividida em duas partes. A primeira parte corresponde ao vetor de entrada propriamente dito. A segunda parte, chamada de unidades de contexto, consiste na cópia das saídas dos neurônios da camada escondida no instante  $n-1$ . O termo cópia, na verdade, é implementado computacionalmente através de um conjunto de conexões de realimentação com pesos fixos e iguais a 1, fato que também possibilita a utilização do algoritmo *backpropagation*. Desta forma, os valores exatos das ativações dos neurônios da camada intermediária no instante  $n-1$  são utilizados pelas unidades de contexto para compor o vetor de entrada no instante  $n$ . A camada de contexto da rede de Elman é responsável por armazenar a informação temporal relativa as entradas anteriores, o que a diferencia da rede MPL. A entrada e a saída da rede estão definidas de acordo com o problema de predição.

As unidades de contexto são usadas apenas para memorizar as ativações anteriores das unidades intermediárias. Podem ser consideradas como atraso no tempo em um passo, constituindo a memória da rede e complementam as entradas dinâmicas. As conexões *feedforward* são modificáveis e as conexões recorrentes podem ser fixas, motivo pelo qual a rede de Elman é apenas parcialmente recorrente. Do ponto de vista das conexões, a rede de Elman possui apenas recorrências globais. As RNA's parcialmente recorrentes surgiram a partir de abordagens que utilizam ligações recorrentes em alguns neurônios pela necessidade de resolver problemas de instabilidade e complexidade do aprendizado das RNA's totalmente recorrentes (SCHÜHLI, 2005). Redes totalmente recorrentes são também conhecidas como redes simétricas e caracterizam-se por apresentar conexões entre os neurônios nos dois sentidos, como é o caso das redes de Hopfield.

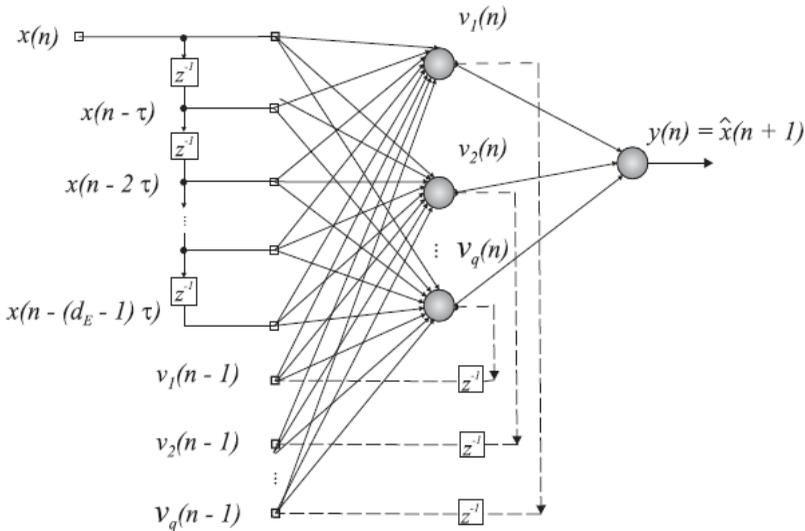


Figura 2.6. Arquitetura da rede neural artificial recorrente de Elman.

Fonte: Menezes Junior (2006).

Segundo Menezes Junior (2006), para redes com duas camadas intermediárias pode-se escolher de qual camada intermediária realimentar as ativações neurais. Desta forma, seriam três as opções possíveis:

- retroalimentação das ativações da primeira camada intermediária para as unidades de contexto,
- retroalimentação das ativações da segunda camada intermediária para as unidades de contexto, ou
- retroalimentação das ativações de ambas as camadas intermediárias para as unidades de contexto.

Com relação à quantidade de unidades de contexto, esta depende diretamente de qual das opções listadas no parágrafo anterior for escolhida. No primeiro caso, a quantidade de unidades de contexto é igual ao número de neurônios da primeira camada intermediária ( $q_1$ ). No segundo caso, é igual ao número de neurônios da segunda camada ( $q_2$ ) e assim consecutivamente. Na terceira situação, a quantidade de unidades de contexto é igual a  $q_1 + q_2$ .

Neste trabalho utilizou-se a terceira opção, ou seja, as ativações de cada uma das camadas intermediárias foram realimentadas para as unidades de contexto. Nos últimos anos essa arquitetura de rede vem sendo usada frequentemente em diversas áreas da indústria nas mais

diversas aplicações (BAHAR e OZGEN, 2008; ARDALANI-FARSA e ZOLFAGHARI, 2010; ÜBEYLİ e ÜBEYLİ, 2010; PASQUOTTO, 2010; CHENG *et al.*, 2011; PALADE e AO, 2011).

### 2.3.1.3.2 Rede de Jordan

A rede de Jordan (1986) é outra arquitetura recorrente clássica, sendo inicialmente usada para reconhecimento de sequências temporais. Assim como a rede de Elman, a rede de Jordan também não possui recorrência entre neurônios da mesma camada, sendo por isso enquadrada entre as redes globalmente recorrentes. Pertence a classe de redes com treinamento supervisionado.

A diferença entre a rede de Elman e a de Jordan é que, na rede de Jordan, a realimentação das ativações dos neurônios ocorre dos nós de saída para os nós das unidades de contexto, enquanto que na rede de Elman a realimentação sai dos nós das camadas intermediárias. Além disso, este tipo de rede recorrente possui auto-conexões ou auto-realimentações, em que a saída de uma unidade de contexto é realimentada para sua entrada, ou seja, as unidades de contexto são localmente recorrentes. A Figura 2.7 ilustra uma rede recorrente de Jordan com uma camada intermediária.

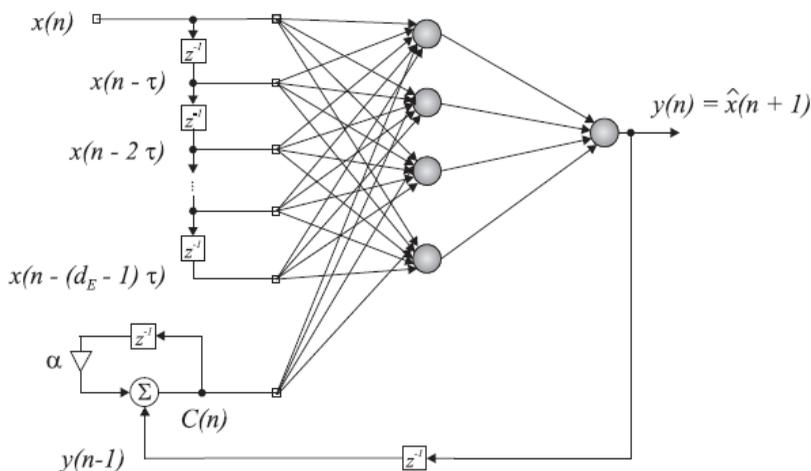


Figura 2.7. Arquitetura da rede neural artificial recorrente de Jordan.

Fonte: Menezes Junior (2006).

Um exemplo interessante de aplicação deste tipo de rede pode ser verificado em Carcano *et al.* (2008). Os autores propuseram o uso de uma rede de Jordan e de rede MPL, visando demonstrar a capacidade destas redes para simular o fluxo diário dos rios, para fins de recursos hídricos, comparando os resultados com uma metodologia conceitual. Para tal, usaram dados de precipitação diárias da região e de temperatura como entradas. Os resultados confirmaram que as RNA's são capazes de suportar a comparação com um processo conceitual. A obtenção de bons desempenhos foi atribuída ao treinamento eficaz da rede e ao fornecimento de quantidade suficiente de informações com qualidade adequada.

### 2.3.1.3.3 Funções de ativação das redes neurais

A função de ativação tem por atribuição repassar o sinal para a saída do neurônio, sendo uma função de ordem interna. A função de ativação determina o valor de saída de cada neurônio em função do valor local induzido. Pode ser também referida como função restritiva, já que limita o intervalo permissível de amplitude de sinal de saída a um valor finito (TERÁN, 2009).

O primeiro grupo compreende as funções de transferência tipo degrau, que permitem que os neurônios de um “*perceptron*” escolham uma dentre duas possíveis saídas. Este tipo de função é largamente empregada em problemas de classificação. Já as funções sigmoidais são empregadas com sucesso como funções de transferências, porque apresentam a propriedade de ser derivável em todos os pontos.

A função logarítmica gera saídas entre 0 e 1, sendo definida como uma função assimétrica estritamente crescente que exibe um balanceamento adequado ente comportamento linear e não linear. Algumas vezes é interessante que a função de transferência se estenda de -1 a +1 assumindo, neste caso, uma forma simétrica. Neste caso, a função de ativação mais adequada é a tangente hiperbólica. Algumas destas formas são representadas na Figura 2.8, e estão disponíveis no software MATLAB<sup>®</sup>.

Refenes *et al.* (1993) mostraram que as funções simétricas geralmente possuem uma convergência bem mais rápida que as assimétricas, mas o aprendizado pode ser extremamente lento se os pesos forem muitos pequenos.

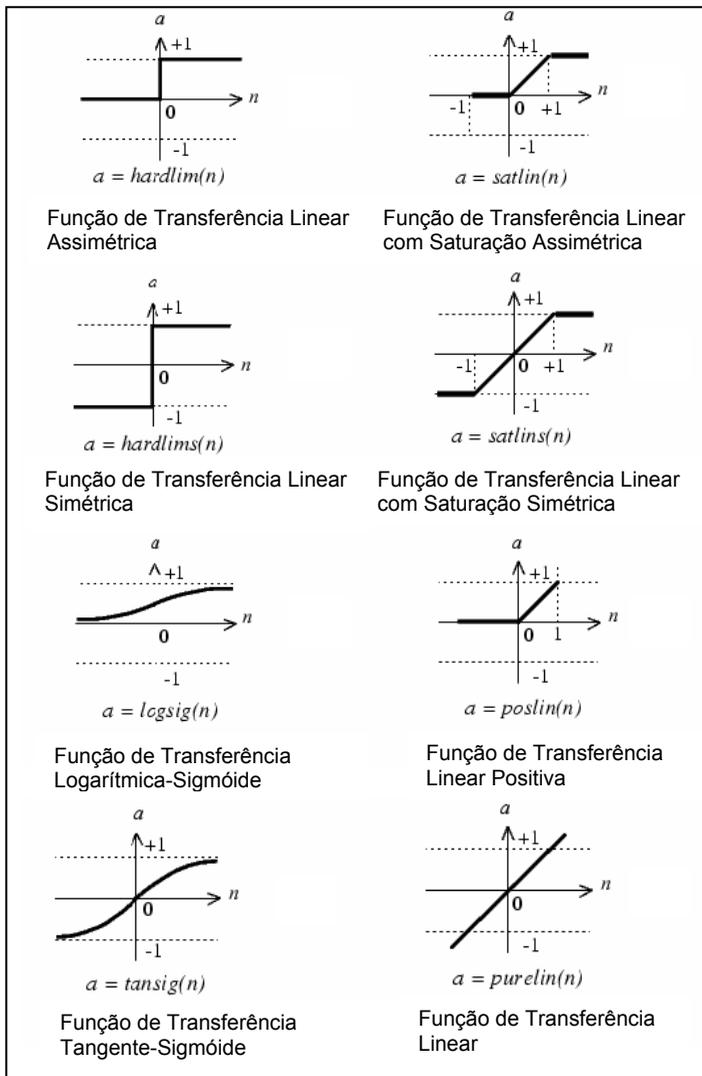


Figura 2.8. Principais funções de transferências aplicáveis às redes neurais.  
 Fonte: Adaptado de Zanata (2005).

As funções de ativação a serem utilizadas nas unidades (neurônios) que compõem a rede exercem um papel muito importante no desempenho do modelo. Alguns aspectos que devem ser considerados

na escolha da função de ativação são, por exemplo, satisfazer o critério de aproximação universal e facilidade de computação (esforço computacional). As funções de ativação são geralmente não-lineares permitindo que o mapeamento feito pela rede seja não-linear. Mapeamentos multivariáveis não-lineares são ideais para descrever uma grande variedade de processos não-lineares como, por exemplo, colunas de destilação e reatores químicos (VEAUX *et al.*, 1993).

#### 2.3.1.3.4 Algoritmos de treinamentos

Existem vários algoritmos para realizar o treinamento de uma rede *feedforward*. De maneira geral, segundo Tarsinaffo (2003), estes algoritmos podem ser divididos em duas categorias distintas: métodos de treinamento determinísticos e métodos de treinamento estocásticos. Entre os métodos determinísticos citam-se alguns: gradiente descendente com e sem *momentum*, gradiente conjugado e de *quasi-Newton*, no qual se pode incluir o método de *Levenberg-Marquardt*. Entre os métodos estocásticos estão: a filtragem de Kalman na versão *full* ou com processamento paralelo e recursivo (JAZWINSKI, 1970; TARSINAFFO, 2003). Todos estes métodos possuem em comum o emprego da retro-propagação para o cálculo das derivadas parciais das variáveis de saída da rede em relação aos parâmetros de ajuste ou pesos sinápticos. Neste trabalho será abordado apenas o estudo dos métodos determinísticos.

O treinamento com o método *backpropagation*, detalhado no artigo original de Rumelhart (RUMELHART e McCLELLAND, 1986), pode ser realizado por vários algoritmos diferentes. Há, entre eles, grande variedade de requisitos computacionais e de exigências de armazenamento de dados. Segundo Zanata (2005) nenhum deles pode ser considerado como melhor do que os demais, mas sim o mais adaptado para uma situação específica. É preciso analisar dentre estes quais apresentam melhor resultado e não comprometem a estrutura do sistema de desenvolvimento. Na Tabela 2.1 são apresentados alguns algoritmos de treinamentos, cada um destes difere entre si, pela estratégia de otimização utilizada para realizar os ajustes nos pesos e bias da rede neural.

Tabela 2.1. Relação dos principais algoritmos de treinamentos com seus respectivos métodos de otimização.

<i>Método</i>	<i>Algoritmo de Treinamento</i>	<i>Sigla</i>	<i>Função MATLAB</i>
Gradiente Descendente (1ª ordem)	Gradiente descendente básico (Hagan <i>et al.</i> , 1996)	GD	<i>traingd</i>
	Gradiente descendente com taxa de aprendizado adaptativa (Hagan <i>et al.</i> , 1996)	GDA	<i>traingda</i>
	Gradiente descendente com momento (Hagan <i>et al.</i> , 1996)	GDM	<i>traingdm</i>
	Gradiente descendente com momento e taxa de aprendizado adaptativa (Hagan <i>et al.</i> , 1996)	GDX	<i>traingdx</i>
Gradiente Conjugado (2ª ordem)	<i>Resilient backpropagation</i> (Riedmiller e Braun, 1993)	RP	<i>trainrp</i>
	Gradiente conjugado escalonado (Møller, 1993)	SCG	<i>trainscg</i>
	Gradiente conjugado de <i>Powell-Beale</i> (Beale, 1972; Powell, 1997)	CGB	<i>traingb</i>
	Gradiente conjugado de <i>Polak-Ribière</i> (Fletcher e Reeves, 1964; Hagan <i>et al.</i> , 1996)	CGP	<i>traingp</i>
	Gradiente conjugado de <i>Fletcher-Reeves</i> (Fletcher e Reeves, 1964; Hagan <i>et al.</i> , 1996)	CGF	<i>traingf</i>
quasi-Newton (2ª ordem)	<i>Broyden-Fletcher-Goldfarb-Shanno</i> (BFGS) (Dennis e Schnabel, 1983)	BFG	<i>trainbfg</i>
	<i>Levenberg-Marquardt</i> (Hagan and Menhaj, 1994)	LM	<i>trainlm</i>
	Regularização de <i>Bayesian</i> (MacKay, 1992; Foresee e Hagan, 1997)	BR	<i>trainbr</i>
Gradiente Conjugado e quasi-Newton	Método da secante de um passo (Battiti, 1992) (2ª ordem)	OSS	<i>trainoss</i>

O objetivo desta seção não é indicar a eficiência relativa destes algoritmos em uma dada aplicação, mas sim, analisar suas principais propriedades e definir o escopo de aplicação de alguns dos algoritmos mencionados. Alguns destes métodos requerem poucas modificações no algoritmo de retropropagação padrão. No entanto, aceleram a convergência média da fase de treinamento, além de não mais requererem a escolha de parâmetros críticos da rede neural como a taxa de aprendizagem e o coeficiente de momento. Serão apresentados vários métodos de otimização, sendo que alguns destes foram implementados e testados neste trabalho.

Além dos métodos citados, existe também outra classe, os *Métodos sem Diferenciação*. São algoritmos que não requerem derivação, apenas avaliação da função em diferentes pontos do espaço. Dentre esses métodos podemos citar *Simulated Annealing* – SA (VASAN e RAJU, 2009; DAMODARAN e VÉLEZ-GALLEGO, 2012), *Algoritmos Genéticos* – AG (SU *et al.*, 2011) e *Busca Exaustiva* – BE (PENDHARKAR, 2010), aplicados à construção e determinação dos pesos de uma rede neural. Estes métodos não serão abordados nesta revisão. É possível adiantar que seus desempenhos são inferiores aos métodos de primeira e segunda ordem quando os problemas são diferenciáveis.

Dentre os métodos que utilizam diferenciação e busca, o *método do gradiente* é o mais simples, pois utiliza apenas informações de primeira ordem, ou seja, as primeiras derivadas para minimizar o erro. Nos métodos de primeira ordem, apenas o gradiente local determina a direção de minimização do erro. São conhecidos como métodos da direção de maior decréscimo (*steepest descent ou gradient descent*). No algoritmo padrão, a minimização é geralmente feita com um passo  $\alpha$  fixo. A determinação do passo  $\alpha$  é fundamental, pois para valores muito baixos, o tempo de treinamento pode tornar-se exageradamente alto, e para valores muito altos, os parâmetros podem divergir (HAYKIN, 1999). A velocidade de convergência é geralmente melhorada adicionando-se um termo de momento (RUMELHART e McCLELLAND, 1986). Com este fator *momentum*, o algoritmo responde não somente ao gradiente local, mas também às tendências recentes na superfície de erro, sem que se tenha de aumentar a taxa de aprendizado, evitando com isto, uma eventual instabilidade que este aumento poderia causar (ZANATA, 2005).

O gradiente descendente com taxa de aprendizado adaptativa é similar ao gradiente descendente empregando os mesmos princípios, porém, tem como diferencial o fato da taxa de aprendizagem ser

variável, fazendo com que se tenha um aprendizado mais rápido do que com taxa fixa, evitando instabilidade e oscilações no processo de aprendizagem.

Em relação ao algoritmo *Resilient Backpropagation* (RP), proposto por Riedmiller e Braun (1993) alguns autores, como Anastasiadis *et al.*, (2004) e o próprio Riedmiller e Braun (1993), o definem como sendo uma melhoria sobre a técnica de aprendizagem *backpropagation*. Neste, a alteração dos pesos sinápticos é baseada em informações de gradientes locais. Segundo Soares (2008), que empregou este algoritmo para predição de séries temporais econômicas, é importante citar que o ajuste da taxa de aprendizado e dos pesos sinápticos dependem dos termos do gradiente e não da sua magnitude, como no *backpropagation*. Com a utilização deste algoritmo o tamanho da atualização dos pesos depende apenas da seqüência de troca de sinais das derivadas parciais. Este fato proporciona que ele realize um grande número de interações. É considerado um algoritmo simples com rápida convergência e pouca memória requisitada.

Atualmente os métodos que empregam a segunda derivada são considerados a maneira mais eficiente para realizar o treinamento de redes neurais do tipo MLP.

Enquanto o *backpropagation* padrão emprega o gradiente descendente como método de aproximação do mínimo da função erro, o algoritmo de *Levenberg-Marquardt* (LM) utiliza uma aproximação pelo método de Newton. Esta aproximação é obtida a partir da modificação do método de Gauss-Newton.

O método de Newton pode ser considerado como o método local básico que utiliza informações de segunda ordem. Deve ser enfatizado que a sua aplicação prática aos “*perceptrons*” multicamadas é pouco recomendada, uma vez que o cálculo da matriz Hessiana é complexo e oneroso e representa um elevado custo computacional, pois exige a inversão, análise espectral e armazenagem de uma matriz quadrada que é da ordem do número de parâmetros a serem ajustados (BATTITI, 1992).

Existe uma classe de algoritmos que é baseada no método de Newton, porém utilizam apenas uma aproximação da matriz Hessiana a cada iteração do algoritmo. A atualização é computada como uma função do gradiente. Estes algoritmos são chamados de métodos quasi-Newton ou secante.

O método *Levenberg-Marquardt* (LM) é considerado do tipo “*quasi-Newton*” utilizado para mínimos quadrados não lineares. Foi incorporado ao algoritmo de retropropagação do erro para resolver

problemas de otimização que aparecem no treinamento de redes multicamadas (KERMANI *et al.*, 2005; SINGH *et al.*, 2007; MUKHERJEE e SRIKANTA, 2012). O processo do método LM é descrito em Bishop (1995). É uma técnica de otimização que torna a convergência rápida.

Este algoritmo, assim como o método de Newton, é bastante eficiente quando trata-se de redes que não possuam mais do que algumas centenas de conexões a serem ajustadas (HAGAN *et al.*, 1996). Isto se deve, principalmente, ao fato de que estes algoritmos necessitam armazenar uma matriz quadrada cuja dimensão é da ordem do número de conexões da rede. Dentre as principais desvantagens do método de *Levenberg-Marquardt* está a necessidade de uma grande quantidade de memória para que ele possa ser executado e o fato do treinamento de cada época ser demorado por exigir uma grande capacidade de processamento (quando comparado com os demais algoritmos), o que requer recurso computacional moderado.

Outro método de quasi-Newton que tem sido mais bem sucedido em estudos publicados é o de Broyden, Fletcher, Goldfarb, e Shanno (BFGS). O algoritmo também possui como base para os ajustes dos pesos o método de Newton, ou seja, utiliza uma aproximação da matriz Hessiana para fazer a correção dos pesos e bias da rede. Requer mais recurso computacional para efetuar as iterações do que os algoritmos de gradiente conjugado, porém geralmente converge em poucas interações. Os detalhes do procedimento BFGS são apresentados em Dennis e Schnabel (1983), Fletcher (1987) e Nocedal (1991).

O algoritmo de treinamento Regularização de *Bayesian* (BR) é uma técnica mais recente para treinar o modelo neural. É considerado uma modificação do algoritmo *Levenberg-Marquardt*, para gerar dados com melhor generalização. Nesta técnica, a quantidade de neurônios da camada intermediária não é estimada por um processo de tentativa e erro. Na regularização *bayesiana* o processo de treinamento faz com que os parâmetros (pesos sinápticos) dos neurônios ocultos desnecessários fiquem com valor próximo a zero. O objetivo principal é encontrar um equilíbrio entre a quantidade de parâmetros e a complexidade do modelo neural exigida pelo problema.

A Regularização de *Bayesian* não requer um conjunto teste para minimizar o excesso de treinamento. Pelo contrário, os pesos e bias em cada nó são assumidos como sendo variáveis aleatórias com distribuições especificadas (MACKAY, 1992; FORESEE e HAGAN, 1997). Medeiros e Pedreira (2001) descrevem o processo de otimização dos parâmetros de regularização aplicando a regra de Bayes. Uma discussão detalhada do uso da regularização *bayesiana*, em combinação

com o treinamento de *Levenberg-Marquardt*, pode ser encontrado em Foresee e Hagan (1997).

O método de Gradiente Conjugado (CG) é um método alternativo aos métodos de quasi-Newton e Gradiente descendente. É projetado para exigir menos cálculos que o método de Newton e apresentar taxas de convergência maiores que as do método do gradiente, mesmo utilizando informações de segunda ordem. Possui requisitos de memória similares ao método de quasi-Newton. Neste grupo, pode-se citar o algoritmo SCG – Gradiente Conjugado em Escala (Møller, 1993), utilizado neste estudo, que possui como característica adaptar-se com facilidade a uma grande variedade de problemas. Os métodos de Gradiente Conjugado são bastante aplicados e tratam de forma efetiva problemas de grande escala.

Como os algoritmos de quasi-Newton requerem mais armazenamento e recurso computacional em cada iteração que os algoritmos de gradiente conjugado, existe a necessidade de uma aproximação secante com menor armazenamento e computação. O método de secante em um passo (*One Step Secant*), detalhado em Battiti (1992), é uma tentativa intermediária entre os algoritmos de gradiente conjugado e os algoritmos quasi-Newton. Esse algoritmo não armazena a matriz hessiana completa. Ele assume que a cada iteração, a Hessiana anterior é uma matriz identidade. Existe uma vantagem adicional de que a nova direção de pesquisa seja calculada sem computar a matriz inversa.

No método da secante de um passo ou *One Step Secant* (OSS), o termo secante provém do fato de que as derivadas são aproximadas por secantes avaliadas em dois pontos da função, onde neste caso a função é o gradiente. A principal razão da redução do esforço computacional deste método em relação aos anteriores, é que a direção de atualização é calculada somente a partir de vetores determinados pelos gradientes, e não há mais a armazenagem da aproximação da inversa da matriz Hessiana.

Segundo Plumb *et al.* (2005), nos últimos anos pouca atenção tem sido dada ao efeito do algoritmo de treinamento sobre a capacidade preditiva do modelo resultante. Alguns pesquisadores como, por exemplo, Murtoniemi *et al.* (1993) avaliaram o algoritmo *backpropagation* padrão e concluíram que a seleção correta do algoritmo de treinamento foi necessário a fim de maximizar a capacidade preditiva.

Em uma pesquisa inovadora, Dai *et al.* (2011) propuseram um novo algoritmo estocástico de otimização que denominaram SOA

(*Seeker Optimization Algorithm*), o qual é baseado no ato de busca humana. A direção da busca é determinada pelo comportamento egoísta, altruísta e pró-ativo do ser humano. O novo método foi empregado para ajustar estruturas e parâmetros de redes neurais artificiais. Os resultados obtidos com o novo método foram comparados com algoritmos *backpropagation* e algoritmos evolucionários para a classificação de padrões e aproximação de funções. Demonstraram que a nova técnica obteve desempenho superior ou pelo menos equivalente. Além disso, as redes obtiveram melhor capacidade de aprendizagem com menos ligações. Apesar da desvantagem de necessitar de alto poder computacional, os autores acreditam que o algoritmo possa se tornar um método revolucionário de treinamento de redes.

O estudo realizado nesta tese, com avaliações do comportamento de distintos algoritmos de treinamento, visou contribuir cientificamente para as pesquisas na área, além de encontrar o algoritmo que mais se adapte ao conjunto de dados fornecendo modelos altamente preditivos com respostas adequadas. Para tal, seis algoritmos de treinamento, que empregam quatro métodos de otimização foram avaliados e são detalhados no *Capítulo III, seção 4.3.2.3.2*. Todos os algoritmos de treinamentos citados na Tabela 2.1 que podem ser empregados para treinamento das RNA's, já estão implementados no pacote de redes neurais do programa MATLAB<sup>®</sup>.

## **2.4 REDES NEURAIAS COMO MODELO DE ANALISADORES VIRTUAIS EMPREGADOS EM COLUNAS DE DESTILAÇÃO**

Uma das ferramentas mais utilizadas no projeto e implementação dos analisadores virtuais são as redes neurais artificiais, pela capacidade de aprendizado e generalização de funções não lineares (SILVA, 2005). Este trabalho traz como objetivo principal, o desenvolvimento de um analisador virtual capaz de realizar a predição da composição do topo de uma coluna de destilação através do emprego de redes neurais artificiais. Vários trabalhos têm sido publicados na área de modelagem inferencial utilizando RNA's. A seguir, alguns trabalhos publicados nesta área são sumarizados.

Holcomb e Morari (1992) propuseram uma arquitetura de modelo híbrido onde é empregado o algoritmo PLS com uma rede neural *feedforward*. O PLS/neural é um aperfeiçoamento de métodos

que combinam o princípio PCA (*principal component analysis*) com redes *feedforward*, sendo particularmente útil no caso de supressão do ruído e de conjuntos de dados esparsos. O PLS tradicionalmente lida com conjuntos de valores mal-condicionados. Estas são algumas vantagens ressaltadas pelo autor dessa abordagem.

Em Vanleeuwen *et al.* (1994) foi desenvolvida uma modelagem para obtenção da octanagem de combustíveis a partir de dados de cromatografia. A rede utilizada possuía 1 camada intermediária com função de ativação tangente hiperbólica. Neste mesmo ano, MacMurray e Himmelblau (1994) descreveram uma modelagem de colunas de destilação empacotada com emprego de redes neurais artificiais, e forneceram um exemplo de modelagem complexa.

Ramchandran e Rhinehart (1995) desenvolveram modelos neurais de torres de destilação utilizando dados obtidos com o simulador HYSIM. Mejdeel e Andersson (1993) citam a utilização do simulador PROCESS para a obtenção de dados para um modelo inferencial baseado em redes neurais artificiais para uma coluna de destilação.

Em Baratti *et al.* (1995) foram desenvolvidos dois modelos inferenciais, baseados em RNA's para aplicação em controle e monitoramento da composição dos produtos de torres de destilação. Esses modelos foram implementados a partir de dados coletados em duas torres de destilação da refinaria de SARAS (Itália). Após vários testes, foram obtidos os arranjos ótimos das duas redes, topologia *feedforward* e configurações 8-2-4 e 7-2-2, respectivamente. Os autores concluíram que dois neurônios na camada intermediária foram suficientes para os problemas estudados. O treinamento foi efetuado com o algoritmo *backpropagation*.

Wang *et al.* (1996) desenvolveram um observador de estado utilizando redes neurais com hierarquia clusterizada, com lógica *fuzzy* para uma coluna de destilação de alta pureza e compararam os resultados com um filtro de Kalman. Os resultados mostram que a rede neural RBF (*radial basis function*) *fuzzy* distribuída usada para projetar um *soft-sensor* e obter dados de pureza dos produtos obteve desempenho adequado, se mostrando muito promissora para ser empregada no controle do processo. Os autores citam a construção de um novo tipo de rede neural *feedforward*.

Fileti *et al.* (1999) aplicaram um sensor virtual para uma coluna de destilação multicomponente em batelada de uma planta piloto empregando redes neurais *feedforward*, as quais utilizam como entradas para a rede basicamente as temperaturas de topo e base e compararam os

resultados de composição obtidos com medições laboratoriais por cromatografia gasosa.

Wang e Huihe (2000) utilizaram um observador em uma unidade de craqueamento catalítico (FCC). Neste trabalho os autores propõem a utilização de uma rede neural com arquitetura *feedforward*. Observaram que os resultados encontrados pela rede neural foram satisfatórios, porém sempre salientando as limitações de extrapolação da rede.

Em Barbosa (2002), pode-se averiguar o potencial da utilização de redes neurais na modelagem do refino de petróleo. Neste trabalho encontra-se ilustrado a inferência de produtos de destilação utilizando dados experimentais obtidos na refinaria REPAR, operada pela PETROBRAS S.A.

Moura (2003) e Meneguelo (2007) utilizaram uma rede neural com função de ativação *wavelet* desenvolvida por Claumann (2003), para simplificar a modelagem fenomenológica de uma coluna de destilação. Os resultados obtidos mostram que a rede utilizada reproduz de forma muito precisa o comportamento dinâmico do processo, podendo ser utilizada como preditor para fins de otimização e controle. Além de poder ser utilizada como um preditor, a rede provê informações sobre a relevância de cada variável do processo.

Zamprogna, Barolo e Seborg (2004) empregaram a técnica de redes neurais recorrentes que recebe informações de temperatura e fluxos da coluna para aplicação de um sensor virtual para uma coluna de destilação operando em regime de batelada. Neste trabalho foram usadas técnicas estatísticas para a redução do número de entradas através de Análise dos Componentes Principais (PCA) e efeitos de ruídos no conjunto de dados.

Zanata (2005), empregando redes neurais artificiais, desenvolveu um *soft-sensor* para uma coluna de destilação, capaz de estimar, em tempo real, a composição dos produtos da mesma, a partir de informações do tipo temperaturas e pressões em diversos pontos da coluna e vazões de entrada, de saída e de reciclo. No presente trabalho emprega-se metodologia similar a utilizada por Zanata (2005), porém validou-se experimentalmente o modelo no qual foram obtidos os dados simulados, diferentemente do autor, que restringiu seu trabalho às simulações. Outros diferenciais são referentes ao processo modelado, que é baseado em ações de calor distribuídas ao longo da coluna, e na proposta da utilização das inferências obtidas pelo sensor, em conjunto com ações das resistências elétricas em cada estágio, o que possibilita o

desenvolvimento de um controle inovador na unidade, possibilitando implementar uma estratégia de controle distinta.

Singh *et al.* (2005) desenvolveram um modelo rigoroso do processo, em estado estacionário, para simulação e geração de dados à rede. Os autores testaram uma rede neural do tipo *feedforward* com realimentação do erro por *backpropagation* em uma coluna de 15 pratos. O objetivo foi estimar a composição de líquido e vapor para uma mistura de cinco componentes e comparar com resultados do modelo. A rede possuía 17 entradas, sendo elas: as temperaturas dos quinze pratos mais a temperatura do refluxo e do refeedor e 10 saídas (5 composições para fase líquida e 5 para fase vapor). Era constituída de 3 camadas intermediárias, sendo composta 35 neurônios em cada uma. Os resultados apresentados pelos autores mostraram que as previsões feitas pela rede neural estão em boa concordância com os resultados da simulação, e que, utilizando a rede neural atinge-se o mesmo resultado com economia de tempo de 68,75% em relação ao modelo, sem perder a precisão da estimativa.

Singh *et al.* (2007), mesmos autores anteriormente citados, utilizaram o algoritmo de *Levenberg-Marquardt* (LM) no treinamento da rede para estimação da composição do destilado utilizando como entradas a pressão, vazão de refluxo e perfil de temperatura ao longo da coluna. Os autores comparam os resultados com outro algoritmo de treinamento, o *Steepest Descendent Back Propagation* (SDBP) para a mistura binária metanol/água (configuração da rede: 14-11-4) e também para mistura multicomponente (5) (configuração: 20-15-15-10) e observaram que para ambos os casos obtiveram um número de interações e erros significativamente reduzidos no treinamento e respostas da predição mais precisas com a utilização de LM do que os obtidos com o SDBP, os quais resultaram em saídas saturadas e que diferem bastante dos resultados desejados.

Linhares *et al.* (2007) e Araújo Junior (2007) aplicaram um sistema *fuzzy* em uma malha de controle em conjunto com uma RNA, também em uma unidade debutanizadora simulada no HYSYS<sup>®</sup>. O sistema *fuzzy* foi projetado para gerar, dinamicamente, os *sets points* dos controladores, a partir de medidas de vazão de refluxo e temperatura de fundo da coluna. A RNA estima a fração molar do *i-pentano* presente no GLP para controle de qualidade da coluna. A estrutura utilizada para a rede neural foi baseada no modelo NNARX, com 3 camadas intermediárias, 8 entradas e 1 saída. Os resultados obtidos demonstraram que o sistema de inferência em conjunto com o controlador *fuzzy* se mostrou válido para aplicação em processos de petróleo e gás natural.

Bahar e Özgen (2008) projetaram uma rede neural artificial (RNA) como estimador para prever os valores da composição dos produtos de uma coluna de destilação reativa em batelada, os quais possam ser empregados em um sistema inferencial de controle. Eles usaram a razão de refluxo da coluna de destilação como entrada para a RNA, bem como os valores de temperatura de alguns pratos. O desempenho da rede foi testada em malha aberta e as composições dos produtos puderam ser previstas. Desta forma, os autores concluíram que o estimador baseado em RNA tem potencial e que pode ser usado no controle de composições de produtos da torre de destilação em batelada.

Um sensor por software para estimativa da concentração *on-line* de etanol em destilarias foi desenvolvido por Osório *et al.* (2008), baseado em medidas de quatro temperaturas de colunas, tanto em escala laboratorial quanto industrial. O *soft-sensor* foi calibrado com dados do protótipo em laboratório e dados da unidade industrial. Consiste de um modelo neural desenvolvido com redes com treinamento não-supervisionado. Segundo os autores simplicidade e boa performance foram as métricas adotadas para testes com diferentes estruturas de redes. O sensor obteve boa performance com erros (RMSE) de predição de  $\pm 0,6\%$  v/v para escala de laboratório e  $\pm 1,6\%$  v/v para coluna industrial, mostrando-se como uma alternativa viável e com baixo custo, quando comparado com sensores físicos para análise da composição.

Patil e Nigam (2009) projetaram um estimador híbrido empregando redes neurais (RNA) e algoritmos genéticos (AG) para estimar a composição de uma coluna de destilação multicomponente através de medidas secundárias de temperaturas. Os autores entendem que as RNA's possuem potencial para lidar com problemas não lineares do processo. Por outro lado, acreditam que os AG sejam capazes de melhorar o desempenho da RNA através da introdução da evolução para os pesos de conexão. Com isso entendem que o estimador deve fornecer uma estimativa precisa e confiável, mesmo quando distúrbios não-mensuráveis estão presentes. O sensor híbrido obteve melhor desempenho do que o sensor com modelo somente com RNA. Porém, os autores consideram que ambos os estimadores são adequados para uso *on-line*, principalmente por exigirem baixo recurso computacional, comparativamente com outras técnicas.

Linhares (2010) simulou uma unidade de processamento de gás natural (UPGN) com uma coluna de destilação deetanizadora e outra debutanizadora com emprego do software HYSYS<sup>®</sup>. O objetivo do autor foi obter dados para desenvolver um sistema híbrido de inferência. O sistema proposto foi baseado em rede neural (MPL – com apenas 1

camada intermediária) treinadas com algoritmo Gradiente Conjugado Escalonado (SCG) e auxiliado com a técnica de Análise dos Componentes Principais (PCA) para estimar as frações molares de etano e pentano no GLP e a fração molar de propano no gás residual. O autor também desenvolveu uma estratégia simples para correção em tempo real do sistema de inferência. Os resultados mostraram que o sistema de inferência proposto demonstrou desempenho satisfatório, tendo funcionalidade para estimar as variáveis definidas.

Rogina *et al.* (2011) desenvolveram um sensor por software para estimar continuamente a qualidade da nafta leve de uma unidade de destilação, avaliando dois tipos de óleos. O sensor faz a estimativa da pressão de vapor da nafta leve (RPV), usando como parâmetros medidas de temperaturas, pressões e vazões. O modelo do sensor foi desenvolvido aplicando regressão linear múltipla e usando redes neurais, as quais foram testadas em três topologias: linear perceptron (LNN), multi-layer perceptron (MPL) e redes de função base radial (RBF), com diferentes algoritmos. Sendo os melhores resultados obtidos com redes MPL e RBF. Com isso, os autores sugerem que é possível o uso de sensores virtuais não lineares com modelos neurais para estimar a qualidade da nafta leve (RVP) como alternativa às análises de laboratório.

Nos últimos anos, grande parte das publicações que tratam do desenvolvimento de *soft-sensors* vêm empregando a tecnologia baseada em redes neurais artificiais. Este fato pode ser explicado devido aos bons resultados apresentados quando esta tecnologia é aplicada a sistemas complexos e não-lineares, como é o caso da destilação. Porém, a maioria destes estudos baseia-se em dados e estratégias simuladas e não são aplicados industrialmente. A real implementação de sensores virtuais com estratégias avançadas de controle em ambientes industriais, mesmo que vantajosa, ainda é lenta e requer uma série de avaliações prévias e um planejamento adequado das empresas. Assim sendo, ainda existe grande relevância dos estudos destes modelos empíricos com processos simulados, de modo a buscarem confiança e espaço dentro dos processos industriais.

Além disso, o grande diferencial deste projeto em relação ao controle de colunas convencionais é a estratégia que ele proporcionará que seja implementada. A partir do modelo empírico desenvolvido, representando o processo de uma unidade piloto de destilação diferenciada com ações distribuídas de controle baseadas em aquecimentos em pontos intermediários, torna-se possível realizar a inferência do produto de topo desta unidade, e utilizá-la com variável

controlada *on-line* atuando localmente em cada estágio da coluna, desenvolvendo, desta forma, um ambiente inovador de controle nesse tipo de processo.

## 2.5 CONCLUSÕES

Neste capítulo foram apresentadas algumas características do processo de destilação e seu controle enfatizando o emprego de modelos inferenciais. As topologias existentes e o princípio das redes neurais artificiais foram abordados mais detalhadamente. Na revisão destacaram-se alguns trabalhos publicados que utilizam modelo de redes neurais como analisadores virtuais. No caso de colunas de destilação, observa-se que o principal objetivo dos sensores por software foi de estimar composições e contribuir no controle da unidade.

Apesar das técnicas empíricas apresentarem a limitação de extrapolação, elas são amplamente utilizadas para geração de modelos, conforme foi constatado na literatura. No caso de utilizar modelos fenomenológicos rigorosos, como planta de colunas de destilação, existe o problema de se trabalhar com um sistema de equações algébrico-diferenciais (para os quais existe uma classe distinta de observadores), que geralmente demanda um elevado esforço e, conseqüentemente, muito tempo para seu desenvolvimento, além de não serem apropriados para aplicações em tempo real.

Desta forma, observa-se que cada ferramenta possui vantagens e desvantagens. A seleção da técnica mais adequada com o objetivo de operar a planta como um sensor por software deve, então, considerar a complexidade de implementação, os limites de operação, a fundamentação disponível na literatura e a aplicabilidade no processo, considerando o tempo de cálculo de cada técnica. Com base nos trabalhos publicados e no problema que se deseja resolver, entende-se que a técnica de Redes Neurais Artificiais seja a mais adequada e, pelos motivos expostos, foi a selecionada para ser aplicada neste trabalho na geração do modelo empírico do processo.

No próximo capítulo descrevem-se os materiais e métodos utilizados no desenvolvimento dos analisadores virtuais para monitoramento e controle da unidade piloto de destilação.

# Capítulo III

## MATERIAL E MÉTODOS

Este capítulo será dividido em duas seções distintas. Na primeira será apresentada uma breve descrição e caracterização da unidade piloto e indicadas as rotinas analíticas e condições operacionais que foram utilizadas nos experimentos. Na segunda, é descrita a metodologia adotada para identificação de sistemas para emprego de redes neurais. Esta etapa, esta subdividida em três seções, conforme é apresentado na representação esquemática em forma de diagrama de blocos da Figura 3.1, ilustrando a sequencia em que foi organizada a metodologia deste trabalho.

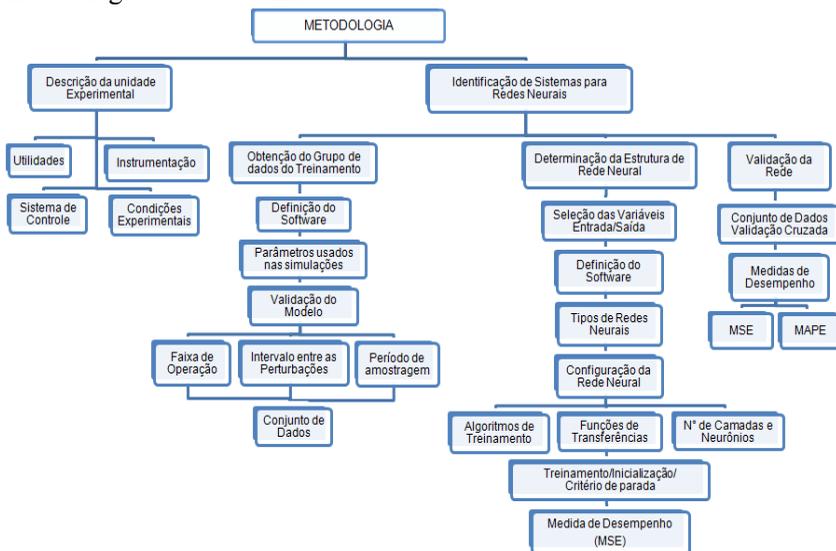


Figura 3.1. Representação esquemática geral em forma de diagrama de blocos, com organização da metodologia empregada.

A metodologia adotada envolve desde o desenvolvimento de um modelo da coluna de destilação, através do software de simulação comercial HYSYS<sup>®</sup> para gerar os dados utilizados como entradas/saída que serão utilizados para treinamento das redes, até a determinação da configuração das redes neurais empregadas neste trabalho e os critérios de desempenho aplicados para a escolha da rede melhor adaptada ao conjunto de dados e com boa capacidade de generalização.

### 3.1 INTRODUÇÃO

Os resultados obtidos com os ensaios experimentais, realizados anteriormente por Werle (2007) em uma coluna de destilação piloto, foram utilizados como base para emprego de parâmetros e ajustes necessários nas simulações, inicialmente desenvolvidas em estado estacionário. Em seguida, migrou-se para o modo dinâmico do software. Estes dados experimentais foram essenciais para validação do modelo desenvolvido em distintas situações de transientes de operação da unidade, depois de aplicadas algumas perturbações na carga. Através dos dados obtidos destas simulações foi possível treinar uma rede neural empregada como sensor virtual para inferência da composição de etanol no topo da coluna, com a finalidade de realizar, além do monitoramento da planta, empregar esta informação para desenvolver um ambiente de controle avançado.

Os controladores e malhas de controle utilizadas tanto no modelo quanto experimentalmente, bem como a forma com que foram conduzidos os testes para obtenção dos dados experimentais serão descritos brevemente neste capítulo. Maiores detalhes dos resultados experimentais podem ser obtidos em Werle (2007).

A coluna de destilação piloto e os demais equipamentos estão instalados no Laboratório de Controle de Processos do Departamento de Engenharia Química e Engenharia de Alimentos da Universidade Federal de Santa Catarina. Os recursos necessários à manutenção da unidade piloto são fornecidos pelo Programa: PRH-ANP/MCT N<sup>o</sup> 34: Formação de Engenheiros nas Áreas de Automação, Controle e Instrumentação para a Indústria do Petróleo e Gás.

### 3.2 DESCRIÇÃO DA UNIDADE EXPERIMENTAL E TESTES

A unidade piloto de destilação em operação foi construída utilizando os mesmos equipamentos e ferramentas de configuração em software desenvolvidos para aplicação industrial. A unidade piloto é representada pela Figura 3.2.



Figura 3.2 Unidade piloto utilizada para obtenção de dados experimentais e parâmetros para simulações.

A unidade de destilação foi construída de forma modular (aço inox 304). Cada módulo (com 0,15 m de altura e 0,20 m de diâmetro) contém um prato perfurado (diâmetro de 0,006 m, com passo triangular), conforme ilustra a Figura 3.4a. Os valores da altura e comprimento do vertedouro são 0,03 e 0,10 m, respectivamente. A coluna é composta de 13 pratos, sendo a alimentação no prato 4 (quatro).

A planta piloto opera em regime contínuo, como ilustrado na Figura 3.3, e possui uma resistência elétrica em cada prato para implementação do controle distribuído (diferencial da unidade e que será descrito detalhadamente nesta seção) com potência de 3,5 kW, detalhado na Figura 3.4b. Dois trocadores de calor a placas são utilizados na unidade piloto: um como refeedor, e outro na corrente de alimentação,

para garantir a temperatura da mesma. O condensador utiliza água como fluido refrigerante, proveniente de uma torre de refrigeração. A coluna possui um acumulador com um volume de 60 L, responsável por receber o vapor condensado e fazer a recirculação desse líquido para garantir a razão de refluxo, aumentando, desta forma, a eficiência da separação.

O processo se torna contínuo com a utilização de um tanque pulmão, onde se armazenam 600 L de mistura, que recebe o produto de topo e de fundo, e é utilizado para prover a alimentação da coluna.

Sensores do tipo Pt100 a 3 fios são utilizados para monitorar a temperatura em todos os estágios de equilíbrio, na corrente de alimentação e de *boilup*, bem como no tanque acumulador. Sensores de pressão manométrica na base e no topo também estão instalados.

Um esquema geral de todo processo, com os equipamentos que o compõem, pode ser visualizado na Figura 3.3.

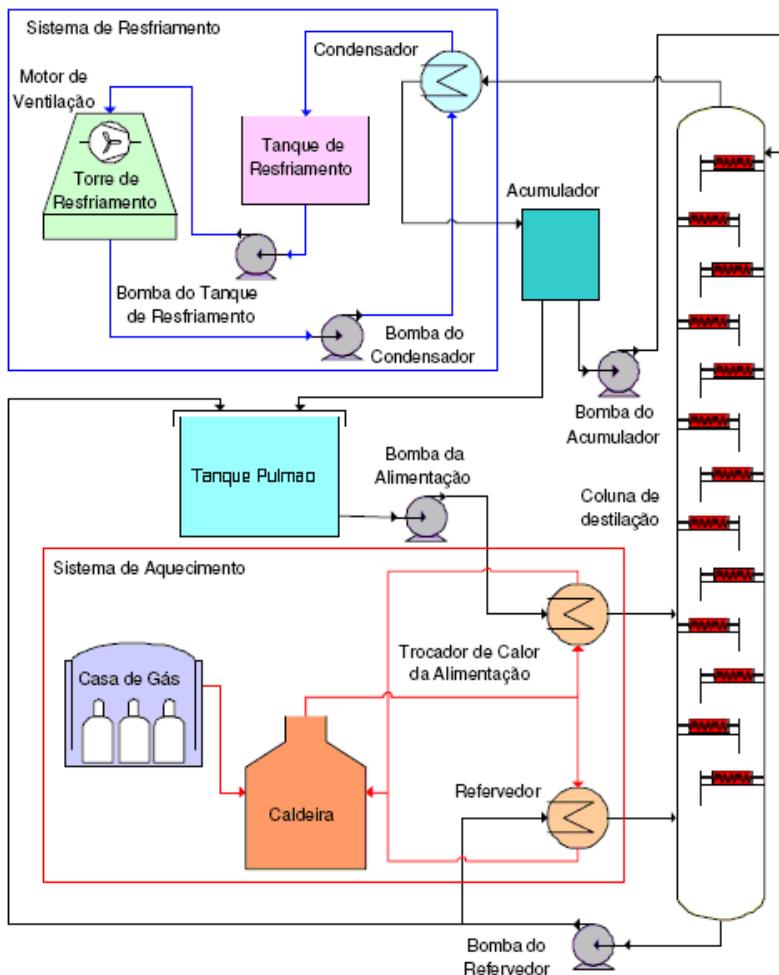


Figura 3.3 Esquema geral da unidade de destilação.

Fonte: Pasetti (2005).

As resistências instaladas em cada prato para a implementação do controle distribuído são ilustradas na Figura 3.4b. Teleken (2009) realizou um estudo no qual avaliou a influência da resistência elétrica sobre o prato perfurado. Sendo que o autor concluiu que não ocorre influência significativa sobre a fluidodinâmica do processo com o emprego de tal dispositivo físico.

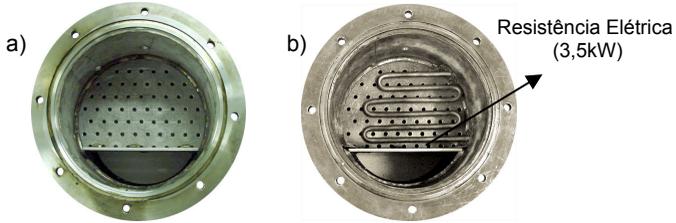


Figura 3.4 Vistas dos módulos com detalhamento do prato. (a) Convencional, (b) Com resistência elétrica, para o aquecimento distribuído.

Essas resistências permitem a implementação de malhas de controle em vários pratos da coluna de destilação para estudar a influência do aquecimento distribuído ao longo da coluna. Convém ressaltar que, ao invés de realizar o aquecimento por meio da resistência, poder-se-ia fazê-lo através do uso de trocadores de calor ou condensadores, para realizar a troca de calor. Um esquema geral do interior da coluna com apresentação dos módulos de alimentação e do topo é apresentado na Figura 3.5.

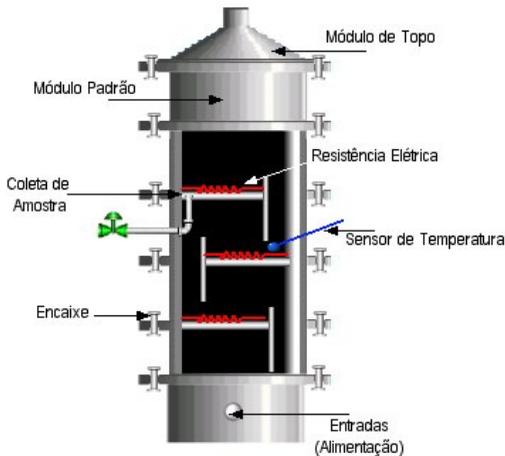


Figura 3.5 Representação geral do projeto

Fonte: Marangoni (2005).

Cada módulo possui um orifício para medição de temperatura, outro para a coleta de amostra e um terceiro onde foi colocada a resistência elétrica.

### 3.2.1 Instrumentação

A instrumentação da coluna de destilação foi implementada com tecnologia FOUNDATION *Fieldbus*. Estes instrumentos são capazes de executar, distribuídamente e de forma dedicada, todo o controle da coluna de destilação (CONSTANTINO, 2010). A FOUNDATION *Fieldbus* (FF) foi projetada especificamente para dispositivos de campo com parâmetros predefinidos e comandos para funções, tais como identificação, controle, diagnóstico, calibração e comissionamento (PIN CETI, 2004; LEE e HONG, 2011).

A unidade experimental possui 15 instrumentos inteligentes (Smar – *Fieldbus*), um controlador lógico programável (LC700), dois inversores de frequência e quatro variadores de potência (com possibilidade de instalar um variador para cada prato) para acionamento das resistências e os sinais de liga e desliga das bombas. Na Figura 3.6 é apresentada, resumidamente, a seqüência da aplicação do controle distribuído.

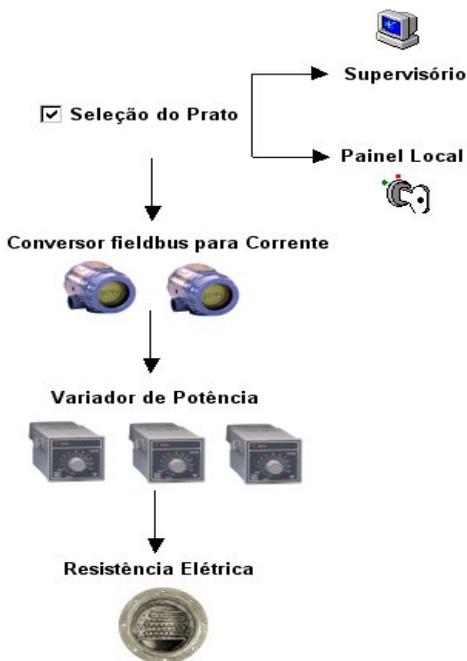


Figura 3.6. Sequencia de operações para aplicação do controle distribuído.

Na unidade experimental, o LC700 é interconectado com o DFI302 (*Fieldbus Universal Bridge*). O DFI302 é um componente de *hardware* multifunção integrado ao sistema, sendo o elemento chave na arquitetura distribuída dos sistemas de controle de campo, executando a maioria das funções exigidas pelo sistema de controle. Portanto, o DFI é o componente que controla as ações relacionadas ao sistema *fieldbus*, sendo capaz de gerenciar, monitorar, controlar, manter e operar a planta. Maiores detalhes da instrumentação da unidade e da forma de comunicação e controle dos instrumentos podem ser obtidos em Pasetti (2005), Marangoni (2005) e Werle (2007).

### 3.2.2 Sistemas de controle

A configuração das malhas de controle implementadas na unidade experimental foi definida por Marangoni (2005) com base nos estudos de Nooraii *et al.* (1999).

A coluna é composta pelas seguintes malhas de controle:

1. Controle do nível da base através do ajuste da corrente de produto de fundo;
2. Controle do nível do acumulador através da manipulação da vazão do destilado;
3. Controle da vazão da alimentação através do ajuste da vazão desta corrente;
4. Controle da temperatura de alimentação através do ajuste da vazão de vapor no trocador de calor deste estágio;
5. Controle da temperatura do último prato através da manipulação da vazão de refluxo;
6. Controle da temperatura do refeedor através do ajuste da vazão de vapor no trocador de calor deste estágio.
7. Controle da temperatura dos estágios intermediários da coluna através da manipulação de um variador de potência acoplado a uma resistência elétrica.

Na Figura 3.7 são representadas as estruturas das malhas de controle implementadas.

O controle da temperatura dos estágios intermediários é realizado objetivando o emprego do controle distribuído. Com a inclusão desta malha pretende-se compensar mais rapidamente perturbações que afetem o perfil de temperatura. Este controle é realizado através de um variador de potência, que é acionado dissipando potência pela resistência elétrica situada no prato e ajustando sua temperatura de maneira a mantê-la no valor desejado. Pode-se observar

que existem apenas quatro controladores de temperaturas para os 13 estágios da coluna, pois a unidade experimental possui quatro variadores de potência instalados. Desta forma, apenas 4 resistências podem ser manipuladas simultaneamente. As resistências são escolhidas pelo operador de acordo com um critério de seleção pré-definido. Desta forma, cada atuador destina-se a manipular uma, dentre duas (ou três) resistências que lhes são permitidas. Para efetivo controle de temperatura de todos os pratos simultaneamente, basta apenas a aquisição de mais variadores de potência.

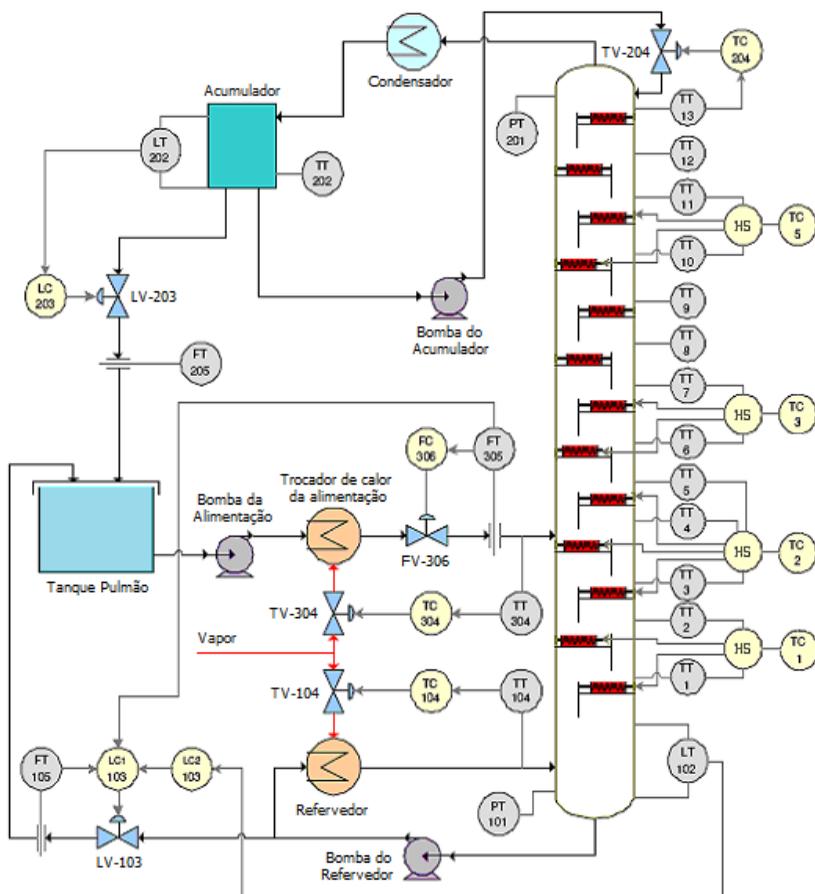


Figura 3.7 Estrutura das malhas de controle da coluna de destilação.

Fonte: Pasetti (2005).

As *tag's* referentes as malhas de controle implementadas na unidade piloto de destilação, ilustradas na Figura 3.7, são apresentadas na Tabela 3.1.

Tabela 3.1 Lista das *Tag's* das malhas de controle da unidade.

<i>Tag's</i>	<i>Denominação/Função</i>
TV104	Válvula de vapor do refeedor
TV204	Válvula de refluxo
TV304	Válvula de vapor do trocador de calor da alimentação
LV103	Válvula da corrente de produto de fundo
LV203	Válvula da corrente de produto de topo
FV306	Válvula na corrente de alimentação
LT102	Transmissor de nível da base da coluna
LT202	Transmissor de nível do tanque acumulador
FT105	Transmissor de vazão do produto de fundo
FT205	Transmissor de vazão do produto de topo
FT305	Transmissor de vazão da corrente de alimentação
PT101	Transmissor de pressão da base da coluna
PT201	Transmissor de pressão do topo da coluna
LC203	Controlador de nível do tanque acumulador
LC <sub>1</sub> 103	Controlador de nível da base da coluna (escravo)
LC <sub>2</sub> 103	Controlador de nível da base da coluna (mestre)
TC104	Controlador da temperatura da base da coluna
TC204	Controlador da temperatura do topo da coluna
TC304	Controlador da temperatura da corrente de alimentação
TC1	Controlador da temperatura do prato 1 ou 2 através do variador de potência da resistência
TC2	Controlador da temperatura do prato 3 ou 4 ou 5 através do variador de potência da resistência
TC3	Controlador da temperatura do prato 6 ou 7 através do variador de potência da resistência
TC5	Controlador da temperatura do prato 10 ou 11 através do variador de potência da resistência
FC306	Controlador de vazão da corrente de alimentação
TT104	Transmissor da temperatura da base ( <i>boilup</i> ) da coluna
TT202	Transmissor da temperatura do líquido do tanque acumulador
TT304	Transmissor da temperatura da corrente da alimentação
TT1 a	Transmissores da temperatura dos pratos da coluna
TT13	

### 3.2.3 Condições experimentais e operacionais

Os experimentos realizados por Werle (2007) foram conduzidos com uma mistura binária água/etanol com composições volumétricas iniciais de etanol na alimentação em torno de 10% (v/v). A temperatura da alimentação foi controlada para entrar na coluna na temperatura de equilíbrio do prato (em torno de 90°C, porém abaixo da temperatura de bolha da mistura). A vazão de alimentação em condições normais de operação foi de 300 L.h<sup>-1</sup> e as pressões da base e do topo variaram entre 1,4 e 1,1 bar, respectivamente. A vazão de destilado ficou em torno de 10 L.h<sup>-1</sup> com valor médio de razão de refluxo (R/D) igual a 6.

Durantes os experimentos as condições da alimentação da coluna foram pouco variadas. Em geral trabalhou-se apenas com duas condições distintas de alimentação, tanto em termos de temperatura, fração de etanol e vazão (uma na condição nominal de operação – detalhada acima, e outra após a realização de uma perturbação em cada uma destas variáveis separadamente). Em termos de análise de composição, as mesmas somente eram realizadas para verificação do alcance do estado estacionário nas extremidades da coluna, com medidas *off-line*. Com os experimentos realizados, foi constituído um banco de dados com perfis experimentais de todas as variáveis que possuíam sensores para aquisição. Porém, estes dados dinâmicos não são suficientes para treinamento e geração de padrões das redes neurais empregadas posteriormente, visto a pequena faixa operacional. Este foi o motivo principal para o uso de um simulador.

## 3.3 METODOLOGIA PARA IDENTIFICAÇÃO DE SISTEMAS

A identificação de sistemas utilizando redes neurais consiste basicamente dos seguintes passos, sumarizado por Claumann (2003):

- Obtenção do grupo de dados de treinamento: aplicando um conjunto de perturbações ao processo que se deseja modelar, gera-se um conjunto de dados para treinamento. Os alvos são as saídas do processo no tempo discreto (k+1). As entradas da rede são os valores atrasados das saídas e das perturbações, ou seja, correspondem aos tempos discretos (k),(k-1),(k-2).

- Determinação da melhor estrutura da rede: esta etapa é relativa tanto ao estudo do conjunto de variáveis de entrada que melhor identifica o processo, quanto a arquitetura da rede propriamente dita e

compreende também a etapa do treinamento. Um critério simples para indicar o melhor conjunto seria aquele que produzisse um menor erro de predição com o menor número de variáveis de entrada. Para determinação da arquitetura da rede não existem regras bem definidas.

- Validação: o procedimento mais comum para validar modelos é utilizar um novo grupo de dados (dados teste) relacionando as entradas às saídas. Pode ser utilizado para verificar a eficiência da rede quanto a sua capacidade de generalização durante o treinamento, podendo ser empregado como critério de parada do treinamento. O desempenho da rede deve ser avaliado com relação a um grupo de dados diferente do utilizado no treinamento (validação cruzada).

Os procedimentos adotados em cada uma destas etapas serão detalhados no transcorrer deste capítulo.

### 3.3.1 Obtenção do grupo de dados de treinamento

Ferramentas computacionais, como softwares para a simulação de processos, são extremamente úteis para obtenção de conjuntos de dados, uma vez que possibilitam resultados muito próximos do real com relativa rapidez. Atualmente, com a elevada capacidade de cálculo e baixo custo, os programas de simulação são altamente difundidos, tanto na análise como no projeto de processos químicos. Adicionalmente ao projeto, a simulação é aplicada ao treinamento de operadores, otimizações de processos, estudos de sensibilidade e, mais recentemente, no apoio ao controle avançado de processos (ONDREY, 2005). Em Seider *et al.* (2003) pode ser encontrada uma boa revisão contendo uma visão geral de simulação de processos com os simuladores HYSYS® e ASPEN PLUS®.

#### 3.3.1.1 Software usado para construção do modelo

Dentre os simuladores comerciais mais conhecidos destaca-se o HYSYS®, o qual apresenta uma linguagem de simulação avançada orientada a objetos. Premkumar e Rangaiah (2009) denominam este software como sendo um rigoroso simulador de processo. O software dispõe de uma biblioteca de componentes padronizados que podem ser interligados e ajustados conforme o processo.

Neste trabalho o software utilizado para simulação, tanto em estado estacionário quanto dinâmico, foi o HYSYS® versão 7.0. O simulador utiliza um método de resolução iterativo amplamente utilizado em fracionadoras. Este algoritmo foi desenvolvido por Boston e Sullivan (1972). O método consiste em utilizar equações para o

cálculo da constante de equilíbrio e da entalpia para resolver as equações de balanço de massa, energia e equilíbrio na iteração interna e, após atualizar o cálculo da constante de equilíbrio e da entalpia em uma iteração externa. Os balanços de massa, energia e composição são resolvidos simultaneamente.

A coluna de destilação utilizada neste projeto pode ser encontrada no diretório de demonstração do “*Simulador Dinâmico HYSYS*”. Empregou-se o simulador comercial para geração de dados de temperatura e composição, em regime transiente. Foram realizadas perturbações na vazão, temperatura e composição, todos aplicados na corrente de alimentação da coluna, sendo que os dados obtidos no período transiente foram usados para o treinamento da rede neural dinâmica.

A eficácia do software em representar uma coluna de destilação experimental será verificada posteriormente com a validação dos dados simulados, na qual consiste na comparação dos resultados simulados no software com os experimentais de perfis de temperatura na base, no prato 2 e no topo da coluna. Seus resultados também serão avaliados de forma global, analisando as respostas obtidas pelo modelo, comparativamente com o comportamento já conhecido e amplamente disponível em literatura, da mistura utilizada. Em Rocha (2003) é exposto o balanço de massa no estado transitório que é usado pelo software HYSYS<sup>®</sup> para uma coluna de destilação.

Atualmente existem vários programas dinâmicos sendo utilizados em processos específicos, equipamentos individualizados e, mais recentemente, programas comerciais para uso geral. Langerhorst (2000) relata os principais simuladores dinâmicos existentes no mercado, alguns deles são ilustrados na Tabela 3.2.

Tabela 3.2. Principais softwares comerciais de simulação dinâmica.

<i>Finalidade</i>	<i>Programa</i>	<i>Instituição</i>
<i>Simuladores dinâmicos (operações unitárias)</i>	BATCHFRAC	Aspentech
	DISBATHC	Calsep
	BARESI	Henkel
	ELCOPOLY	Henkel
	KOSI	Henkel
<i>Simuladores dinâmicos de processo</i>	<b>HYSYS</b>	<b>Aspentech</b>
	ASPEN Dynamics	Aspentech
	DIVA	Uni Stuttgart
	SPEEDUP	Prosys
	OPTSIM	Linde
	SuperPro	Intelligent, Inc
	UniSim	Honeywell
	WSim	Open source
	DWSIM	Open source
ENZO	COPPE/Petrobrás	

Fonte: Adaptado de Langerhorst (2000).

Apesar da existência de diversos programas de simulação dinâmica, existem particularidades inerentes a determinados processos que não são contempladas em versões comerciais. Para estes casos, a alternativa recomendada é o desenvolvimento de programas específicos e orientados para a simulação exclusiva de determinado processo.

### 3.3.1.2 Parâmetros usados nas simulações

Nas simulações foi utilizado um sistema binário, composto por água e etanol, o mesmo usado experimentalmente. A escolha do modelo termodinâmico é fundamental, pois se refere à modelagem do equilíbrio de fases que pode apresentar desvios da idealidade, e o modelo deve prever tal situação. O software HYSYS<sup>®</sup> possui modelos de equação de estado e de coeficiente de atividade. Em geral, para a maior parte das destilações um modelo termodinâmico baseado em equações de estado para fase líquida e vapor já é suficiente. Porém, quando existem não-linearidades, alguns modelos que utilizam o cálculo do coeficiente de atividade para fase líquida são necessários.

A coluna experimental piloto é simulada utilizando-se os 13 estágios teóricos, representando os pratos de uma coluna real, e mais o condensador e o refeedor que são equivalentes a outros dois estágios.

Na Tabela 3.3 são listadas algumas condições de operação utilizadas na simulação da coluna, quando empregado o programa de simulação de processo HYSYS<sup>®</sup>. As condições de operação utilizadas nas simulações foram as mesmas empregadas nos testes experimentais.

Tabela 3.3. Condições de operação empregadas na simulação estática e dinâmica.

<i>Variável</i>	<i>Valor (SI)</i>
Fração volumétrica de etanol na alimentação	0,10
Fração volumétrica de etanol no topo	0,865
Temperatura da alimentação	90 °C
Temperatura da base	101 °C
Vazão da alimentação	300 L.h <sup>-1</sup>
Pressão na base	1,4 bar
Queda de pressão ao longo da coluna	0,3 bar
Nível de líquido da base	35%
Nível de líquido do condensador	70%

O modelo termodinâmico para a fase líquida empregado foi o UNIQUAC, baseado nas recomendações contidas no software, sendo indicado para processos com não-linearidade e não-idealidade. O simulador foi utilizado inicialmente no seu modo estático e depois migrou-se para o modo dinâmico.

As simulações realizadas para contemplar o escopo deste projeto podem ser divididas em três etapas:

- Simulações estáticas – realizadas inicialmente para obter conhecimento do software e verificar se suas respostas são coerentes com comportamentos conhecidos da mistura em questão;
- Simulações Dinâmicas em Malha Fechada – realizadas para reproduzir a operação da coluna piloto, incluindo riqueza de detalhes em relação a equipamentos e malhas de controle, com o objetivo principal de validação do modelo com dados experimentais da planta piloto. Foram realizadas simulações adicionais com esta configuração para gerar dados do processo em malha fechada, com a intenção de desenvolver uma rede neural que possa ser usada como um analisador virtual *on-line* para monitoramento da composição de topo, com o sistema de controle atual da unidade piloto.
- Simulações Dinâmicas em Malha Aberta – realizadas para obtenção do conjunto de dados de entradas/saídas, conforme exigido para a formação de redes neurais artificiais. A rede desenvolvida

com este conjunto de dados será usada como sensor por software, onde sua informação poderá ser empregada posteriormente para projeto de controles inferenciais na unidade piloto de destilação, buscando a otimização da planta com técnicas avançadas de controle.

O intervalo de aquisição dos dados simulados nas simulações estáticas e nas simulações em malha fechada para validação foi de 1 segundo. Este mesmo intervalo também foi usado na aquisição dos dados experimentais com o intuito de reproduzir instantaneamente os efeitos da perturbação.

Nas simulações em malha aberta, o intervalo foi determinado pela avaliação da menor constante de tempo do processo em malha aberta, entre as variáveis avaliadas como possíveis candidatas a entradas/saídas da rede neural. Foi empregado o método de identificação proposto por Sundaresan e Krishnaswamy (1977) que utiliza dois tempos ( $t_1$  e  $t_2$ ) estimados a partir da curva de resposta à perturbação, correspondendo a 35,3% e 85,3% do tempo de resposta, respectivamente. O tempo de amostragem usado nesta etapa foi de 10 segundos e sua obtenção é detalhada no *Capítulo IV, seção 4.2.2.6*.

### 3.3.1.3 Sistemas de controle da unidade simulada

Neste trabalho é abordada a aplicação de uma nova configuração de controle baseada em ação distribuída com aquecimentos em pontos intermediários de uma coluna de destilação. A estratégia é avaliada do ponto de vista da minimização dos transientes da unidade bem como a avaliação do fornecimento e distribuição de calor desta nova concepção. Para tal, o modelo foi desenvolvido buscando representar esta nova concepção.

A diferença entre a configuração convencional de controle e a configuração com aquecimento distribuído é ilustrada de forma geral na Figura 3.8. Na unidade piloto foram ajustados controladores PID (multivariável com desacoplamento estático) para as malhas da temperatura da base e do último prato, caracterizando o controle convencional.

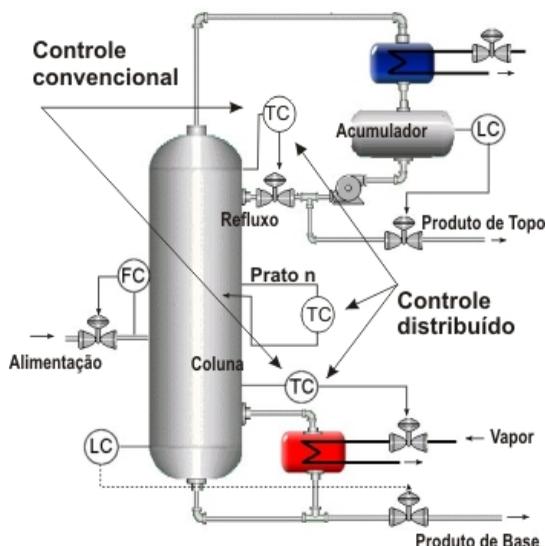


Figura 3.8. Diferença entre o controle convencional e o distribuído.

Com os resultados experimentais foi realizada uma comparação do desempenho da configuração convencional com a abordagem distribuída (controle na base, no topo - usando o mesmo ajuste que o aplicado na estratégia convencional – adicionando a malha de controle de temperatura do prato 2 com ajuste descentralizado). Nas simulações realizadas em malha fechada, o objetivo foi reproduzir tais comparações, com o intuito de validar o modelo. Desta forma, foram desenvolvidos 2 modelos do processo de destilação: 1) modelo do processo convencional e, 2) modelo do processo com abordagem distribuída.

O controle da temperatura do prato 2 foi realizado objetivando a implementação do controle distribuído. Com a inclusão desta malha pretende-se rejeitar mais rapidamente perturbações que afetem o perfil de temperatura, além da economia energética. A arquitetura das malhas de controle aplicadas no software HYSYS<sup>®</sup> são similares às implementadas na unidade experimental, com exceção do controle de composição volumétrica de etanol do topo, o qual é realizado em cascata com o controle de vazão de refluxo. Conforme mencionado na descrição da unidade experimental, este controle é realizado indiretamente através da temperatura do topo. No modelo não foi aplicado o controle de temperatura do topo, pois com essa configuração de malha não foi atingida estabilidade do processo. Com o emprego desta configuração

no topo, verificou-se que o comportamento do processo simulado é semelhante ao experimental, pois o objetivo de controle é o mesmo, apenas em um dos casos é realizado inferencialmente.

Existem muitas maneiras de controlar a qualidade do topo. A temperatura tem sido a variável medida mais utilizada no monitoramento e controle da separação, pois os sensores são simples, baratos e seguros. Contudo, para muitas colunas, o controle de temperatura não satisfaz os objetivos da separação. Nesses casos, a medição direta da composição deve ser considerada.

Outra diferença entre o experimental e o modelo desenvolvido, foi que no software não houve necessidade de implementar a malha de controle da temperatura da alimentação, pois este valor pode ser fixado na condição de operação do processo. Já no experimental, o controle é realizado com a manipulação da válvula de vapor/água desta corrente.

A configuração das malhas de controle desta nova concepção, implementadas no simulador HYSYS<sup>®</sup> com controle distribuído aplicado no prato 2, é descrita abaixo.

1. LIC-101: Controle do nível do refeedor através do ajuste da corrente de produto de fundo;
2. LIC-100: Controle do nível do acumulador através da manipulação da vazão do destilado;
3. FIC-102: Controle da vazão da alimentação através do ajuste da vazão desta corrente;
4. TIC-101: Controle da temperatura do refeedor através do ajuste do calor no trocador de calor deste estágio.
5. FIC-100: Controle da vazão de refluxo através da manipulação da vazão desta corrente;
6. XIC-100: Controle da fração volumétrica de etanol da corrente de destilado em cascata com o controle de vazão de refluxo, através da manipulação do *set point* do controlador FIC-100 (escravo);
7. TIC-R2: Controle da temperatura do prato 2 da coluna através da manipulação de uma fonte de calor.

Como nos experimentos, nas simulações com controle convencional, também foram utilizadas as mesmas malhas e ajustes dos parâmetros dos controladores utilizadas na simulação com o controle distribuído, com exceção da malha de controle de temperatura do prato 2, que não foi empregada.

As malhas descritas acima são controladas por meio de PID, cujo algoritmo é residente no programa HYSYS<sup>®</sup>. A maioria dos controladores foi sintonizada utilizando-se o módulo *autotuner*

*parameters* do software. Porém, para alguns esta sintonia não obteve resultados satisfatórios. Para a obtenção dos parâmetros destes, aplicou-se o Método da Sensibilidade Limite Modificado, proposto por Ziegler e Nichols (1942) (SEBORG *et al.*, 1989), baseado no ajuste de uma malha fechada.

Foram realizadas várias simulações dinâmicas onde foram introduzidas perturbações ao sistema, gerando-se um conjunto de dados em regime transiente. Foram simuladas situações distintas de condições de entrada na coluna com vazões, temperaturas e composições variadas, analisando-se as variáveis de saída do processo. Estas condições serão descritas na próxima seção.

Nesta fase do projeto, a etapa de simulações possui um caráter fundamental no desenvolvimento do trabalho, pois é nesta etapa que importantes decisões são tomadas, com o intuito de representar o processo de forma mais fidedigna possível. Neste sentido, o HYSYS<sup>®</sup> foi usado como ferramenta estratégica, pois o software utiliza modelos termodinâmicos apropriados, balanços rigorosos e cálculos para a mistura em questão.

#### 3.3.1.4 Validação do modelo

Em uma primeira etapa, elaborou-se uma simulação estática padrão, usando parâmetros e variáveis de processo idênticos aos obtidos experimentalmente, sendo a mesma validada através da comparação do perfil de temperatura experimental dos pratos, em estado estacionário. Nesta simulação foram aplicadas várias perturbações nas variáveis da carga. Foi realizada uma análise dos resultados obtidos nas simulações para verificar as respostas obtidas pelo software HYSYS<sup>®</sup>. Nesta etapa foram avaliadas conceitualmente as respostas do modelo, verificando se ele tem a capacidade de descrever de forma coerente alguns comportamentos conhecidos do processo de destilação da mistura binária etanol/água, vastamente descritos na literatura.

Na etapa seguinte, migrou-se para a plataforma dinâmica do HYSYS<sup>®</sup>, onde foi desenvolvido o modelo dinâmico, acrescentando as malhas de controle do processo. No simulador realizaram-se perturbações na temperatura e vazão da corrente de alimentação. Para validar o modelo, as respostas do processo frente às perturbações (analisadas em termos de temperaturas da base, topo e prato 2 e suas respectivas variáveis manipuladas) foram confrontadas com as respostas experimentais, quando submetidas as mesmas perturbações na carga.

Os perfis de temperatura experimentais foram utilizados para validação das simulações visto que, no período transiente, não se tinha coleta de amostras de produtos de base e topo, para realização de análises de composição. Historicamente a temperatura tem sido a variável medida mais utilizada no monitoramento (inferência) e controle da separação, pois os instrumentos de medição de temperatura são simples, tem resposta rápida, são relativamente baratos e seguros, possuem alta padronizações e não requerem um complicado sistema de amostragem. Segundo Kalid (1999), em torno de 75% das colunas de destilação utilizam a temperatura para inferir a qualidade da separação. Baseado nestas constatações, considera-se que o método de validação das simulações, através dos perfis de temperatura utilizados neste trabalho, é consistente e adequado, devido à escassez de dados experimentais de composição.

#### 3.3.1.5. Faixa de operação e perturbações realizadas

No presente trabalho, um simulador de processo foi empregado para gerar dados dinâmicos para posterior treinamento de redes neurais. Desta forma, variáveis manipuladas da corrente de alimentação do processo, como composição volumétrica de etanol ( $X_a$ ), vazão ( $Q_a$ ) e temperatura ( $T_a$ ) são perturbadas para obter conjunto de dados de entradas/saídas, conforme exigido para a formação de redes neurais artificiais. A abrangência das análises de sensibilidade do modelo foi determinada pelas variáveis manipuladas na carga, ou seja, as mesmas foram variadas ao máximo, até a simulação apresentar problemas de convergência.

Em virtude das diferenças operacionais observadas da coluna em malha fechada e aberta, a faixa de operação foi diferenciada para cada modo. O desempenho do HYSYS<sup>®</sup> em malha fechada (submetido a perturbações com grande magnitude) foi considerado adequado, sendo possível operar a planta na faixa de operação desejada. Já em malha aberta, a faixa de operação teve que ser restringida, pois quando aplicadas perturbações de graus com grande magnitude o processo migrava para situações que as integrações e os cálculos não convergiam.

A faixa de operação foi limitada a convergência do modelo simulado desenvolvido. No entanto, as faixas de trabalho empregadas, nos dois casos, podem ser consideradas representativas em relação aos pontos de operação da unidade piloto experimental.

Os dados obtidos deste processo devem ser representativos de toda amplitude na qual se deseja que o sensor venha a operar e devem

ser em número suficiente para que a rede neural consiga gerar os padrões de treinamento necessários. Visando obter dados que satisfaçam as exigências anteriores, aplicaram-se perturbações nas variáveis de entrada da coluna, usando combinações entre as variáveis: vazão, temperatura e composição volumétrica de etanol, dentro dos intervalos representados nas Tabelas 3.4 e 3.5 para a coluna em malha fechada e aberta, respectivamente.

Para o processo operando em malha fechada, a multiplicação dos valores do número de variações ( $Q_a, T_a, X_a$ ) da Tabela 3.4, resulta em 125 combinações possíveis, para distintas condições de operação da coluna. Foram realizadas o dobro deste número em simulações (250) com o software HYSYS<sup>®</sup>, visto que foram aplicadas as perturbações, e em seguida, após o sistema de controle atuar, estas eram retiradas. Cada perturbação é designada neste trabalho como sendo uma simulação no software. Vale ressaltar que, no caso em malha fechada, o processo sempre retorna ao ponto de operação, em virtude da ação de controle, onde ocorrem ajustes nas variáveis manipuladas para rejeitar as perturbações.

Seguindo o mesmo procedimento, porém em condições diferentes daquelas nas quais foram gerados os dados para treinamento da rede (ou seja, diferentes  $Q_a, T_a, X_a$  - dentro da faixa estipulada) foi usada a função “*rand*” do MATLAB<sup>®</sup> para gerar um conjunto aleatório de dados. A partir destes valores foram realizadas mais 30 simulações (ou 60 considerando aplicação e retirada) para obtenção de um conjunto de dados de validação para a rede em malha fechada, representando um percentual em torno de 25% dos dados de treinamento.

O processo se repete para as simulações realizadas em malha aberta. Na Tabela 3.5, a multiplicação dos valores do número de variações resulta em 54 condições distintas de operação da coluna simulada. Desta forma, foram realizadas 108 simulações, aplicando as perturbações e após retirando-as (retornando ao estado estacionário anterior). O diferencial do processo em malha aberta é que agora o sistema migra para estados estacionários distintos.

Para gerar o conjunto de dados de validação, novamente, foi usada a função “*rand*” do MATLAB<sup>®</sup> com valores intermediários das variáveis manipuladas da alimentação, gerando combinações distintas das usadas para gerar os dados de treinamento. A partir destes valores foram realizadas mais 15 simulações (ou 30 considerando aplicação e retirada da perturbação) com o intuito de realizar a validação da rede. Estas novas simulações também representam um percentual em torno de 25% dos dados de treinamento.

Tabela 3.4. Intervalos adotados para as combinações usadas nas simulações em malha fechada.

<i>Variáveis</i>	<i>Intervalos</i>	<i>Valores usados</i>	<i>Nº variações</i>
Vazão (L.h <sup>-1</sup> )	$250 \leq Q_a \leq 450$	250/300/330/375/450	5
Temperatura (°C)	$75 \leq T_a \leq 94$	75/82/86/90/94	5
Composição de etanol (v/v)	$0,08 \leq X_a \leq 0,22$	0,08/0,10/0,15/0,20/0,22	5

Tabela 3.5. Intervalos adotados para as combinações usadas nas simulações em malha aberta.

<i>Variáveis</i>	<i>Intervalos</i>	<i>Valores usados</i>	<i>Nº variações</i>
Vazão (L.h <sup>-1</sup> )	$250 \leq Q_a \leq 350$	250/300/350	3
Temperatura (°C)	$86 \leq T_a \leq 94$	86/90/94	3
Composição de etanol (v/v)	$0,05 \leq X_a \leq 0,22$	0,05/0,10/0,15/0,20/0,22	6

### 3.3.1.6 Intervalo entre as aplicações das perturbações

O intervalo de aplicação entre uma perturbação e outra, ou o tempo de permanência da perturbação, foi determinado em função do tempo necessário até o processo atingir o novo estado estacionário (malha aberta) ou retornar ao ponto de operação (malha fechada). Este tempo é dependente da magnitude da perturbação aplicada e é variável para cada perturbação. Desta forma, o critério para a retirada da perturbação foi visual, até a composição do topo da coluna estabilizar em um dos patamares citados (atingindo novo valor ou retornando ao valor do *set point*).

Foram realizados estudos para conhecimento do comportamento dinâmico do processo tanto em malha fechada (*Capítulo IV, seção 4.2.1*), quanto em malha aberta (*Capítulo IV, seção 4.2.2*) antes de realizar as simulações definitivas. Neste estudo, observaram-se as diferentes respostas do sistema, frente às distintas perturbações. Estas análises foram corroboradas com a avaliação dos resultados das derivadas das curvas, para a certificação do alcance do estado estacionário. Nas perturbações mais severas que a unidade foi submetida em malha aberta, o processo necessitou em torno de 3,5 h para atingir estado estacionário. Em média, o tempo necessário no simulador foi de 2,5 h, para cada perturbação aplicada e mais 2,5 h para retornar ao estado estacionário anterior.

Em malha fechada, em média, o tempo necessário para estabilização do processo foi de 2 h, para cada perturbação aplicada e mais 2 h para retirada da perturbação. Este tempo foi menor quando comparado ao processo em malha aberta, o que já era esperado, em função da ação do sistema de controle.

### 3.3.1.7 Dimensão do conjunto de dados entrada/saída

Redes Neurais requerem grande quantidade de dados de processo para realizar o treinamento de forma satisfatória. Analogamente ao que ocorre com modelos lineares, é importante que os dados contenham informações suficientemente relevantes para permitir a reprodução do processo pelo modelo.

Para a confecção do conjunto de dados, é necessário que as variáveis independentes da alimentação ( $T_a$ ,  $Q_a$ ,  $X_a$ ) sejam persistentemente perturbadas. Desta forma, na obtenção dos dados de treinamento para rede em malha aberta, com emprego do simulador, foi gerado um conjunto de 78.833 pontos, referentes a aproximadamente 220 h de simulação da operação da coluna, com aquisição das variáveis

de entradas e saída da rede. Neste período, a cada 2,5h em média foi aplicada uma perturbação. Já para obtenção dos dados de validação o conjunto foi menor, 18.610 pontos, referentes a aproximadamente 51,7 h de simulação da operação da coluna.

No modelo desenvolvido representando a operação da unidade em malha fechada, foi gerado um conjunto com grande dimensão de dados, contendo 151.902 pontos, referentes a aproximadamente 425 h de simulação do funcionamento da coluna. Neste período, a cada 2 h em média foi aplicada uma perturbação. Já para obtenção dos dados de validação o conjunto obtido foi menor, 36.000 pontos, referentes a aproximadamente 100 h de simulação do processo modelado.

### 3.3.1.8 Tratamento dos dados

No tratamento dos dados utilizou-se o software *ORIGIN*<sup>®</sup> 7.5 para elaboração dos gráficos com os quais foram realizadas as comparações entre o desempenho do processo com a configuração convencional e a distribuída, além da comparação entre os resultados dos modelos frente aos dados experimentais.

Para verificar o alcance do estado estacionário em algumas situações, aplicou-se a derivada nos dados relativos às curvas de interesse, geralmente composição ou a temperatura, para verificar quando essas propriedades não variavam mais em relação ao tempo. Esta ferramenta também está presente no software *ORIGIN*<sup>®</sup>. O alcance do estado estacionário foi determinado quando a variável de interesse atinge em torno de 95% do valor do estado final.

Para verificação da precisão da estimativa (diferença entre os valores reais e estimados) das redes frente aos dados simulados, existe uma série de abordagens na literatura disponíveis. O software *STATISTICA*<sup>®</sup> 7.0 serviu como base na escolha do método de avaliação de desempenho utilizado na rede neural. Optou-se pelas medidas de erro médio quadrático (MSE) e erro percentual médio absoluto (MAPE).

## 3.3.2 Determinação da estrutura da rede neural

### 3.3.2.1 Seleção das variáveis de entrada/saída

Em geral, nem todas as variáveis adquiridas em um processo são igualmente úteis no desenvolvimento de um modelo que empregue Redes Neurais Artificiais (RNA's). De fato, a presença de variáveis irrelevantes representa uma fonte de perturbação e ruído, e pode

conduzir a um sistema com deficiência na sua representação (ZAMPROGNA *et al.*, 2001a).

Nos processos químicos e petroquímicos, especificamente no caso de colunas de destilação as quais geralmente recebem materiais provenientes de processos anteriores, existem variações tanto na carga térmica, composição e vazão, uma vez que a produção na fase anterior nunca é constante. Modificações nestas variáveis de entrada refletem em alterações na temperatura de todos os estágios, afetando diretamente a separação. Esta situação conduz a um novo estado de equilíbrio.

Logo, as temperaturas dos pratos, que são variáveis internas do processo (variáveis de estado) influenciam fortemente a composição dos produtos, e são usualmente empregadas em sensores virtuais como variáveis de entrada (VENKATESWARLU e KUMAR, 2006; CANETE *et al.*, 2008; BAHAR e OZGEN, 2008; KHAZRAEE e JAHANMIRI, 2010). Neste trabalho, têm-se disponíveis medidas de temperaturas de todos os estágios de equilíbrio tanto oriundas de simulações, quanto de aquisições na unidade piloto. Porém, neste último caso, em poucas condições de operação (insuficientes para treinamento da rede).

As medições de temperatura que são utilizados como entradas para a rede devem ser devidamente selecionadas a fim de fornecer estimativas precisas das composições, pois a localização de tais medidas têm efeitos significativos na performance da estimação da rede. Skogestad e Hori (2007) apresentam uma pesquisa abordando este tema. Não é adequado usar todas as medidas de temperatura disponíveis como entradas para o estimador por causa da redundância de medição. O uso inadequado de medidas podem levar a problemas numéricos como singularidade, sobre-parametrização e redução de precisão da estimativa (VENKATESWARLU e KUMAR, 2006).

### 3.3.2.1.1 Variáveis de entrada

Para emprego no sensor por software em *malha aberta* serão escolhidas como entradas além das variáveis manipuláveis da corrente de alimentação: composição volumétrica de etanol ( $X_a$  [v/v]), vazão ( $Q_a$  [ $L \cdot h^{-1}$ ]) e temperatura ( $T_a$  [°C]), também variáveis de estado (não diretamente manipuláveis) como temperaturas da base ( $T_b$  [°C]) e topo ( $T_t$  [°C]) e de alguns pratos intermediários da coluna, baseados na análise de sensibilidade (*Capítulo IV, seção 4.2.2.7*), como temperaturas do prato 2 ( $T_2$ ), prato 5 ( $T_5$ ) e prato 7 ( $T_7$ ) todas em °C.

Logo, temos definidas 8 variáveis de entrada para a rede em malha aberta:

- Composição volumétrica de etanol da alimentação ( $X_a$  [v/v]);
- Vazão volumétrica da alimentação ( $Q_a$  [L.h<sup>-1</sup>]);
- Temperatura da alimentação ( $T_a$  [°C]);
- Temperatura dos estágios ( $T_b, T_t, T_2, T_5, T_7$  [°C]).

As temperaturas da base e topo foram escolhidas por serem temperaturas de referências da coluna, bem como, a temperatura de alimentação do processo. Venkateswarlu e Kumar (2006) estudaram uma coluna de destilação em batelada. Observaram em sua pesquisa que o refeedor e o prato superior são os locais de medidas de temperaturas mais sensíveis. Da mesma forma, Yıldız *et al.* (2005) concluiu que, as posições de medição de temperatura devem ser distribuídos de forma homogênea em toda a coluna, desde o refeedor até o prato superior. Em ambas as pesquisas citadas, foram construídos estimadores baseados em modelo empírico para a composição.

Considerando essas discussões, três locais intermediários de medição de temperatura foram selecionados. A temperatura do prato 2 foi escolhida, representando a seção de esgotamento, devido a ser neste prato que está situada a resistência elétrica, que junto com o calor da base caracterizam o aquecimento distribuído da unidade. As temperaturas dos pratos 5 e 7, por representarem a seção de enriquecimento da coluna, e por serem as temperaturas mais sensíveis da unidade nesta seção (baseado em estudo de Marangoni, 2005). Foram escolhidas 2 temperaturas na seção de enriquecimento, de forma a conseguir uma boa representatividade de seu comportamento, em virtude desta seção conter uma quantidade maior de pratos. As temperaturas 1 e 3 foram descartadas pois seus comportamentos são muito similares a  $T_2$  (possivelmente apresentam linearidade com  $T_2$ ), devido a propagação do efeito.

Atenção especial foi dada em relação às variáveis de entrada para o monitoramento do processo com o analisador virtual de composição em *malha fechada*, pois, nesta condição de operação, a composição do topo sofre influência de outras variáveis do processo, principalmente variáveis manipuladas do sistema de controle. Desta forma, além das 8 variáveis selecionadas anteriormente, foram incluídas como entradas da rede o calor do refeedor ( $\Theta_{ref}$ ) e a vazão de refluxo (R), pois espera-se que as informações fornecidas por estas variáveis sejam relevantes para a construção dos padrões de treinamento pelo algoritmo empregado.

Logo, temos definidas 10 variáveis de entrada da rede em malha fechada:

- Composição volumétrica da alimentação ( $X_a$  [v/v]);
- Vazão volumétrica da alimentação ( $Q_a$  [L.h<sup>-1</sup>]);
- Temperatura da alimentação ( $T_a$  [°C]);
- Temperatura dos estágios ( $T_b, T_t, T_2, T_5, T_7$  [°C]);
- Calor do refeedor ( $\Phi_{ref}$  [kW]);
- Vazão de refluxo ( $R$  [L.h<sup>-1</sup>]).

A variável manipulada vazão de refluxo ( $R$ ) foi escolhida em virtude da grande influência que possui sobre a pureza do produto de topo. Vários autores propõem o uso desta variável para controlar a composição do destilado, seja diretamente ou indiretamente pelo controle da temperatura de topo (SINGH *et al.*, 2005). Da mesma forma o calor cedido pelo refeedor ( $\Phi_{ref}$ ), sendo esta variável uma das principais fontes de calor do processo e podendo afetar diretamente as vazões de vapor e indiretamente as composições dos pratos.

Desta forma, como entrada da rede neural foi selecionado um número de variáveis (8 – malha aberta e 10 – malha fechada) do qual se entende serem capazes de fornecer informações suficientes em relação ao processo, para geração de padrões de treinamento da rede. O emprego das temperaturas de todos os pratos não traria informações novas à rede, pelo contrário, geraria um conjunto de dados demasiadamente grande, com informações repetitivas, com provável linearidade entre elas, o que traria prejuízos e lentidão à rede.

### 3.3.2.1.2 Variáveis de saída

São os valores fornecidos como resposta do simulador, frente às perturbações aplicadas nas variáveis de entrada, que representa o modelo dinâmico da coluna de destilação da unidade piloto. As saídas mais relevantes do processo de destilação são as composições dos produtos. No caso específico da unidade piloto, é a composição de topo de etanol. Logo, a composição volumétrica de etanol no topo ( $X_t$  [v/v]) será a variável escolhida de saída (tanto em malha aberta, quanto em malha fechada), a qual será obtida com o uso do simulador e fornecida à rede neural durante a etapa do treinamento. Posteriormente, será a variável estimada pelo sensor e usada para monitoramento e controle do processo.

### 3.3.2.2 Ambiente de desenvolvimento da rede neural

Visando avaliar os diversos métodos e algoritmos de treinamento e aprendizado, foi feito uso do ambiente do pacote MATLAB 7.1, versão R2010a<sup>®</sup>, que é um ambiente computacional para visualização e computação numérica de alto desempenho. Uma tela da interface do programa usado é mostrada na Figura 3.9. Segundo Plumb *et al.* (2005) outros pacotes comerciais, além do MATLAB<sup>®</sup>, também são empregados com sucesso para desenvolvimento de redes neurais, entre eles pode-se citar INForm e CAD/Chem.

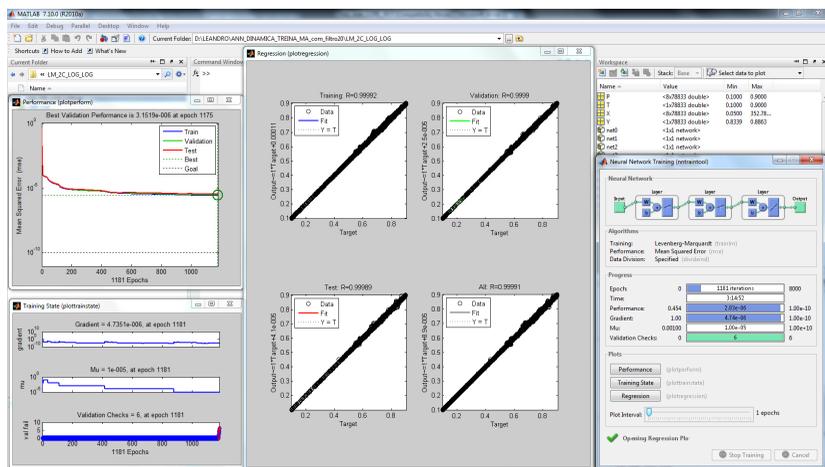


Figura 3.9. Tela do programa MATLAB 7.6 R2010a<sup>®</sup> utilizado para realizar o treinamento da rede.

Foram utilizadas as bibliotecas (*Neural Network Toolbox*) de redes neurais e de interface gráfica. Essa biblioteca utiliza programação orientada a objeto e interage diretamente com o ambiente MS-Windows. O pacote de redes neurais abrange diversos modelos de redes (*Backpropagation, Elman, Hopfield, entre outros*) e funções correlatas. O software MATLAB<sup>®</sup> integra análise numérica, processamento de sinais e gráficos, em que o problema e as soluções são expressos tal como são formulados matematicamente, sem a programação tradicional.

### 3.3.2.3 Configuração da rede neural

Esta etapa pode ser dividida em três passos: 1) seleção do paradigma neural apropriado à aplicação – aprendizado supervisionado

ou não e tipo de algoritmo de otimização; 2) determinação da topologia da rede a ser utilizada – o número de camadas, o número de neurônios em cada camada; 3) determinação de parâmetros do algoritmo de treinamento, como a escolha das funções de ativação, etc. Estes passos possuem um grande impacto nas capacidades e limitações do desempenho do sistema resultante.

Existe grande dificuldade de definir a arquitetura ideal da rede. Não existem regras claras para a definição de quantos neurônios devem existir nas camadas intermediárias, quantas camadas, ou como devem ser as conexões entre essas unidades. Normalmente, estas escolhas são feitas de forma empírica por tentativa e erro, testando algumas variações de arquiteturas (topologia e parâmetros). A definição da configuração de redes neurais é considerada uma arte que requer grande experiência dos projetistas. Para tentar contornar essa dificuldade, algumas regras foram propostas por alguns pesquisadores. Porém, estas regras são muitas vezes contraditórias e pouco práticas. Outra forma possível é o emprego de técnica de inteligência artificial, denominada de Algoritmo Genético, para auxiliar na busca de arquiteturas de redes neurais otimizadas. No entanto, nenhuma delas garante que a melhor opção seja encontrada. Por esses motivos, estas regras e técnicas (MENDES *et. al*, 1997; MENDES e CARVALHO, 1998; SON *et al.*, 2004) não foram empregadas neste trabalho. Optou-se por testar várias arquiteturas, até encontrar redes que obtivessem bom desempenho frente aos dados utilizados e com boa capacidade de generalização.

### 3.3.2.3.1 Tipos de rede neural avaliadas

Redes neurais podem ser classificadas de acordo com as suas estruturas e algoritmos de aprendizagem como redes: *feedforward* e recorrentes. Neste trabalho serão avaliadas estes dois tipos de redes neurais, *feedforward* do tipo *Multi-Layer Perceptron* (MPL) – *Backpropagation* e redes recorrentes do tipo Elman. Optou-se por selecionar dois tipos diferentes de arquiteturas e avaliá-las segundo o seu desempenho com os conjuntos de dados obtidos pelo simulador, referente ao processo que se deseja modelar. Esta seleção teve como intuito permitir a comparação de uma rede considerada padrão para problemas de identificação de sistemas, a *Multi-Layer Perceptron* (MPL), com uma rede muito utilizada em processos dinâmicos, a rede recorrente de Elman.

A arquitetura das redes *Multi-Layer Perceptron* (MPL) se caracteriza pelo seu alto grau de conectividade, com neurônios dispostos

em uma série de camadas. Pelo menos uma camada oculta e uma camada de saída e são conhecidas também como redes alimentadas adiante (redes *feedforward*), pois a informação é propagada em apenas um sentido, da primeira para a última camada. Em uma rede de Elman, além de camadas de entrada, intermediárias e saída, há também unidades de contexto. As unidades de contexto são usadas para memorizar as ativações anteriores das unidades intermediárias. Em um momento específico,  $k$ , as ativações anteriores das camadas intermediárias (em tempo de  $k-1$ ) e as correntes de entradas propriamente ditas (em  $k$  tempo) são usadas como entradas para a rede.

#### 4.3.2.3.2 Métodos e algoritmos de treinamento avaliados

Para redes *backpropagation* do tipo MPL, vários tipos de algoritmos de treinamentos são disponibilizados, e diferem entre si pela estratégia de otimização empregada ao realizar ajustes dos pesos e bias. Cada um possui requisitos diferenciados em termos de memória computacional exigida para processamento dos cálculos e tempo de convergência. Geralmente são divididos em grupos que utilizam treinamentos baseados em: *Métodos de gradiente descendente com e sem momento*, *Gradiente conjugado* e *Método de quasi-Newton*.

Segue abaixo descrito, primeiramente, o método que o algoritmo emprega e, em seguida o algoritmo de treinamento propriamente dito e que serão avaliados neste trabalho:

*Métodos de gradiente descendente sem momento:*

- *Taxa de Aprendizado Adaptativa (GDx)*;
- *Resilient Backpropagation (RP)*;

*Método de gradiente conjugado:*

- *Gradiente Conjugado em Escala (SCG)*;

*Métodos quasi-Newton:*

- *Levenberg-Marquardt (LM)*
- *Regularização de Bayesian (BR)*

*Método intermediário entre gradiente conjugado e quasi-Newton:*

- *One Step Secant Method (OSS)*.

O estudo com as avaliações do comportamento de distintos algoritmos de treinamento foi realizado visando, além de encontrar o algoritmo que mais se adapte e forneça as melhores respostas ao conjunto de dados, contribuir cientificamente para as pesquisas na área. Grande parte dos autores empregam o *Método quasi-Newton* com o algoritmo *Levenberg-Marquardt* (HAGAN *et al.*, 1994; ZHAO e XU,

2004; KERMANI *et al.*, 2005; SINGH *et al.*, 2007; MUKHERJEE e SRIKANTA, 2012) em seus trabalhos. Porém, são poucas as pesquisas que relatam em quais situações cada algoritmo é recomendado, e que apresentam um comparativo do desempenho entre os algoritmos de treinamento empregando diferentes métodos. Logo, este estudo visa dar sua contribuição neste aspecto.

### 3.3.2.3.3 Funções de transferência avaliadas

Para as redes *backpropagation* as funções de transferência aplicadas nas camadas intermediárias e na camada de saída serão uma combinação entre funções *tangente hiperbólica*, *logarítmica* e *linear*. Estas foram escolhidas pois apresentam o melhor resultado na transferência dos pesos ponderados. Sendo que a função linear foi aplicada apenas na camada de saída das redes, usando-se variantes com padronizações distintas: *purelin* (-1 a +1) e *poslin* (0 a +1). O software MATLAB<sup>®</sup> é usado para realizar esse pré-processamento e padronização dos dados.

Em relação às redes de Elman, serão testadas distintas arquiteturas e configurações de redes, para comparar seu desempenho frente às redes *feedforward* - MPL (sem realimentação). Nestas redes recorrentes será aplicado apenas o algoritmo de treinamento que obteve o melhor desempenho nas redes MPL que foi a *Regularização de Bayesian* (BR), para o conjunto de dados específico. Foi empregada a função de transferência padrão para este tipo de rede: *tangente hiperbólica* nas camadas intermediárias e *linear* na camada de saída. Porém, também foram testadas redes usando função *logarítmica* tanto nas camadas intermediárias quanto na saída. Estas escolhas foram feitas para as redes de Elman visando diminuir o universo dos testes.

Pelo fato das redes recorrentes reconhecerem padrões variantes no tempo e possuírem capacidade de armazenar informações passadas, espera-se obter redes recorrentes com bom desempenho, para o processo que se deseja modelar e aplicar o sensor virtual.

### 3.3.2.3.4 Número de camadas e neurônios

No presente trabalho serão empregadas redes contendo 1, 2 e 3 camadas intermediárias, com número de neurônios variando de 3 a 30 em cada camada. A Tabela 3.6 ilustra as situações avaliadas. Foram testadas 27 combinações. Nas redes em malha aberta foram empregadas arquiteturas com 8 entradas e uma saída. Para redes em malha fechada foram testadas 8 e 10 entradas, ambas com 1 saída. Para todas as redes,

o número de camadas intermediárias e de neurônios variou de acordo com a Tabela 3.6.

Tabela 3.6. Arquiteturas das redes neurais testadas em relação ao número de camadas e neurônios.

<i>Nº de camadas Intermediárias</i>	<i>Nº de neurônios</i>
1	[5]
1	[10]
1	[15]
1	[20]
1	[25]
2	[3 3]
2	[5 5]
2	[10 5]
2	[5 10]
2	[10 10]
2	[15 10]
2	[15 15]
2	[20 15]
2	[20 20]
2	[30 30]
3	[3 3 3]
3	[5 5 5]
3	[8 8 8]
3	[5 10 5]
3	[10 5 5]
3	[10 10 10]
3	[15 10 5]
3	[10 15 10]
3	[15 15 15]
3	[20 15 10]
3	[20 20 20]
3	[30 30 30]

A determinação do número de neurônios ideal na camada intermediária foi feita através de tentativas. Não existe um critério geral que permita definir o número de neurônios na camada intermediária. Em geral, redes neurais artificiais com poucos neurônios nas camadas intermediárias são preferidos, visto que elas tendem a possuir melhor poder de generalização, reduzindo o problema de sobre-ajuste ou excesso de treinamento (*overfitting*). Entretanto, redes com poucos neurônios intermediários podem não possuir a habilidade suficiente para modelar e aprender os dados em problemas complexos, podendo ocorrer *underfitting* (PIAIA *et al.*, 2008). Desta forma, é necessário encontrar um ponto ótimo, com redes de tamanho moderado.

Com o objetivo de que as redes obtivessem boa capacidade de generalização, a estratégia de escolha da quantidade de neurônios na camada intermediária foi encontrar a quantidade mínima suficiente para um desempenho adequado. Avaliando os procedimentos empregados com as variações entre algoritmos de treinamento, número de camadas, número de neurônios e combinações das funções de transferências, acredita-se que com este nível de complexidade serão encontradas redes que consigam modelar o processo de forma coerente e com isso obtenham excelente capacidade de predição da propriedade desejada.

### 3.3.2.3.5 Procedimento de treinamento, inicialização e critério de parada

As redes neurais necessitam de um grande número de dados do processo para seu treinamento. É importante, entretanto, que esses dados contenham informações suficientemente relevantes de forma a permitir a correta modelagem do processo.

Na Tabela 3.7 estão contidas as matrizes do banco de dados formado com o auxílio do simulador de processos, empregado para geração dos conjuntos de dados entrada/saída. Estes dados são usados tanto na etapa de treinamento quanto na validação, e estão separados de acordo com o modo de operação do simulador: malha aberta e malha fechada. Com objetivo de formar redes com diferentes finalidades em cada caso, porém com a mesma propriedade estimada. Nas matrizes da Tabela 3.7 o primeiro número corresponde ao número de linhas (pontos/dados coletados) e o segundo ao número de colunas (variáveis).

Tabela 3.7. Banco de dados e dimensões das matrizes de entrada/saída nas etapas de treinamento e validação, com simulador operando em malha aberta e fechada.

<i>Arquitetura</i>	<i>Malha Aberta</i> <i>8 entradas</i>		<i>Malha Fechada</i> <i>8 entradas</i>		<i>Malha Fechada</i> <i>10 entradas</i>	
	<i>Conjunto de</i> <i>Entrada (X)</i>	<i>Conjunto de</i> <i>Saída(Y)</i>	<i>Conjunto de</i> <i>Entrada(X)</i>	<i>Conjunto de</i> <i>Saída(Y)</i>	<i>Conjunto de</i> <i>Entrada(X)</i>	<i>Conjunto de</i> <i>Saída(Y)</i>
ETAPA						
TREINAMENTO	$\langle 78.833 \times 8 \rangle$	$\langle 78.833 \times 1 \rangle$	$\langle 151.902 \times 8 \rangle$	$\langle 151.902 \times 1 \rangle$	$\langle 151.902 \times 10 \rangle$	$\langle 151.902 \times 1 \rangle$
VALIDAÇÃO	$\langle 18.610 \times 8 \rangle$	$\langle 18.610 \times 1 \rangle$	$\langle 36.000 \times 8 \rangle$	$\langle 36.000 \times 1 \rangle$	$\langle 36.000 \times 10 \rangle$	$\langle 36.000 \times 1 \rangle$

Nesta etapa, seguindo o algoritmo de treinamento escolhido, são ajustados os pesos das conexões. É importante considerar, nesta fase, alguns aspectos, tais como a inicialização da rede, o modo de treinamento e o tempo de treinamento.

Normalmente, os valores iniciais dos pesos da rede são números aleatórios uniformemente distribuídos, em um intervalo definido. A escolha errada destes pesos pode levar a uma saturação prematura. Nguyen e Widrow (1990) encontraram uma função que pode ser utilizada para determinar valores iniciais melhores que valores puramente aleatórios. Esta função denominada “*initnw*” está disponibilizada pelo MATLAB<sup>®</sup> e foi empregada nas redes desenvolvidas neste trabalho. A principal vantagem de uma escolha mais apropriada para a inicialização dos pesos da rede refere-se, principalmente, ao tempo de treinamento necessário e à melhor acuidade dos resultados finais apresentados.

Para obtenção de uma rede com bom desempenho, tanto em termos de treinamento, quanto posteriormente, na etapa da validação, é necessário estabelecer um critério de parada do treinamento. Existem vários fatores que podem influenciar o tempo de treinamento da rede, porém é fundamental determinar um critério de parada. O mais comum e utilizado por diversos autores (PATIL e NIGAM, 2009), é o erro quadrático mínimo apresentado pela rede na sua adequação ao conjunto de dados do treinamento, o qual afeta indiretamente a capacidade de generalização da rede. Este foi um dos critérios de parada adotado.

No treinamento realizado usando-se o software MATLAB<sup>®</sup>, o próprio conjunto de dados de treinamento é dividido em 3 subconjuntos (70%, 15%, 15%). O primeiro realiza o treinamento propriamente dito. É usado para o cálculo do gradiente e a atualização dos ganhos e bias da rede e faz a verificação do ponto de melhor desempenho da rede, através do valor do erro médio quadrático (MSE). O segundo sub-conjunto é usado para validação do modelo paralelamente ao treinamento. O erro do modelo com os últimos ganhos e bias calculados, bem como os dados da validação, são monitorados durante o processo de treinamento da rede. Normalmente este erro decresce na medida em que a rede gera os padrões de aprendizado. Entretanto, quando a rede começa a se tornar sobre treinada, o erro do conjunto de validação começa a crescer. Quando este fato ocorre, é definido um número máximo de iterações com erro crescente (6 épocas), no qual o treinamento é parado, quando este número for atingido, e os ganhos e bias no erro mínimo de validação são os que melhor representam o processo em questão. O terceiro sub-conjunto de dados é usado no teste do modelo. Assim, deve-

se encontrar redes com uma boa capacidade de generalização, um ponto ótimo de parada com erro mínimo. Quando atingido o número máximo de épocas (10.000) o treinamento também é interrompido. Logo, estes foram os critérios de parada adotados.

No código do algoritmo da rede neural vários parâmetros foram definidos, tais como: tempos de visualização, erro mínimo, limite de iterações/épocas, etc. Esses valores, juntamente com os pesos da rede neural e a sua configuração, são salvos para futura utilização.

Para realização dos treinamentos das redes neurais foram usados 5 computadores. Sendo 4 com processador Intel® Core i7 950 3GHz 64bit, cada um com 4GB de RAM. Adicionalmente, para redes com arquiteturas mais robustas, foi usado um computador com processador Intel® Core 2Quad 2,5GHz Q9300 64bit com 8GB de RAM. O tempo de treinamento variou consideravelmente dependendo da arquitetura da rede e da dimensão do conjunto de dados. Foram realizados treinamentos com duração de 30 minutos, para redes com poucas camadas e neurônios, até 70 horas para redes muito robustas.

### 3.3.3 Validação da rede – Medidas de desempenho

Existem várias maneiras de avaliar o desempenho do estimador desenvolvido. A forma mais importante, e talvez mais fácil, seja através da medição da precisão da estimativa. A precisão da estimativa pode ser definida como a diferença entre os valores do simulador e estimados. Há uma série de abordagens que apresentam as medidas de precisão na literatura (KUMAR e PALIWAL, 2009), como SSE (soma de erros quadrados), RMSE (raiz quadrada média do erro), MAPE (erro médio percentual absoluto) e outros. Mais frequentemente utilizado é o MSE (erro médio quadrático), definido pela Equação 3.1.

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2 \quad (\text{Eq. 3.1})$$

Esta última abordagem foi escolhida, sendo então o erro médio quadrático utilizado para determinar o desempenho das redes e auxiliar na busca da topologia ótima das redes avaliadas. A decisão sobre a topologia ideal foi baseada no erro médio quadrático mínimo, observado na etapa da validação. Para auxiliar e corroborar na escolha foi usado outro parâmetro de avaliação, o erro médio percentual absoluto (MAPE), justamente por ser um erro percentual e não levar em

consideração valores positivos e negativos, mas sim seu módulo. Calculou-se o erro percentual (PE) e, através dele, o erro percentual médio absoluto (MAPE). As respectivas expressões (Equação 3.2 e Equação 3.3) usadas para o seu cálculo são mostradas a seguir:

$$PE_t = \left( \frac{F_t - Y_t}{Y_t} \right) * 100 \quad (\text{Eq. 3.2})$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n |PE_t| \quad (\text{Eq. 3.3})$$

sendo:

n: número de pontos do conjunto de dados;

$Y_t$ : valores de saída obtidos pelo HYSYS<sup>®</sup> para os pontos do grupo de dados;

$F_t$ : valores de saída inferidos pela rede neural para os pontos do grupo de dados.

Desta forma, a avaliação da rede com os dados utilizados no treinamento e na validação consiste na verificação do desempenho na previsão de valores da composição do produto de topo, alvo de interesse. Os valores de saída das composições de etanol preditos pela rede são comparados com os valores obtidos nas simulações com o HYSYS<sup>®</sup>.

### 3.4 CONCLUSÕES

O capítulo apresentou desde a caracterização da unidade piloto de onde foram obtidos os dados experimentais que permitiram validar o modelo desenvolvido, bem como todo processo de identificação do sistema que compreende a obtenção do grupo de dados, determinação da estrutura neural e etapa de validação.

Neste trabalho, o simulador HYSYS<sup>®</sup> foi utilizado como ferramenta importante pela confiabilidade que o software possui, uma vez que é largamente empregado na indústria e foi responsável pela geração de dados para o treinamento e validação da rede.

A utilização de softwares comerciais para simulação de processos representa um grande avanço para a pesquisa e ampara o setor industrial, tanto devido à rapidez, como pela facilidade de manipulação, permitindo assim, a execução de testes sem a utilização de plantas reais.

Pode-se ainda concluir que o HYSYS<sup>®</sup> foi adequado e eficiente gerador de dados operacionais, fazendo com que a rede seja treinada contemplando uma faixa operacional compatível com os pontos de operação que se deseja operar a planta.

Com relação ao procedimento apresentado para construção e desenvolvimento da rede neural, buscou-se esclarecer ao leitor a metodologia empregada, detalhando todos os critérios adotados e a fundamentação teórica nas quais certas decisões foram baseadas, buscando oferecer ferramentas suficientes para o entendimento e avaliação do referido trabalho.

No próximo capítulo serão abordados os resultados obtidos.

# Capítulo IV

## RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentados os resultados do modelo da unidade de destilação desenvolvido com o emprego do software HYSYS® 7.0 inicialmente em modo estático, abordando apenas o controle convencional. Posteriormente, será apresentado o modelo dinâmico completo, com a comparação dos resultados simulados com os dados experimentais de perfis de temperatura, com o intuito de validar as simulações, para duas perturbações realizadas experimentalmente.

Para o modelo dinâmico, foi implementado tanto o sistema convencional de controle, quanto a configuração com ação distribuída, com controle aplicado no prato 2 que será o foco principal.

Com o simulador dinâmico em malha aberta, será realizada a caracterização da dinâmica da coluna em regime e em torno do ponto de operação nominal, através de aplicações de degraus na carga processada. Esta etapa fornecerá dados para seleção das variáveis de entrada/saída do sensor e para determinação do período de amostragem e tempo entre as aplicações das perturbações até o sistema entrar em regime.

Na sequência o foco se voltará para a concepção de redes neurais capazes de realizar a inferência da composição de topo da coluna com maior precisão. Com os dados para o treinamento, obtidos no modelo da coluna desenvolvido no simulador dinâmico, foram avaliados dois tipos de redes neurais dinâmicas: redes *Multi-Layer Perceptron* (MLP) e redes recorrentes tipo Elman. Em ambas vários algoritmos de treinamento, com diversas funções de transferência e arquiteturas de redes foram avaliados para verificar qual obtém melhor performance na inferência da composição de produto de topo.

Segue descrito, no diagrama esquemático ilustrado na Figura 4.1, a sequência da apresentação dos resultados deste projeto, ilustrando suas principais etapas para o desenvolvimento dos analisadores virtuais, sendo, o primeiro em malha aberta, no qual sua predição pode ser empregada no controle e otimização futura da unidade piloto e, o segundo obtido com a planta operando em malha fechada, que poderá ser usado para monitoramento do processo em tempo real.

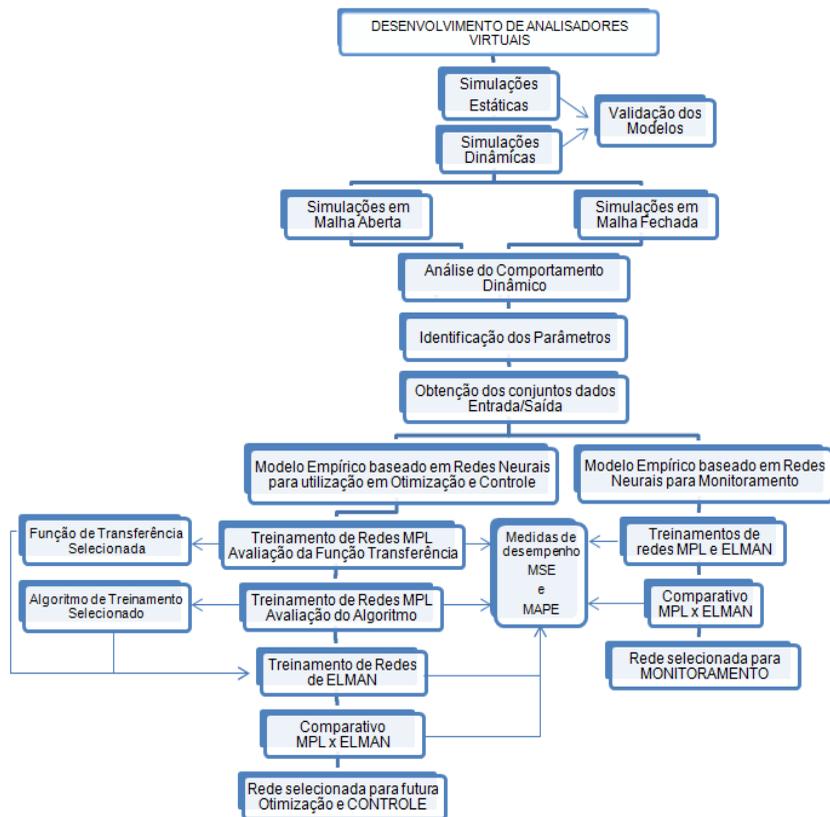


Figura 4.1. Diagrama esquemático para ilustração da sequência da apresentação dos resultados com as etapas do desenvolvimento dos analisadores virtuais para controle e monitoramento.

#### 4.1 SIMULADOR EM MODO ESTÁTICO

Desenvolver uma simulação estática que reproduza com fidelidade o comportamento da unidade alvo do estudo é fundamental para realização de estudos mais aprofundados sobre o comportamento do processo, uma vez que nas simulações podem ser realizados testes e se obter alguns dados que seriam inviáveis experimentalmente. Além disso, a simulação estacionária serve como base para construção do modelo dinâmico, onde são incorporadas as constantes de tempo e o processo passa a ser representado com maior fidelidade.

O simulador estático foi empregado no início deste trabalho. O software apresenta a possibilidade de migrar de uma simulação estática para o estado dinâmico, o que lhe confere um atrativo especial, sendo este um dos motivos pela sua escolha.

O estudo realizado nesta seção teve o intuito de verificar a confiabilidade das respostas obtidas nas simulações com o software HYSYS<sup>®</sup>. Nesta etapa foram avaliados conceitualmente alguns fenômenos do processo de destilação, bem como comportamentos conhecidos da mistura binária etanol/água. Os resultados obtidos através da análise das saídas das simulações também serviram para validação do simulador, e podem ser usados como banco de dados para futuros trabalhos.

Todas as simulações foram realizadas partindo de uma simulação padrão, onde foram usados parâmetros e condições idênticas as obtidas experimentalmente, detalhadas no *Capítulo III, na seção 3.3.1.2*, validando os perfis de temperatura por meio de comparação com os perfis experimentais.

Outras condições da coluna que foram mantidas constantes nesta etapa são as condições do topo, tais como vazão de destilado ( $10 \text{ L.h}^{-1}$ ) e razão de refluxo ( $R/D=6$ ), de acordo com dados experimentais obtidos. Já a carga térmica do refeedor foi calculada pelo software nas simulações, de acordo com alterações na corrente de alimentação e as condições fixadas no topo. Os resultados obtidos nos perfis de composição das figuras desta seção estão representados em base molar.

#### 4.1.1 Análise de uma simulação padrão

Os perfis de separação obtidos para condição de alimentação de  $80^\circ\text{C}$  e  $300 \text{ L.h}^{-1}$  são representados na Figura 4.2 pela fração molar dos componentes água e etanol para a fase líquida. Na legenda da figura estão representadas as frações molares dos respectivos componentes. Vários efeitos podem ser observados, dentre eles a grande quantidade de água e baixa quantidade de etanol na seção de esgotamento. Nesta seção as frações de etanol são pouco incrementadas. Por ser uma seção mais rica em líquido, existe o “enriquecimento” dos componentes menos voláteis, neste caso a água. Este também é o resultado da introdução da corrente de alimentação com temperatura abaixo da temperatura do prato em que é introduzida. Com isso, existe a condensação da fase vapor que ascende a coluna até o prato de alimentação. O incremento de etanol ocorre logo acima do prato da alimentação, como era esperado, onde ocorre a mudança de seção da coluna.

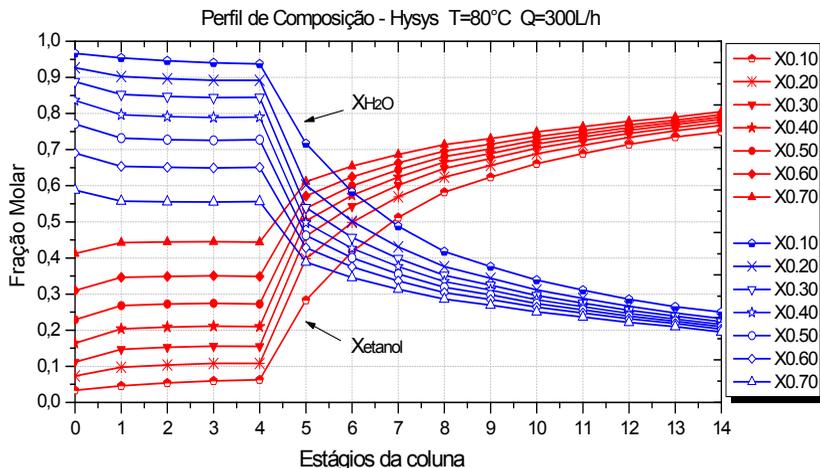


Figura 4.2. Perfis de fração molar de etanol (vermelho) e água (azul) para a fase líquida.

A partir dos dados das simulações foi possível avaliar se os experimentos estavam sendo conduzidos na faixa do azeótropo formado pela mistura etanol-água. Foi construído um gráfico representativo do equilíbrio líquido vapor (ELV) que é apresentado na Figura 4.3, onde as condições da alimentação da coluna são 0,1 de fração molar de etanol, 80°C e 300 L.h<sup>-1</sup>, conforme ilustrado no título do gráfico.

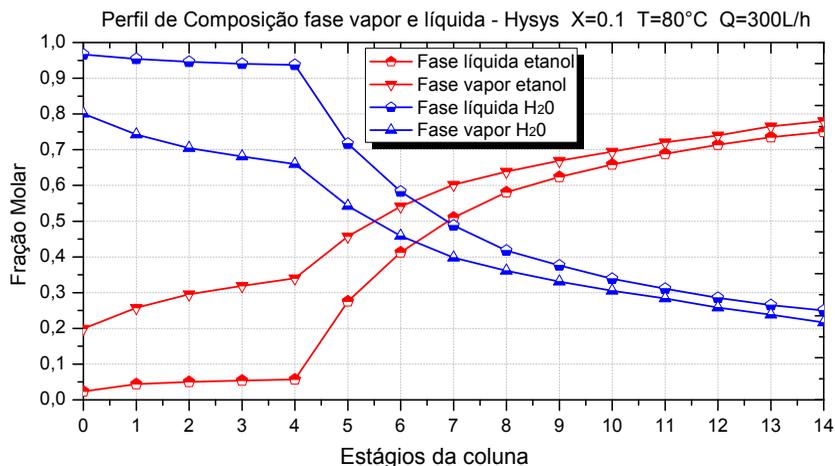


Figura 4.3. Equilíbrio líquido-vapor para o etanol e para a água, na fase vapor e na fase líquida.

Observa-se que para a condição avaliada, as composições da fase líquida e vapor, não se igualam, indicando que a separação está ocorrendo fora da faixa do azeótropo.

#### 4.1.2 Avaliação do efeito da alteração nas condições de alimentação da coluna

A seguir, serão apresentadas as avaliações do efeito da alteração de uma das variáveis da corrente de alimentação sobre o processo de destilação, mantendo-se a temperatura ( $75^{\circ}\text{C}$ ) e a vazão ( $300 \text{ L}\cdot\text{h}^{-1}$ ) desta corrente, constantes. A variável escolhida para este estudo foi a *fração molar de etanol*, em virtude do conhecimento de seu impacto sobre o sistema. O perfil de composição ao longo dos pratos é ilustrado na Figura 4.4 para uma determinada condição de operação da coluna. Vale destacar que o objetivo nesta etapa não é avaliar o efeito em si causado pelas perturbações, pois eles já são conhecidos e disponíveis em vários trabalhos na literatura. O objetivo principal é verificar se a resposta do software é coerente e reproduz tais efeitos.

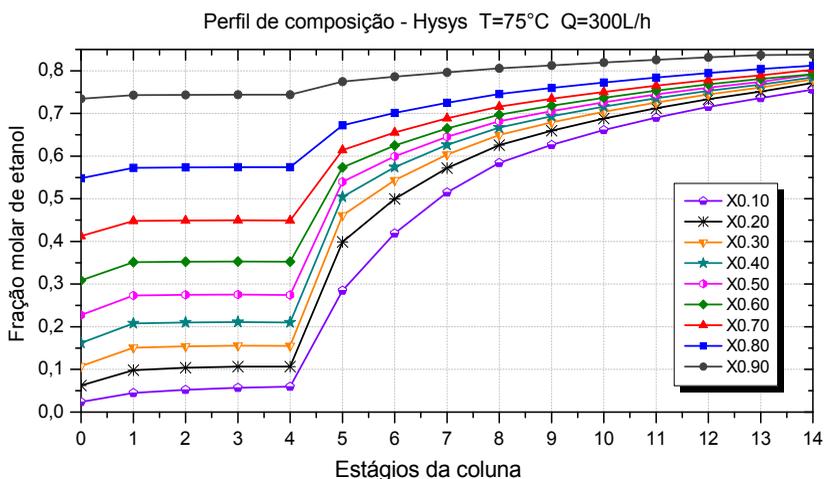


Figura 4.4. Efeitos da alteração da composição da alimentação no perfil de composição de etanol nos pratos da coluna, em estado estacionário.

Percebe-se claramente que quando se trabalha com frações baixas de etanol na alimentação, como por exemplo, na faixa de 0,1 a 0,5 (Figura 4.4), o perfil de composição ao longo dos pratos apresenta maior inclinação a partir do prato da alimentação (4), decorrente da

maior necessidade de energia a ser suprida ao sistema para evaporação da mistura, que contem quantidades elevadas do componente menos volátil. À medida que a fração de etanol na alimentação é incrementada esta inclinação vai diminuindo, até o ponto no qual a fração de etanol é tão alta que praticamente não ocorre mais o processo de separação. Este fato é atribuído a proximidade do ponto de azeótropo e, portanto, o uso da destilação convencional para tal finalidade já não é mais economicamente viável. Estes efeitos são os esperados nestas situações. Portanto, o simulador estático obteve resposta representativa.

Para manter o perfil de temperatura e de composição desejado na coluna, principalmente na seção de retificação, quando se trabalha com temperaturas de alimentação abaixo do ponto nominal de operação, por exemplo, existe a necessidade de grandes aportes de energia na carga térmica do refeedor, como pode ser observado na Figura 4.5 (a,b,c).

Esta figura ilustra o efeito da variação da composição de etanol da alimentação sobre a carga térmica do refeedor, em distintas temperaturas e vazão fixa (Figura 4.5a), em distintas vazões a temperatura da alimentação fixa (Figura 4.5b), e o efeito da variação da temperatura de alimentação em distintas vazões a composição da alimentação fixa (Figura 4c). Logo, para manter a composição desejada no topo constante, quando se trabalha com temperaturas de alimentação baixas, deve-se avaliar o custo energético, pois este pode inviabilizar o processo.

Através da análise das curvas da Figura 4.5a observa-se que, com o aumento da temperatura da alimentação, tem-se menor necessidade de carga térmica do refeedor, ou seja, a condição entálpica da alimentação determina a quantidade de calor que deve ser suprida pelo refeedor. Na Figura 4.5b percebe-se que quando se trabalha com vazão mais elevada e menor fração de etanol desta corrente, tem-se maior necessidade de carga térmica no refeedor, devido a maior quantidade do componente menos volátil e ao grande volume a ser processado. Efeito similar foi observado por Werle (2007) ao realizar uma perturbação positiva na vazão da alimentação e estudar o efeito causado na carga térmica do refeedor.

Verifica-se pela análise da Figura 4.5c, o efeito combinado do aumento da temperatura e vazões baixas na alimentação. Neste caso, tem-se uma baixa necessidade de carga térmica na base da coluna, pois com a alimentação já pré-aquecida, obtêm-se uma situação mais otimizada para operação da unidade, corroborando as análises anteriores.

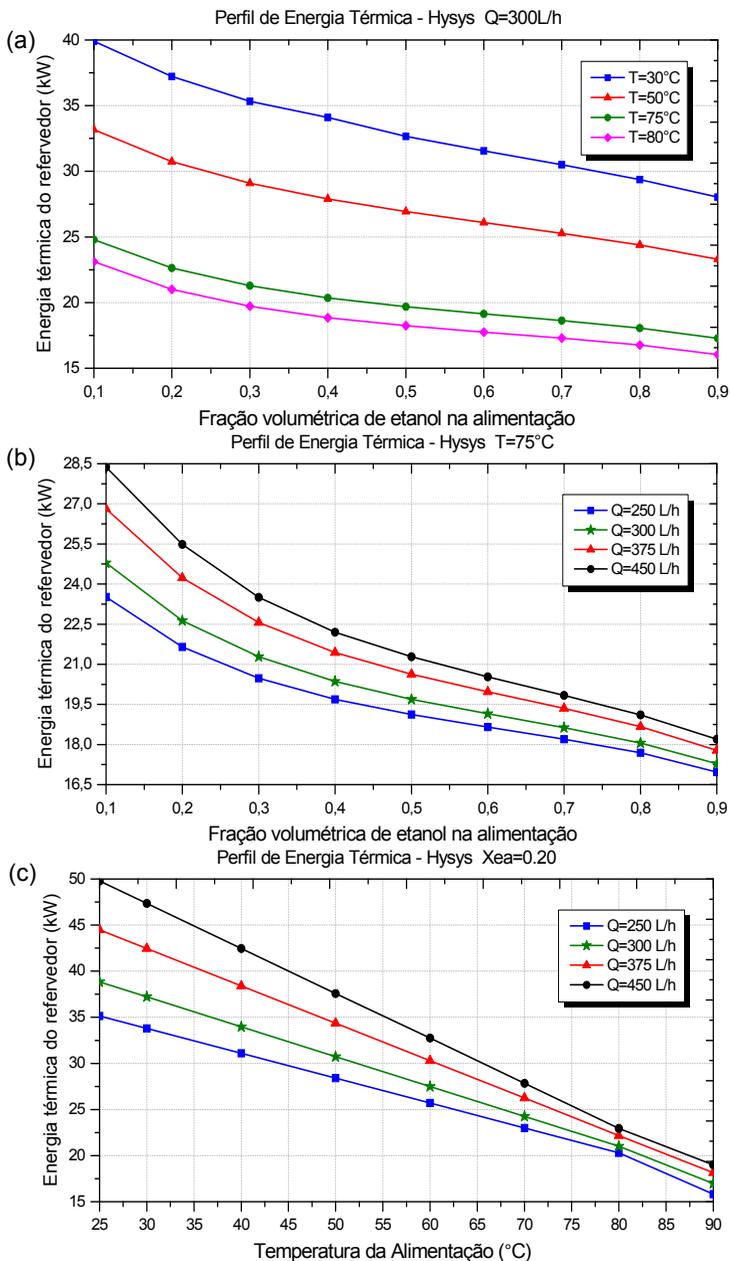


Figura 4.5. Necessidade energética decorrente de alterações da condição de alimentação da coluna.

Os resultados encontrados através destas análises já eram esperados e seguem o comportamento previsto do sistema etanol-água nas condições avaliadas. Mostram a coerência dos dados obtidos através das simulações e servem para comprovar a eficácia do simulador em reproduzir o processo e seus efeitos quando sujeito a perturbações. Com isso, segue-se para a próxima etapa que consiste na migração do software estático para o modo dinâmico de operação, no qual serão gerados os dados para treinamento da rede neural.

## 4.2 SIMULADOR EM MODO DINÂMICO

A simulação dinâmica de um processo possibilita avaliar seu comportamento em regime transitório entre dois estados. O objetivo principal do emprego do software HYSYS<sup>®</sup> em seu modo dinâmico é na geração de dados de composição de topo da coluna de destilação em regime transiente, os quais não se têm em grande número na unidade de destilação piloto do Laboratório de Controle de Processos – LCP/UFSC, pois a mesma não possui analisadores de composição em linha. Os poucos dados de composição que se tem foram obtidos através da coleta de amostras e análise em laboratório e/ou densímetro, porém com a coluna operando em estado estacionário.

Na Figura 4.6a é ilustrada uma das interfaces desenvolvidas no simulador em modo dinâmico para uma coluna de destilação operando nas condições similares as adotadas experimentalmente. Na mesma Figura 4.6b é ilustrada outra interface, detalhando a malha interna de controle da coluna de destilação. Observa-se o controle da fração volumétrica de etanol da corrente de destilado em cascata com o controle de vazão de refluxo.

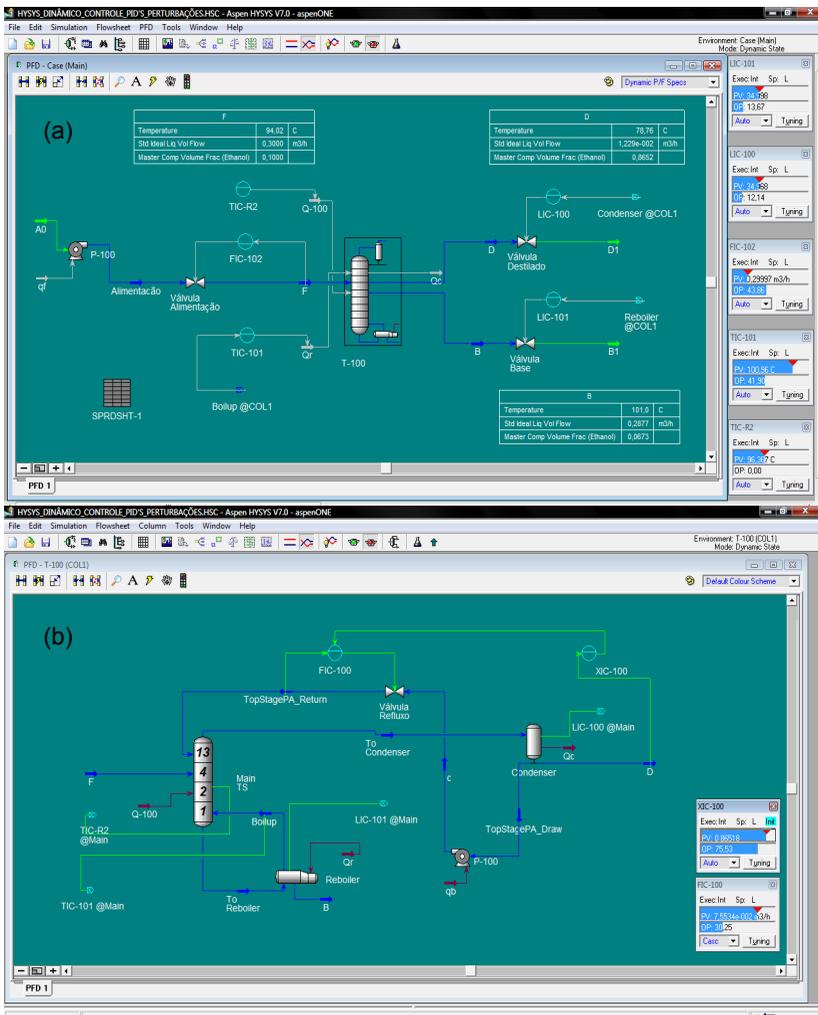


Figura 4.6. Tela do programa HYSYS® em modo dinâmico com (a) esquema da unidade de destilação e seus controladores (b) malha interna da coluna de destilação, com o controle de composição do topo.

Neste estudo empregou-se o simulador dinâmico como uma planta de processo, a partir do qual obteve-se um conjunto de dados de entrada/saída, com os quais pretende-se treinar uma rede neural e desenvolver um sensor por software. A rede neural artificial foi previamente treinada para inferir valores de composição do produto de

topo de uma coluna de destilação piloto. A proposta é utilizar a composição estimada como variável controlada *on-line*.

#### 4.2.1 Validação das simulações dinâmicas e análise do comportamento do processo em malha fechada

Nesta etapa foram desenvolvidas simulações computacionais buscando reproduzir a operação da planta piloto de destilação. Partindo-se da simulação em estado estacionário, a qual serviu de base para desenvolvimento do modelo dinâmico, foram acrescentadas as malhas de controle das variáveis que se deseja controlar e realizada a sintonia destes controladores com ferramentas do próprio software. Realizaram-se várias simulações, ajustando parâmetros de processo, dimensionamento de equipamentos e válvulas e, principalmente, o sistema de controle.

Foi aprimorada também a parte gráfica do software, desenvolvendo telas que permitem o monitoramento *on-line* dos pares de variáveis manipuladas e controladas, bem como visualização de perfis de composição e temperaturas em todos os estágios da coluna e outras variáveis de interesse, conforme ilustrado na Figura 4.7. Desta forma, o modelo além de gerar os dados dinâmicos para treinamento da rede, também pode ser utilizado similarmente como um sistema supervisor do processo simulado, para melhor visualização e avaliação dos efeitos causados no processo quando aplicadas perturbações.

Realizaram-se várias simulações dinâmicas onde foram introduzidas perturbações ao sistema, gerando-se um conjunto de dados em regime transiente. Simulou-se várias condições de entrada na coluna, com vazões, temperaturas e composições distintas, analisando-se as variáveis de saída do processo. Na Figura 4.7 é apresentado o ambiente gráfico desenvolvido com telas ilustrativas do programa HYSYS<sup>®</sup>. Destaca-se o período transiente e o comportamento das malhas de controle, após aplicadas perturbações no processo na temperatura e composição de alimentação, por exemplo.

As variáveis manipuladas na corrente de alimentação, como composição ( $X_a$ ), vazão ( $Q_a$ ) e temperatura ( $T_a$ ), influenciam diretamente as variáveis de estado (temperaturas dos pratos), que por sua vez, afetam as composições da coluna. Foram realizadas perturbações nas variáveis manipuladas da alimentação para conhecermos o comportamento do processo em regiões próximas do ponto de operação em que a unidade poderá operar.

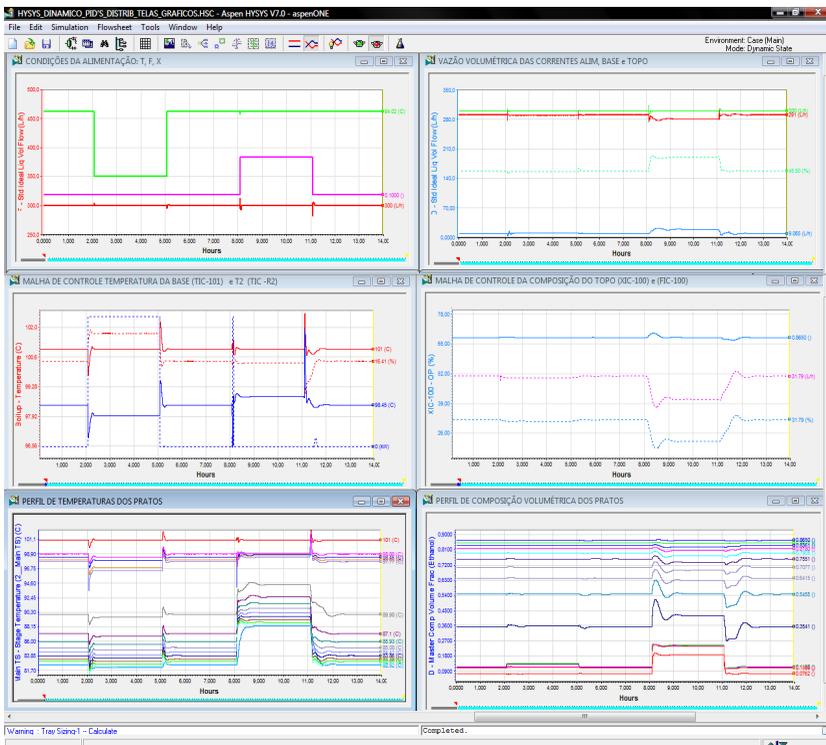


Figura 4.7. Tela do programa HYSYS® com a interface gráfica desenvolvida destacando os períodos transientes após aplicações de perturbações na temperatura e composição de alimentação.

Similarmente ao que foi realizado nas simulações estáticas, utilizou-se os dados experimentais da planta piloto, desta vez no regime transiente, para validar as simulações. Para tal foram utilizados os perfis de temperaturas experimentais, já que no período transiente não se tinha coleta de amostras de produtos de base e topo, para realização de análises de composição. Nesta seção são apresentadas as respostas destas simulações confrontando-as com as respostas obtidas experimentalmente. Tanto na aquisição de dados experimentais, quanto dos dados obtidos pelo simulador, o tempo de amostragem utilizado nesta etapa de validação foi de 1 segundo, para garantir que todos os efeitos sejam ilustrados.

Na Figura 4.8 pode ser observada a comparação dos perfis simulados de temperatura da coluna *versus* resultados experimentais da

unidade piloto de destilação com a configuração de controle distribuído. Os dados referem-se ao período em estado estacionário antes da aplicação das perturbações, porém já em ambiente dinâmico, visando validação das simulações. Neste mesmo gráfico pode-se observar as composições volumétricas de etanol ao longo da coluna na fase líquida, para cada temperatura do prato, apenas no caso simulado. A alteração da inflexão das curvas que os perfis de composição e temperatura apresentam no estágio 4, pode ser explicado pela presença da alimentação neste local. Também caracteriza-se pela diferenciação entre as seções de esgotamento e enriquecimento.

Pela análise dos resultados da Figura 4.8 verifica-se que na seção de esgotamento, onde é inserida maior quantidade de energia, tem-se um perfil de temperaturas levemente superior na simulação, quando comparado com o experimental. Este fato pode ser atribuído as perdas de energia para o ambiente externo, principalmente na base da coluna quando empregado o aquecimento distribuído no prato 2. Estas perdas ocorrem no caso experimental, mesmo utilizando uma camada de manta de lã de vidro como isolante, conforme estudo específico nesta mesma unidade, realizado por Oliveira *et al.* (2007). Nas simulações não foram aplicados modelos de perdas de calor.

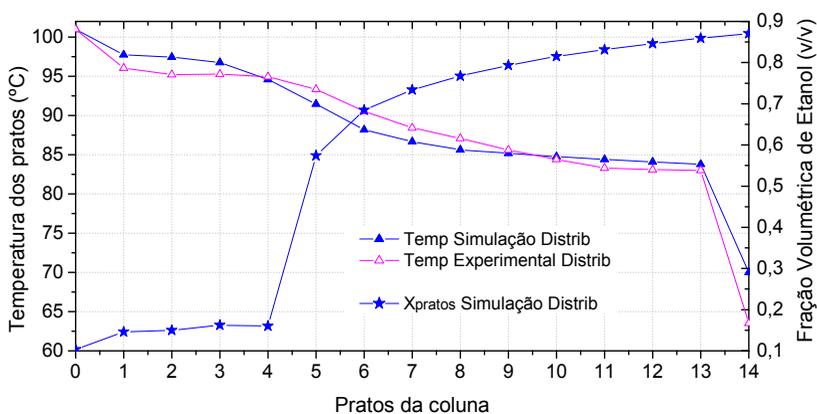


Figura 4.8. Comparação entre os perfis de temperatura da coluna: modelo *versus* experimental, ambos para o controle distribuído, e ilustração do perfil de composição para o modelo.

Além das perdas de calor para o ambiente, outros parâmetros e efeitos experimentais são difíceis de serem reproduzidos através de simuladores na sua totalidade. Como por exemplo, a fluidodinâmica da

coluna e a dinâmica de aquecimento e resfriamento da estrutura de aço da unidade, principalmente no período transiente quando aplicada alguma perturbação no processo. Em função disto, ao se analisar e comparar os perfis obtidos a partir dos resultados experimentais e simulados observa-se concordância entre os dados, com desvios do experimental em torno de 2,5% nos pratos 1 e 2 e de 4% no sexto prato. A diferença obtida entre os perfis de temperatura é considerada pequena, reproduzindo a condição experimental.

Com a simulação ajustada em modo dinâmico e em malha fechada, foram aplicadas duas perturbações na corrente de alimentação de mesma magnitude às realizadas durante os experimentos na unidade piloto, variando-se a temperatura e vazão. Nas simulações foram obtidos os perfis das variáveis de interesse ao longo do tempo no regime transiente, tais como temperaturas da base, do topo e do prato 2. Estes dados foram validados através da comparação com dados experimentais em situação idêntica.

Vale destacar que, também nesta seção, o objetivo maior não é avaliar o efeito causado pela perturbação nas variáveis do processo (detalhado em WERLE, 2007), mas sim, se os modelos são capazes de reproduzi-los. Para tal, foi analisada principalmente a coerência das respostas obtidas pelo software.

No entanto, alguns parâmetros do processo considerados importantes para caracterização das duas abordagens de controle avaliadas, serão discutidos, tais como: verificação de retorno ao ponto de operação e o impacto da perturbação no sistema. No transcorrer da avaliação de cada malha de controle, serão citados os ganhos da configuração de controle distribuída sobre o processo convencional, quando estes ocorrerem.

#### 4.2.1. 1 Perturbação na temperatura de alimentação

Nos processos químicos e petroquímicos, a destilação geralmente recebe o material proveniente de uma etapa anterior. Existem variações na carga térmica, composição e vazão, uma vez que a produção na fase anterior nunca é constante. Modificações na temperatura da alimentação refletem em alterações na temperatura de todos os estágios. Esta situação conduz a um novo estado de equilíbrio, afetando a separação (MARANGONI, 2005).

A perturbação realizada na temperatura da alimentação para comparação entre o comportamento experimental e simulado foi um degrau negativo de 12°C, ou seja, na simulação alterou-se o valor do *set*

point desta corrente de 94 para 82°C (-13%). O momento da aplicação da perturbação na escala de tempo do gráfico é de 0,5 h, conforme Figura 4.9.

Nas legendas das Figuras e ao longo do texto a configuração com o controle distribuído aplicado no prato 2 para o caso experimental será denominado: *experimental distribuído* e, para o caso das simulações com HYSYS® (modelo): *simulação distribuída*. Similarmente, para a configuração de controle convencional: *experimental convencional* e *simulação convencional*.

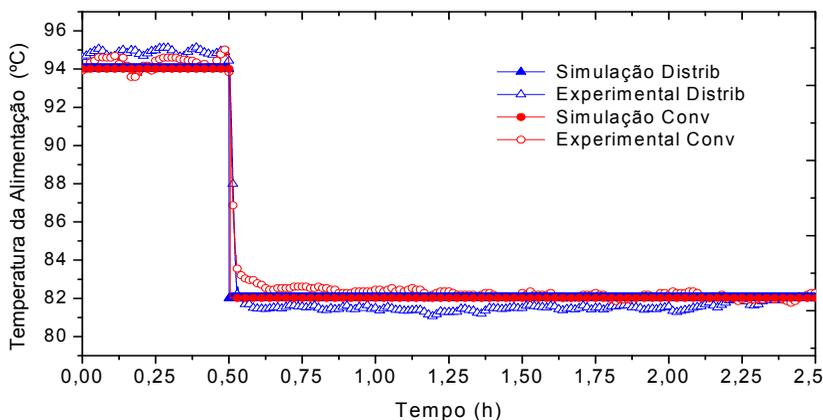


Figura 4.9. Perturbações aplicadas na temperatura da alimentação: comparação entre experimental e simulação dinâmica, para os casos de controle convencional e distribuído.

Neste gráfico não é apresentada a variável manipulada (abertura de válvula de água/vapor), pois nos testes experimentais ela foi controlada manualmente, devido a problemas de super dimensionamento da válvula da coluna, na ocasião dos testes. Já na simulação, não houve necessidade de implementar a malha de controle da temperatura da alimentação, pois este valor foi fixado no software na condição de operação do processo.

Os resultados apresentados a seguir mostram o efeito da perturbação na variável controlada para rejeitar a perturbação nas principais malhas de controle de uma coluna. Nos gráficos subsequentes é apresentado o comportamento do sistema 30 minutos antes da perturbação e 2 horas após. O principal objetivo é avaliar a resposta do modelo desenvolvido no software HYSYS® comparativamente com os dados experimentais. Como os dados que se deseja usar para

treinamento da rede neural são oriundos do modelo com controle distribuído, o foco principal é a validação desta concepção. Porém, para desenvolver a simulação dinâmica com controle distribuído foi necessário a elaboração do modelo convencional, e como os dados experimentais para esta concepção estavam disponíveis, optou-se por validar também a configuração convencional, para demonstrar a acurácia e reprodutibilidade de processos que o software proporciona.

A primeira malha de controle avaliada após a perturbação na temperatura da alimentação é a de controle da temperatura da base, ilustrada na Figura 4.10. Nesta mesma figura e nas subsequentes pode-se observar sempre uma ampliação da região próxima ao momento da perturbação, para melhor observação do efeito da perturbação nas variáveis analisadas e, principalmente, as respostas dos modelos.

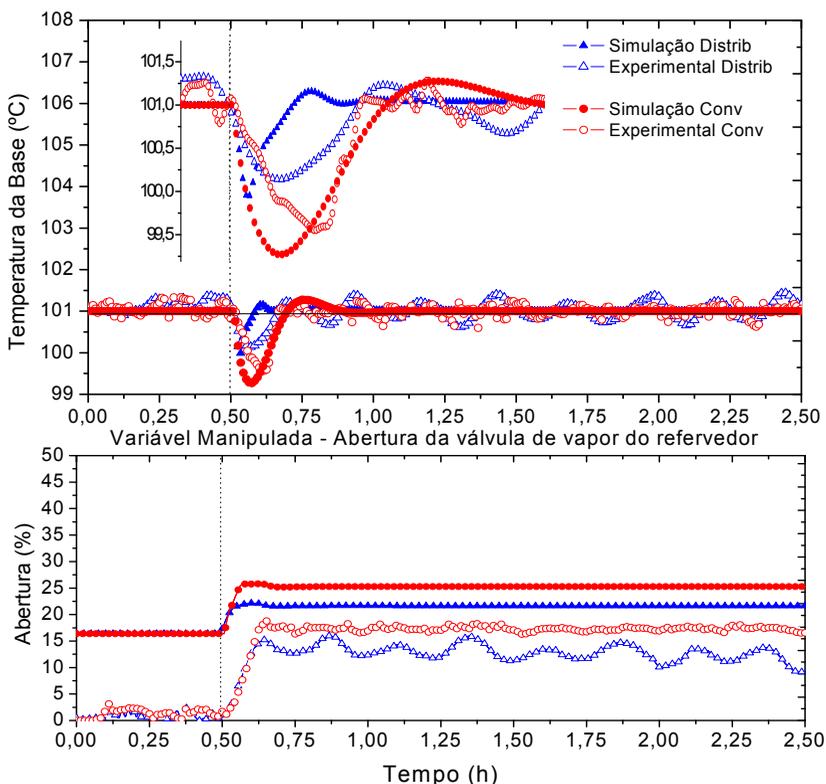


Figura 4.10. Comportamento da malha de controle da temperatura da base, antes e depois de aplicada a perturbação na temperatura da alimentação, comparação entre experimental e simulação dinâmica.

Observa-se na Figura 4.10 que os perfis de temperatura seguem a mesma tendência, quando confronta-se a simulação com aquecimento distribuído com os dados experimentais desta mesma configuração. O mesmo acontece na comparação da resposta do modelo convencional com os dados experimentais do processo com controle convencional. A perturbação do processo foi rejeitada um pouco mais rápido nos modelos, tanto distribuído quanto convencional, quando comparados com o experimental. Este comportamento pode ser atribuído à fluidodinâmica do processo, pois no caso da coluna experimental, as paredes de aço possuem uma dinâmica de aquecimento e resfriamento da estrutura mais lenta, principalmente quando aplicada alguma perturbação, difícil de ser reproduzida com exatidão em simulações.

Na Figura 4.10, observa-se que a variável manipulada do modelo também teve um comportamento similar ao experimental, com grande reprodutibilidade, com aberturas maiores da válvula do refervedor para o caso convencional, onde o decréscimo de temperatura foi maior, em torno de  $1^{\circ}\text{C}$ . Com o uso do controle com aquecimento distribuído, tanto simulado quanto experimental, verifica-se claramente o menor impacto da perturbação, além de retornar ao valor desejado mais rapidamente, reduzindo os transientes de operação.

O efeito da perturbação no controle da temperatura do último estágio é mostrado na Figura 4.11. Observa-se claramente a diferença entre a resposta do sistema frente à perturbação quando aplicado a configuração com o controle distribuído, pois o sistema rejeita a perturbação e retorna ao valor de referência ( $83^{\circ}\text{C}$ ), o que não ocorre na configuração convencional, a qual mantém um desvio permanente em torno de  $1^{\circ}\text{C}$ . O não retorno ao valor de referência é decorrente de uma alimentação subresfriada, o que implica numa quantidade menor de vapor dentro da coluna, consequentemente uma obtenção de produtos fora da especificação em relação ao estado estacionário anterior. Em uma unidade industrial este fato não seria aceitável.

As respostas dos modelos, tanto com controle distribuído quanto convencional, foram coerentes e conseguiram representar a dinâmica do processo com grande reprodutibilidade, como se pode observar na ampliação da Figura 4.11. Observa-se que o decréscimo da temperatura do topo com a simulação distribuída foi similar ao experimental distribuído ( $\pm 1,5^{\circ}\text{C}$ ). O tempo para retornar ao valor de referência foi ligeiramente menor para o modelo quando comparado ao experimental. No processo convencional o modelo também estabilizou mais rapidamente, quando comparado com o experimental, que obteve

uma resposta um pouco mais lenta, porém permaneceram no mesmo patamar.

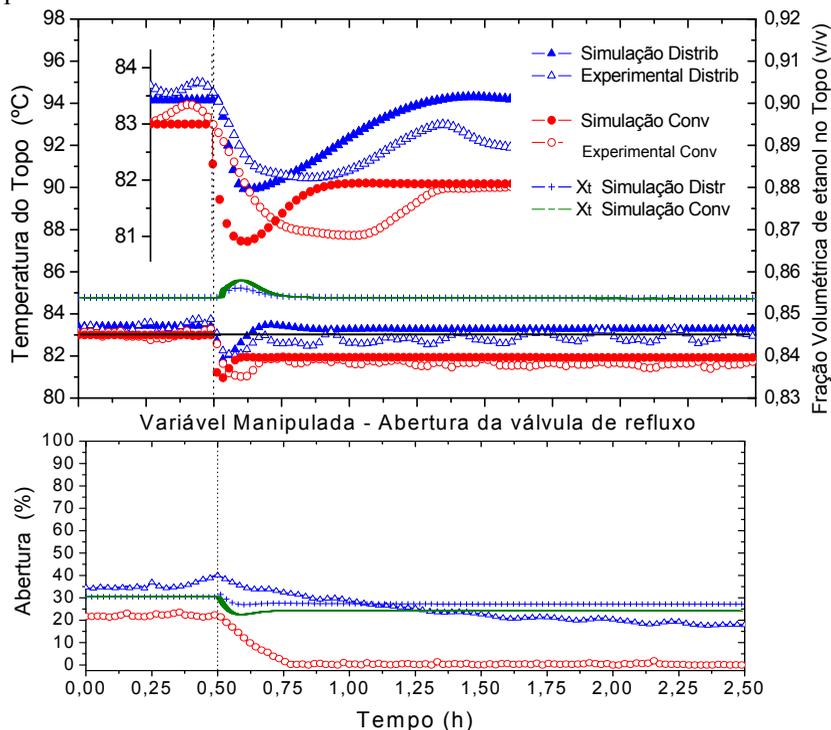


Figura 4.11. Comportamento das malhas de controle da temperatura (experimental) e da fração de etanol (simulação) do topo da coluna, antes e depois de aplicada a perturbação na temperatura da alimentação, comparação entre experimental e simulação dinâmica.

A variável manipulada, tanto da malha de controle da temperatura (experimental) quanto da composição de topo (modelo) é a válvula da vazão de refluxo. Na Figura 4.11, para os casos experimentais (distribuído e convencional) observa-se que devido a diminuição da temperatura de topo, as válvulas fecham gradativamente reduzindo o refluxo. No caso das simulações, também observa-se comportamento similar das variáveis manipuladas, devido à diminuição da temperatura interna. Ocorre um aumento da composição de etanol do produto de topo. Os percentuais que as válvulas fecharam, quando comparados *experimental distribuído* com *experimental convencional* e

*simulação distribuído* com *simulação convencional* são proporcionais ao impacto da perturbação nos sistemas.

A redução da abertura das válvulas de refluxo é menor para a configuração distribuída (experimental e simulação), pois o sistema foi menos afetado que o convencional, mantendo as vazões de refluxo na ordem de  $30 \text{ L.h}^{-1}$ . Neste aspecto, as simulações foram representativas e produziram o efeito que se esperava. Mesmo sendo malhas de controle distintas, pode-se considerar a comparação válida, pois o objetivo do controle é o mesmo, apenas o experimental é realizado via inferência. Outra vantagem observada do controle distribuído nos testes experimentais é o fato de manter uma vazão de refluxo mais elevada em relação ao convencional, que após a perturbação permaneceu com refluxo mínimo.

Nesta situação, de redução da vazão de refluxo, outra malha afetada diretamente é a malha de controle do nível do tanque acumulador ou condensador. Na Figura 4.12, tem-se o comportamento desta malha para o caso simulado com controle distribuído. Verifica-se que ocorre um aumento do nível do condensador, em função do decréscimo do refluxo, fazendo com que o controle atue, abrindo instantaneamente a válvula de produto de topo, até uma posição estável para manter o nível no valor desejado (70%). Este fato implica um aumento de produção de produto de topo, passando de 9 para  $10 \text{ L.h}^{-1}$ , valor similar a produção da unidade experimental.

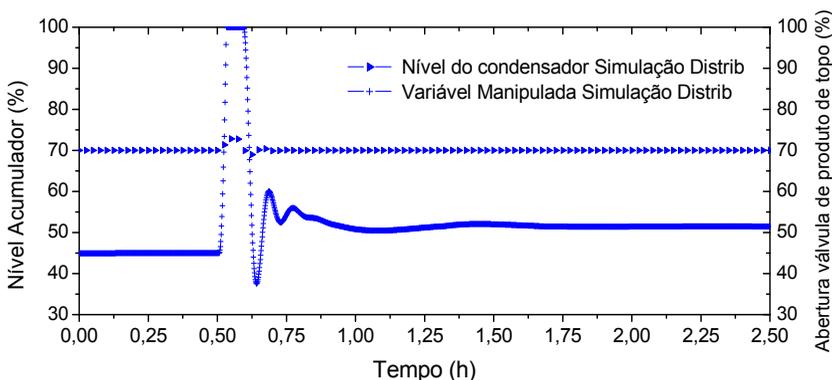


Figura 4.12. Comportamento da malha de controle do nível do condensador, antes e depois de aplicada a perturbação na temperatura da alimentação para a simulação dinâmica distribuída.

Na abordagem distribuída, a adição de energia no prato 2 promove maior vaporização do líquido da alimentação quando em contato com a resistência, resultando em menor perturbação do nível de líquido da base, em comparação com a estratégia convencional, conforme pode-se verificar na Figura 4.13. Pela análise dos resultados obtidos nas simulações, percebe-se que o nível é mais afetado no convencional, necessitando de maior abertura de válvula de produto de fundo, liberando parte do vapor que foi condensado na coluna (mantendo *set point* em 35%) decorrente da perturbação negativa na carga. Com isso conclui-se que a fase vapor no interior da coluna é menos prejudicada com o uso do aquecimento distribuído. Este mesmo efeito foi observado experimentalmente, conforme descrito por Werle (2007).

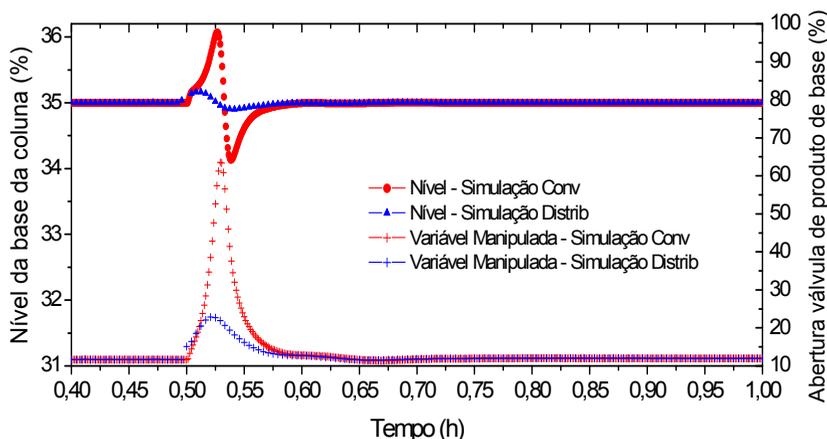


Figura 4.13. Comportamento da malha de controle do nível da base, antes e depois de aplicada a perturbação na temperatura da alimentação, para a simulação dinâmica distribuída e convencional.

Os perfis de temperatura do prato 2 no interior da coluna e a potência dissipada pela resistência perante a perturbação na temperatura da alimentação são ilustrados na Figura 4.14, para os casos simulados e experimentais. Observa-se que os perfis experimentais são inferiores aos perfis do modelo, estando defasados em aproximadamente 1°C. Este fato pode ser atribuído às perdas de energia para o ambiente externo que ocorrem ao longo da coluna experimental, conforme já comentado anteriormente na análise da Figura 4.8.

Analisando-se os perfis de temperaturas do prato 2, observa-se que logo após a perturbação, ocorre um decréscimo da temperatura do

prato para todas as situações analisadas, seguindo o mesmo comportamento da base e topo. Porém, verifica-se que a configuração que emprega o controle distribuído rejeita a perturbação e retorna ao valor desejado, tanto no caso *experimental* quanto no *modelo*, sendo similar também o tempo deste retorno (12 min), fato que não ocorre com a configuração convencional (*experimental e modelo*). O motivo se torna aparente ao analisar a variável manipulada do controle distribuído. Para manter o valor de referência do prato 2, a resistência dissipa calor imediatamente após a perturbação, mantendo-se com aproximadamente 70% (*experimental distribuído*) e 100% (*modelo distribuído*) de sua potência.

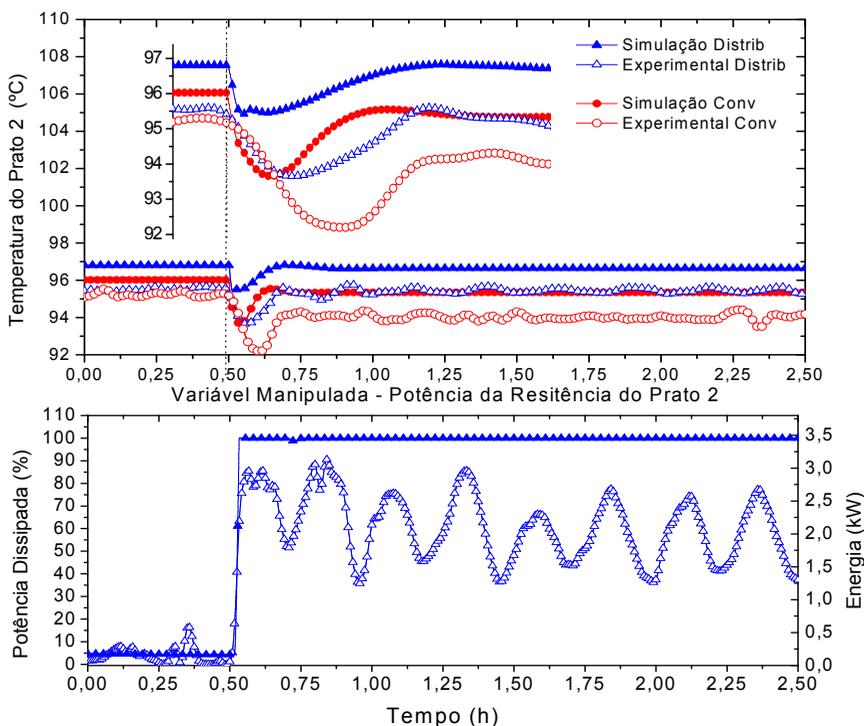


Figura 4.14. Comportamento da malha de controle da temperatura do prato 2 com controle distribuído aplicado no prato 2 e sem ação de controle no caso convencional, para a perturbação na temperatura da alimentação, comparação entre experimental e simulação dinâmica.

Quando comparado as respostas do modelo com os resultados experimentais, verifica-se novamente que eles reproduzem a tendência

das curvas e refletem o realismo do modelo empregado, tanto para reproduzir a coluna convencional, quanto a coluna com a concepção de ação distribuída de calor.

Na Figura 4.15 observa-se os perfis simulados das temperaturas dos pratos 2, 4 e 7 e suas respectivas derivadas em relação ao tempo, para o caso da simulação distribuída. Os pratos 2 e 7 foram escolhidos de acordo com os resultados de análise de sensibilidade. Desta forma, cada um representa uma seção da coluna e o prato 4 devido a ser o prato onde ocorrem as maiores flutuações decorrente da alimentação do processo ser justamente neste estágio. Pela análise das derivadas têm-se uma estimativa da duração do período transiente do processo. Como estes são os pratos mais sensíveis da coluna, o período de tempo no qual a derivada variar substancialmente é equivalente ao período transiente onde as alterações das variáveis são mais significativas.

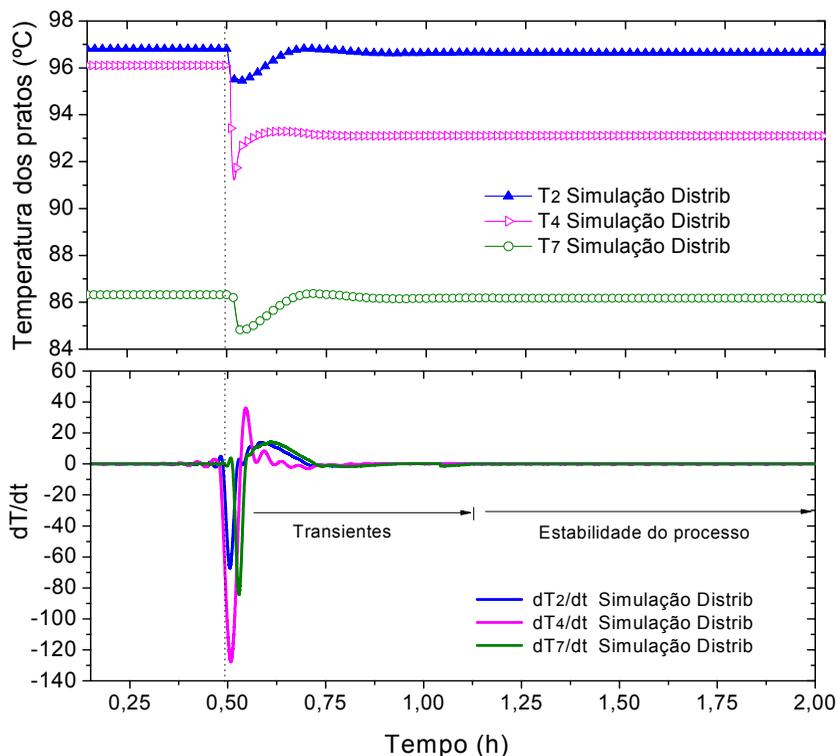


Figura 4.15. Perfis de temperatura dos pratos 2, 4 e 7 e suas respectivas derivadas em relação ao tempo para a simulação distribuída, quando aplicada a perturbação na temperatura de alimentação.

Pela análise dos resultados da Figura 4.15, verifica-se que em aproximadamente 30 minutos após a perturbação, as derivadas das temperaturas dos pratos 2, 4 e 7, em relação ao tempo, não apresentam mais variações significativas, ou seja, as variáveis alcançaram em torno de 95% do valor do estado estacionário. Desta forma, pode-se considerar atingido o regime permanente. Percebe-se também certo atraso na resposta do prato 7, em relação ao prato 2 e 4, em função do efeito de propagação prato a prato.

Na Figura 4.16 estão representados os perfis simulados de frações volumétricas de etanol em todos os pratos da coluna, antes e após a perturbação, obtidos a partir do modelo com controle distribuído. Devido ao decréscimo das temperaturas internas da coluna logo após a perturbação, verifica-se um aumento da composição de etanol neste período, propagado ao longo da coluna com um fator de amortecimento da base para o topo, em função de ser uma operação em estágios. A diminuição da temperatura dos pratos favorece o enriquecimento da fase vapor do componente mais volátil. O efeito é mais rápido na seção de esgotamento em virtude da propagação na fase líquida, a qual é maior nesta região da coluna. Na base e no topo da coluna, onde existe o controle de temperatura e composição, não ocorre grande variação e a visualização do efeito no topo é prejudicada em virtude da ampla escala do gráfico.

Para corroborar com a análise dos resultados da Figura 4.15 e Figura 4.16, construiu-se um gráfico com as derivadas das composições de etanol ao longo do tempo, nos pratos 2, 4 e 7, conforme ilustrado pela Figura 4.17.

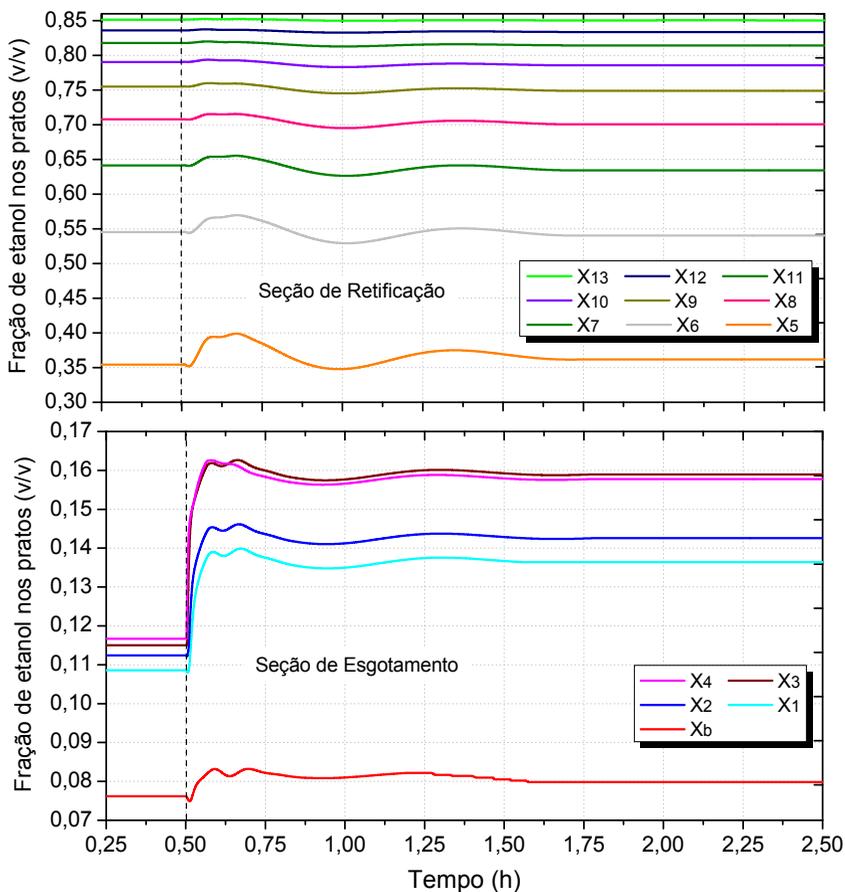


Figura 4.16. Comportamento das frações de etanol em todos os pratos da coluna, antes e depois da aplicada a perturbação na temperatura da alimentação para simulação com controle distribuído.

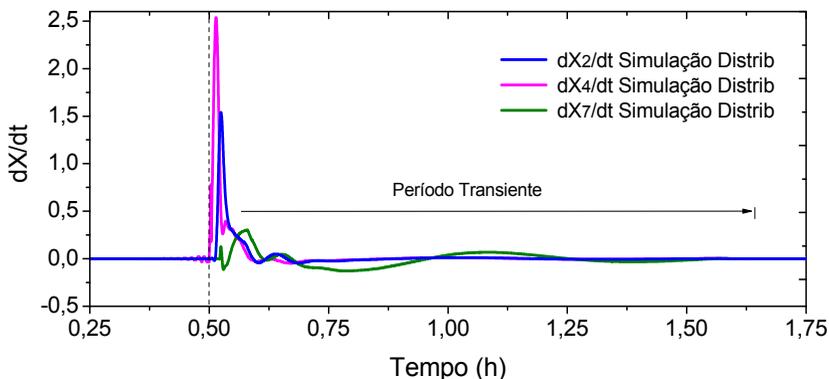


Figura 4.17. Derivada das frações volumétricas de etanol em relação ao tempo, antes e depois de aplicada a perturbação na temperatura da alimentação, para simulação com controle distribuído.

A partir deste gráfico, pode-se definir com maior precisão o alcance do regime permanente de operação. Percebe-se que não ocorrem grandes variações na composição destes pratos aproximadamente 45 minutos após a perturbação. Novamente este tempo é referente ao alcance de 95% do valor do regime estacionário.

#### 4.2.1.2 Perturbação na vazão de alimentação

Este tipo de perturbação pode ocorrer constantemente em colunas de destilação, em virtude da demanda a ser processada dependendo das condições de mercado, ou ainda, mais freqüentemente, decorrentes das oscilações de processos que antecedem seu processamento. O estado estacionário de uma coluna, considerando uma eficiência de pratos constante, está diretamente relacionado com a vazão de carga (RIGGS, 1998).

A perturbação realizada na vazão da alimentação foi um degrau positivo com intensidade de  $150\text{L}\cdot\text{h}^{-1}$ , alterando-se o valor do *set point* desta corrente de 300 para  $450\text{L}\cdot\text{h}^{-1}$  (+50%), conforme ilustrado na Figura 4.18.

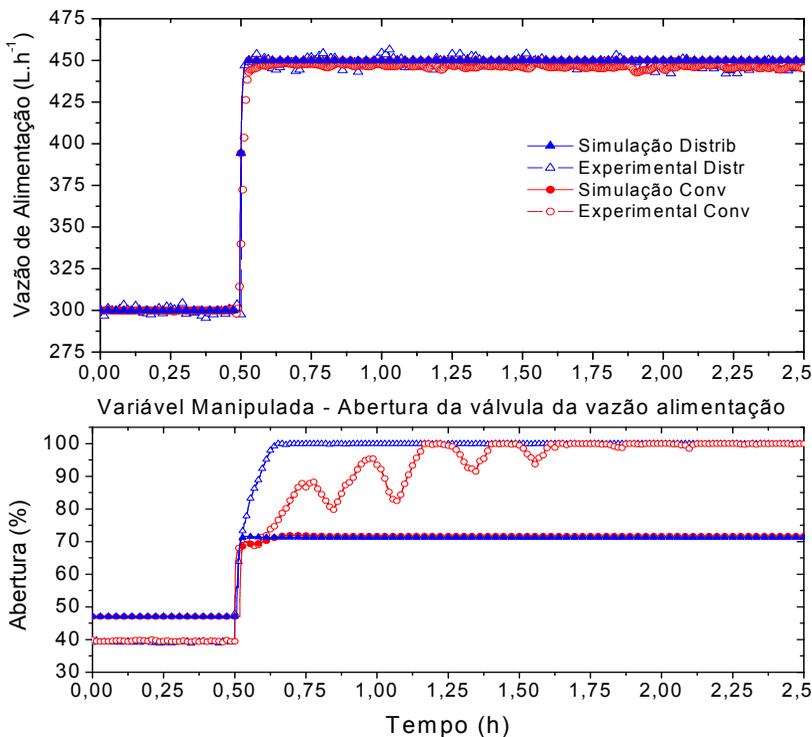


Figura 4.18. Perturbações aplicadas na vazão da alimentação, comparação entre experimental e simulação dinâmica para os casos de controle convencional e distribuído.

Para este tipo de perturbação, ao aumentar-se a vazão e manter a temperatura de alimentação constante, é inserida uma maior carga térmica no interior da coluna. O impacto inicial da perturbação no sistema é caracterizado pelo decréscimo das temperaturas em função do aumento da vazão. Porém, após o sistema assimilar a perturbação, os perfis de temperatura tendem a aumentar para um patamar levemente superior ao anterior. A ação das malhas de controle deve ser efetiva o suficiente para restabelecer o perfil desejado.

O efeito da perturbação da vazão da alimentação na malha de controle da temperatura do refeedor pode ser observado pela análise dos resultados da Figura 4.19. Verifica-se o decréscimo da temperatura da base para todas as situações estudadas. Porém, observa-se que o mesmo foi maior quando empregada a configuração convencional (em torno de 1,75°C) se comparado com o resultado obtido pelo controle

distribuído (em torno de  $0,7^{\circ}\text{C}$ ), como se pode observar na ampliação da seção. Verifica-se que aproximadamente  $0,25$  h depois da perturbação os valores se estabilizaram em torno do *set point* ( $101^{\circ}\text{C}$ ), para ambas as configurações, com os dados experimentais mais oscilatórios devido à característica do ajuste dos controladores da unidade experimental.

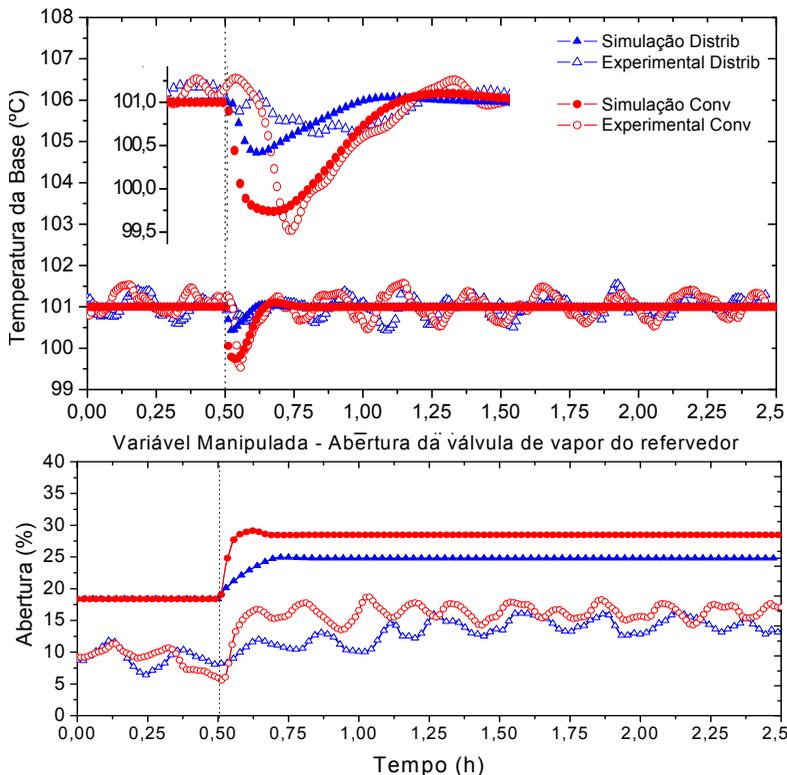


Figura 4.19. Comportamento da malha de controle da temperatura da base, antes e depois de aplicada a perturbação na vazão da alimentação, comparação entre experimental e simulação dinâmica.

O modelo conseguiu representar de forma coerente estes efeitos para as duas configurações. A variável manipulada do modelo também reproduziu a tendência do experimental após a perturbação, mantendo aberturas menores na válvula de vapor para o caso distribuído, fato justificado pela fonte adicional de calor no prato 2 (Figura 4.23, variável manipulada). No caso das simulações, o patamar de abertura da válvula (em torno de 18%) antes da perturbação, foi diferente dos 10% no caso

experimental. Porém este fato é aceitável, pois o dimensionamento das válvulas é distinto. O importante nesta análise é o comportamento da variável.

Da mesma forma que na seção de esgotamento, a seção de retificação é afetada pela perturbação da vazão de alimentação em função do decréscimo do perfil de temperatura. Na Figura 4.20, verifica-se o efeito da perturbação na malha de controle da temperatura do último estágio (para o caso experimental) e na malha de controle da composição de etanol no topo (para o modelo).

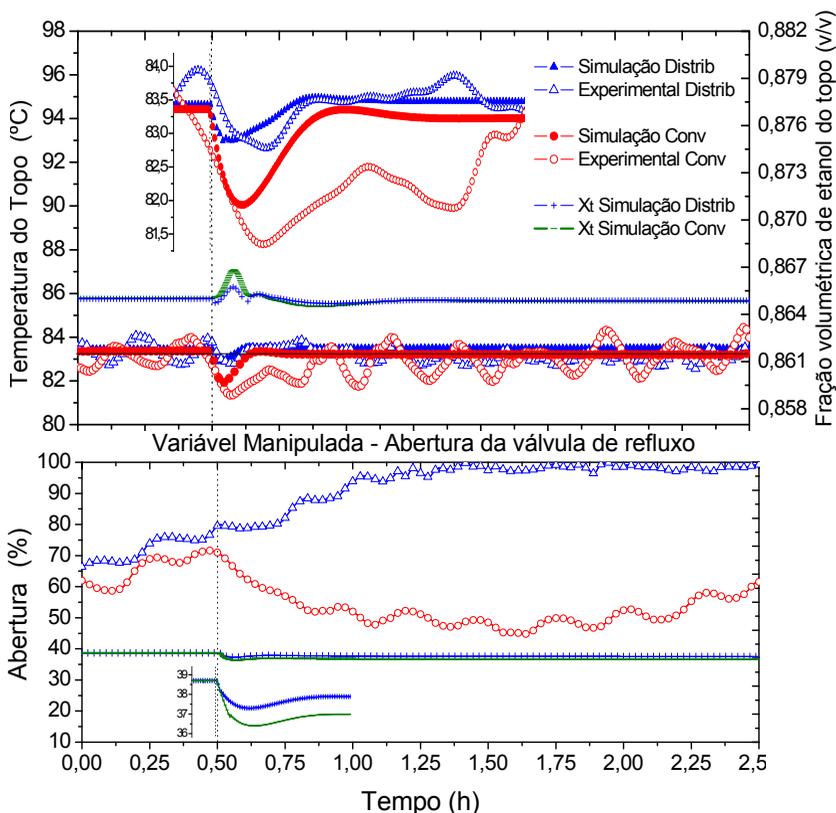


Figura 4.20. Comportamento das malhas de controle da temperatura (experimental) e da fração de etanol (simulação) do topo da coluna, antes e depois de aplicada a perturbação na vazão da alimentação, comparação entre experimental e simulação dinâmica.

A perturbação na vazão da alimentação afeta de forma mais acentuada a vazão de vapor que ascende os pratos no interior da coluna e, conseqüentemente, seu equilíbrio. Este fato pode ser observado pela análise da Figura 4.20, para os testes experimentais e simulados com as duas configurações de controle: convencional e distribuída. Novamente percebe-se que o sistema é menos afetado pela perturbação quando se emprega o controle distribuído.

Na comparação entre o controle convencional, observa-se que no caso experimental o retorno ao valor de referência é um pouco mais lento e o decréscimo de temperatura foi ligeiramente maior (0,7 °C), em relação ao modelo do HYSYS<sup>®</sup>. Já no controle distribuído as respostas do modelo frente ao experimental foram bem próximas. Alguns efeitos não foram reproduzidos em sua totalidade pelo modelo, devido à dificuldade em reproduzir a dinâmica, porém as tendências dos perfis novamente foram semelhantes, o que indica o bom desempenho do *modelo* em descrever o comportamento do processo. Salientando-se que o objetivo principal é a validação do modelo com a proposta do aquecimento distribuído, o qual obteve desempenho adequado frente aos dados experimentais.

Na malha de controle de composição de topo, observa-se que o modelo responde rapidamente ao acréscimo de composição de topo, devido à queda momentânea de temperatura, com uma pequena redução da abertura da válvula de refluxo (melhor observado na ampliação da variável manipulada). Observa-se que, durante o período de restabelecimento da temperatura, as composições decrescem gradualmente, o que gera novamente uma ação de controle na vazão de refluxo, verificado com aumento sutil das aberturas de válvulas, suficientes para manter a composição no valor desejado. Efeito verificado nas duas configurações simuladas.

Na Figura 4.21 observa-se o comportamento da malha de controle do nível do condensador e o efeito causado na produção de destilado, para a simulação com controle distribuído.

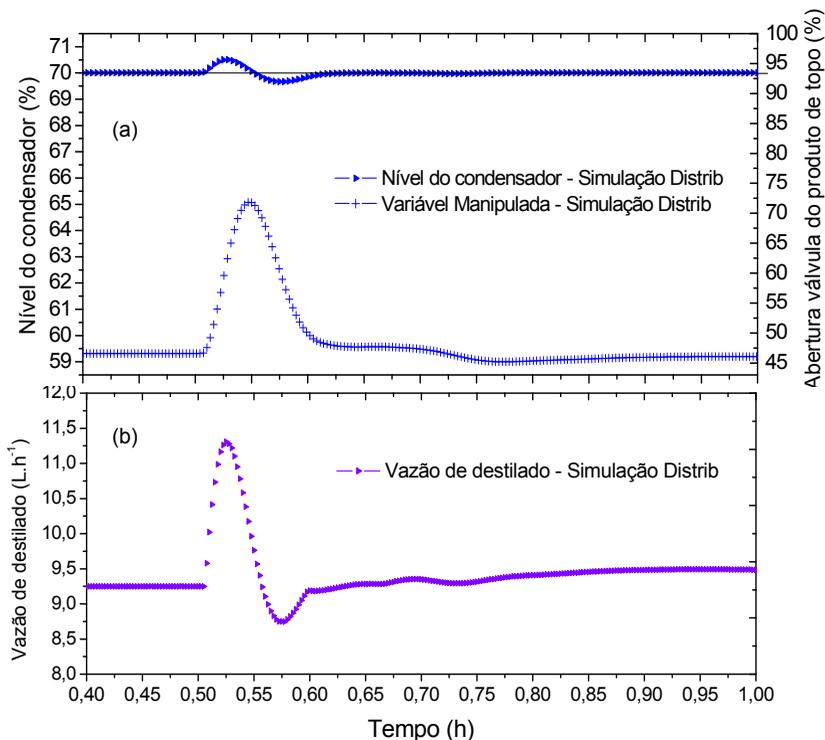


Figura 4.21. (a) Comportamento da malha de controle do nível do condensador e (b) perfil da vazão de destilado, antes e depois de aplicada a perturbação na vazão da alimentação, para a simulação dinâmica distribuída.

A redução do refluxo, causado pela perturbação na vazão, propicia um aumento do nível do acumulador, que por sua vez gera um acréscimo da produção de destilado, para manter o nível do tanque no valor desejado. Porém, este aumento é momentâneo. Após o sistema estabilizar, a vazão de destilado reduz e permanece em um valor levemente maior que antes da perturbação.

Com o aumento da vazão da carga da coluna, o efeito desta perturbação é propagado rapidamente para a seção de esgotamento. Em um primeiro momento, espera-se que o nível da base da coluna aumente, e que o sistema de controle atue de modo a manter o nível constante, abrindo a válvula de produto de fundo. Este comportamento pode ser visualizado na Figura 4.22. Nesta figura percebe-se claramente que o efeito da perturbação é menor na configuração com ação distribuída, a

qual consegue vaporizar parte do líquido que é despejado sobre o prato 2, que contém a resistência elétrica, acarretando uma maior fase vapor, absorvendo melhor o impacto da perturbação.

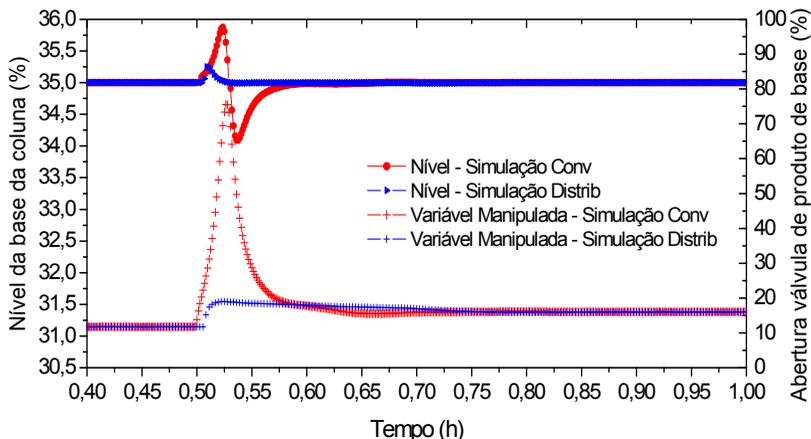


Figura 4.22. Comportamento da malha de controle do nível da base da coluna, antes e depois de aplicada a perturbação na vazão da alimentação para a simulação dinâmica distribuída e convencional.

Os perfis de temperatura do prato 2 e a potência dissipada pela resistência após a perturbação são ilustrados na Figura 4.23, para os casos simulados e experimentais. Novamente observa-se que os perfis obtidos pelo modelo são superiores aos perfis experimentais, fato atribuído as perdas de energia para o ambiente externo que ocorrem na unidade experimental, principalmente na seção de esgotamento, onde se concentra maior quantidade de energia. Estas perdas não conseguem ser reproduzidas pelo software. No modelo foi possível manter as temperaturas das extremidades da coluna idênticas as experimentais, antes e depois da perturbação (Figuras 4.19 e 4.20). No entanto, as temperaturas da seção de esgotamento mantiveram-se superiores às experimentais, o que indica que estas seriam as temperaturas desta seção, caso a unidade experimental fosse perfeitamente isolada e sem perdas para o ambiente externo.

Na ampliação da Figura 4.23 observa-se claramente que a configuração de controle distribuída assimila perfeitamente a perturbação e retorna rapidamente ao valor desejado (em torno de 8 minutos), com temperaturas levemente superiores. Já no processo com controle convencional, ocorre um decréscimo de temperatura de 3°C e

somente depois de transcorridos 20 minutos o processo retorna próximo ao seu ponto de operação. O motivo desta ação rápida e eficaz da configuração com ação distribuída frente à convencional é a adição de energia no prato 2, situado logo abaixo da alimentação, como pode-se verificar pela análise da variável manipulada da Figura 4.23. Para manter o valor de referência do prato 2, a resistência dissipa calor imediatamente após a perturbação fazendo com que a temperatura mantenha-se no valor desejado. Quando comparado as respostas do modelo com os resultados experimentais, verifica-se novamente que eles reproduzem perfeitamente a tendência das curvas, inclusive na magnitude do decréscimo das temperaturas e o tempo de retorno ao valor desejado para as duas configurações.

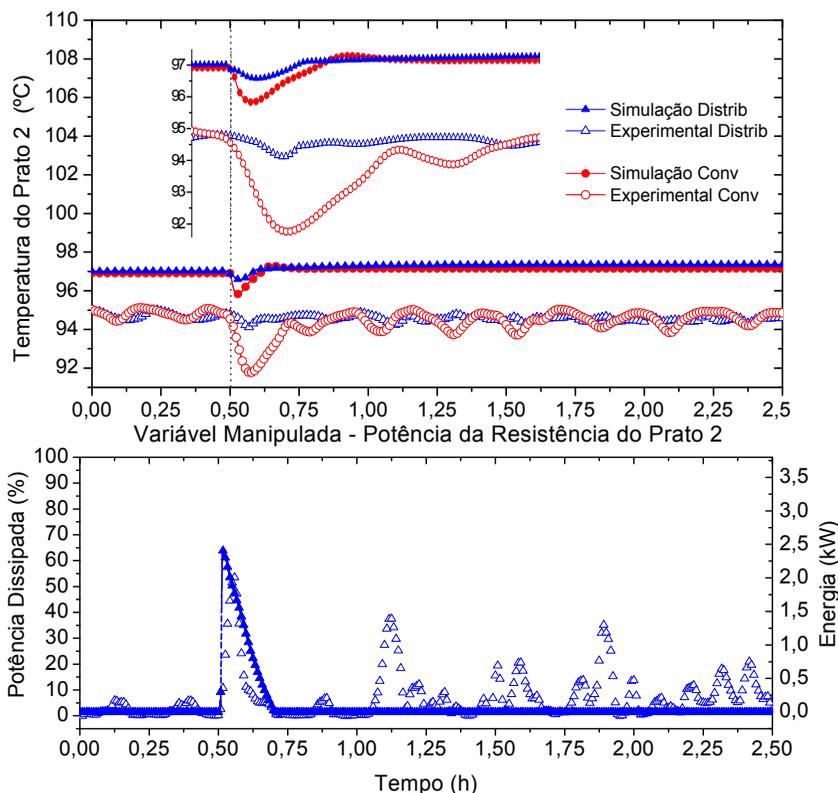


Figura 4.23. Comportamento da malha de controle da temperatura do prato 2 com controle distribuído aplicado no prato 2 e sem ação de controle no caso convencional, para a perturbação na vazão da alimentação, comparação entre experimental e simulação dinâmica.

O comportamento da variável manipulada do modelo no controle distribuído também é semelhante ao experimental, adicionando inicialmente grande quantidade de energia (em ambos os casos, em torno de 65% dos 3,5kW disponíveis). Observa-se pela análise dos dados experimentais, pequenos acréscimos de energia ao longo do tempo, os quais podem ser atribuídos as oscilações da variável controlada e ao próprio ajuste dos parâmetros do controlador. É importante ressaltar que o ajuste da malha de controle do prato 2, nos experimentos, foi realizada de forma descentralizada (ou seja, independente) em relação ao ajuste das malhas de temperatura do refulvedor e último estágio. Assim, é possível que com um controle melhor ajustado, haja diferenças nas cargas térmicas utilizadas na estratégia distribuída experimental, indicando até mesmo a possibilidade de redução da quantidade térmica necessária para rejeitar a perturbação, aproximando-se ainda mais da resposta obtida pelo modelo.

Os perfis das temperaturas dos pratos 2, 4 e 7, e suas respectivas derivadas em relação ao tempo para a simulação distribuída, podem ser visualizados na Figura 4.24.

Pela análise da derivada pode-se verificar o tempo necessário para que o processo rejeite a perturbação e retorne ao ponto de operação nominal. Observa-se que a perturbação é propagada para todos os estágios da coluna. Porém, como era esperado, seu efeito é um pouco mais pronunciado no prato 4, o prato da alimentação, fato que pode ser observado pelo maior pico da derivada.

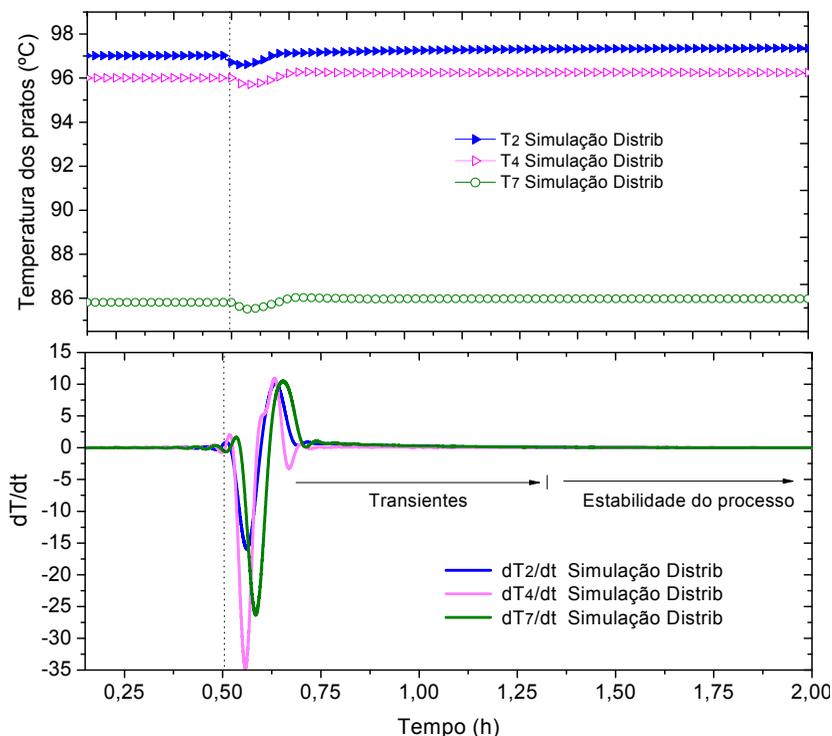


Figura 4.24. Perfis de temperatura dos pratos 2, 4 e 7 e suas respectivas derivadas em relação ao tempo para a simulação distribuída, quando aplicada a perturbação na vazão de alimentação.

Verifica-se que em torno de 30 minutos após a perturbação, a variação das derivadas das temperaturas dos pratos em relação ao tempo não é significativa, indicando que o sistema se encontra estável.

#### 4.2.1.3 Avaliação final em relação a validação das simulações dinâmicas

O estudo realizado nesta seção foi necessário para verificar a confiabilidade das respostas obtidas pelo modelo, o qual foi desenvolvido com o uso do software HYSYS<sup>®</sup> em malha fechada. Nesta etapa do projeto, as respostas obtidas com o emprego das simulações do processo de destilação foram avaliadas conceitualmente e comparadas com o comportamento conhecido da mistura binária etanol/água, frente a aplicações de perturbações na carga. Realizou-se a comparação dos

resultados simulados com os dados experimentais para as duas concepções avaliadas: controle convencional e controle distribuído aplicado no prato 2. Sendo que, a avaliação dos resultados do processo distribuído, é o principal foco da pesquisa. Desta forma, foi possível verificar a reprodutibilidade e validação do modelo.

Quando comparadas as respostas do modelo em malha fechada com as experimentais e analisando as respostas das malhas de controle da base, topo e prato 2, verifica-se que os resultados apresentados reproduzem a tendência das curvas. A magnitude do impacto causado no processo nas variáveis controladas e o tempo necessário para retorno ao valor desejado são muito semelhantes, na comparação dos resultados do modelo com o experimental para as duas configurações confrontadas. O comportamento das variáveis manipuladas do processo simulado também é coerente e se assemelha aos obtidos experimentalmente nas duas configurações. O modelo mostrou vantagem do emprego do controle distribuído sobre o convencional, fato já comprovado nos testes experimentais.

A análise dos resultados simulados mostrou que os mesmos reproduzem os efeitos observados experimentalmente, possuem a mesma tendência e são coerentes com respostas esperadas para o processo frente a determinadas perturbações, dando confiabilidade às simulações. Desta forma, consideram-se validadas as simulações realizadas e que o modelo desenvolvido é capaz de servir como uma planta geradora de dados para treinamento de uma rede neural dinâmica.

Pelos resultados obtidos quando aplicadas as derivadas nas temperaturas e composições de alguns estágios pré-selecionados da coluna verifica-se que, para as perturbações na temperatura e vazão da alimentação, um período de uma hora após aplicada a perturbação é necessário para descrever a dinâmica do processo. A frequência de aquisição dos dados (1segundo) que foi usada nesta etapa deverá ser reduzida. Serão realizados testes para identificação da menor constante de tempo do processo, para as variáveis que poderão ser usadas no treinamento da rede.

#### 4.2.2 Comportamento do processo em malha aberta

Para o desenvolvimento do sensor virtual é necessário conhecer o comportamento do processo e sua dinâmica em malha aberta no regime nominal de operação e em torno dele. A análise dinâmica do processo é importante pois permite averiguar questões relativas ao

tempo morto e tempo de assentamento das variáveis chave e secundárias do processo. Por exemplo, o tempo que a composição de topo levará para atingir o estado estacionário depois de aplicada alguma perturbação na coluna.

Para realizar o estudo dinâmico, as malhas de controle da unidade elaboradas no simulador, foram abertas, sendo implementado somente o controle do balanço de massa da coluna, visto que este possui interações fracas com o controle de composição, conforme apresentado por inúmeros estudos (SKOGESTAD, 1992). Em seguida, foram realizadas perturbações na coluna observando seu comportamento em relação às temperaturas e composições da coluna.

O controle do balanço de massa, aplicado no simulador HYSYS<sup>®</sup> é composto pelas seguintes malhas: 1) controle do nível da base através do ajuste da vazão da corrente de produto de fundo; 2) controle do nível do acumulador por meio da manipulação da vazão do produto de topo; 3) controle da vazão da alimentação em função do ajuste da vazão desta mesma corrente.

As variáveis que serão candidatas para serem empregadas como entradas no sensor para inferir a composição de topo terão atenção especial neste estudo. Estas variáveis devem ser de fácil medição e possuírem precisão nas medidas de aquisição e suas alterações devem obrigatoriamente refletir nas variações das composições dos pratos, especialmente do topo.

As variáveis manipuladas na corrente de alimentação como composição ( $X_a$ ), vazão ( $Q_a$ ) e temperatura ( $T_a$ ) influenciam diretamente as variáveis de estado (temperaturas dos pratos), que por sua vez afetam as composições da coluna. Serão realizadas perturbações nas variáveis manipuladas da alimentação para conhecermos o comportamento do processo em regiões próximas do ponto de operação em que a unidade deverá operar. Ou seja, as perturbações afetarão o desempenho da coluna, a qual migrará para novas condições (distintos estados estacionários), os quais deverão ser previamente conhecidos pela rede, e estarem dentro do conjunto de treinamento.

Os simuladores comerciais são muito eficazes em operar próximo do ponto de operação. Porém, quando submetidos a degraus de grande magnitude, possuem dificuldade na realização dos cálculos e integrações e, conseqüentemente, não convergem. O desempenho do HYSYS<sup>®</sup> em malha fechada (submetido a perturbações com grande magnitude) foi considerado adequado, sendo possível operar a planta na faixa de operação desejada. Já em malha aberta, a faixa de operação teve que ser restringida (detalhes no *Capítulo III, seção 3.3.1.5*), pois quando

aplicadas perturbações degraus com grande magnitude o processo migrava para situações que os cálculos não convergiam.

Será realizada uma análise do comportamento da coluna quando ela for submetida a perturbações do tipo degrau com amplitudes de -50% e +50% aplicadas na variável fração volumétrica de etanol na corrente de alimentação ( $X_a$ ), ou seja, seu valor passará de 0,1 (ponto de operação) para 0,05 e 0,15. Outro ponto a ser destacado nesta análise é o emprego da resistência elétrica situado no prato 2, em modo manual, com 50% de sua capacidade, ou seja 1,75kW, em ambos os casos analisados.

#### 4.2.2.1 Perturbação degrau negativo (-50%) na composição de etanol da corrente de alimentação – Malha Aberta

A primeira perturbação que o processo simulado em malha aberta foi submetido foi uma perturbação degrau negativo, onde a fração volumétrica de etanol na alimentação foi reduzida de 0,1 para 0,05%.

Com este tipo de perturbação, a coluna passa a operar de maneira a minimizar a composição do etanol no topo. Observa-se na Figura 4.25 que, após a aplicação da perturbação (0,5 h), ocorre elevação instantânea das temperaturas dos pratos na seção de esgotamento, em virtude do incremento do componente mais pesado. O efeito é imediato na seção de esgotamento em virtude da propagação na fase líquida, a qual é maior nesta região da coluna.

Devido a elevação da quantidade de água na mistura, o equilíbrio líquido vapor promove a elevação da temperatura de ebulição da mesma, aumentando seu o ponto de bolha. Devido a essa modificação, é importante ressaltar que, perturbações na composição da alimentação resultam em novos ajustes do ponto de operação do processo, pois conduzem a novos valores de composições no produto de base e de topo. Em malha aberta, o calor cedido pelo refeedor e pela resistência do prato 2 (operando com 1,75 kW, ou seja 50% do valor máximo) são mantidos constantes. Porém, esta energia para este novo estado é relativamente baixa para gerar uma vaporização adequada da mistura. As temperaturas decrescem seu valor gradualmente, nas duas seções, até atingirem um novo estado de equilíbrio. Provavelmente, nesta situação também não ocorre arraste de vapor do componente mais pesado aos pratos superiores.

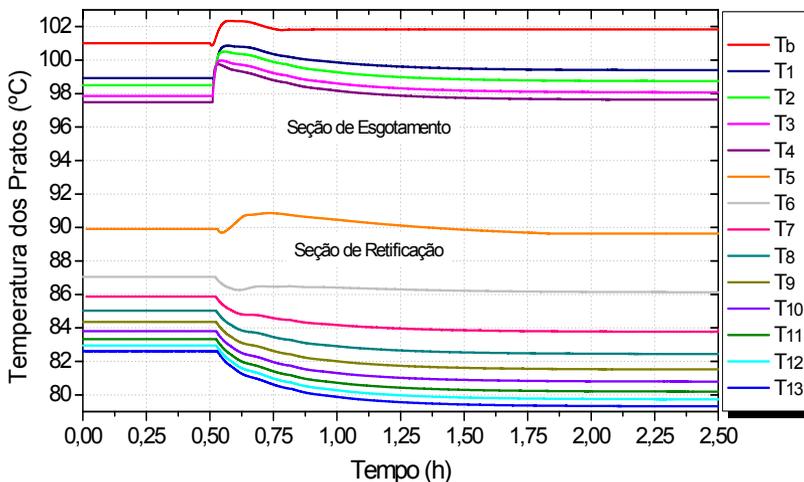


Figura 4.25. Comportamento das temperaturas em todos os pratos da coluna, com aplicação do degrau negativo (-50%) na composição volumétrica da alimentação.

Com o decréscimo das temperaturas, era de se esperar um aumento na composição do topo. Porém, em função da baixa taxa de vaporização e como a perturbação foi justamente na redução deste componente na alimentação, estes dois efeitos são mais pronunciados e a composição dos pratos reduz seu valor, conforme podemos observar na Figura 4.26. Este mesmo efeito pode ser melhor visualizado na análise da Figura 4.27, onde verifica-se que, após a perturbação, a composição de etanol é reduzida ao longo dos pratos nas duas seções. Percebe-se que no topo a redução foi menor. Observa-se também um atraso nos pratos da seção de retificação em relação à seção de esgotamento, devido à propagação do efeito a cada estágio.

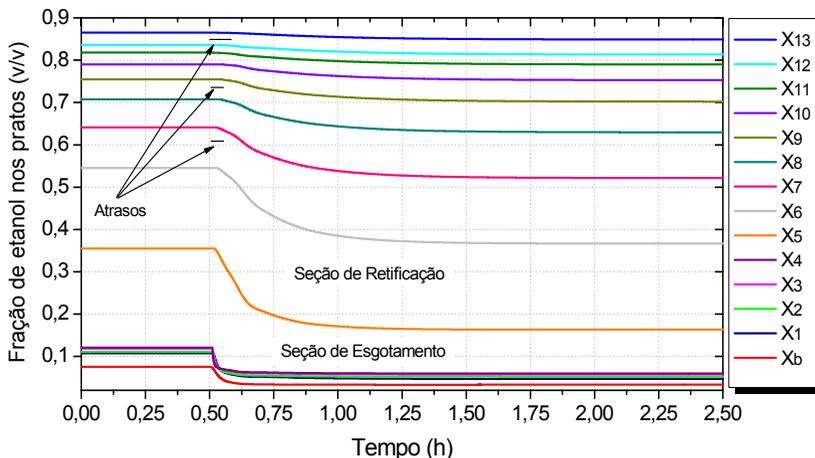


Figura 4.26. Comportamento das frações de etanol em todos os pratos da coluna, com aplicação do degrau negativo (-50%) na composição volumétrica da alimentação.

A Figura 4.27 ilustra os perfis de temperatura e composição antes da perturbação (coluna em regime nominal de operação) e após a aplicação do degrau negativo. Refere-se aos dados com a coluna já em estados estacionários distintos.

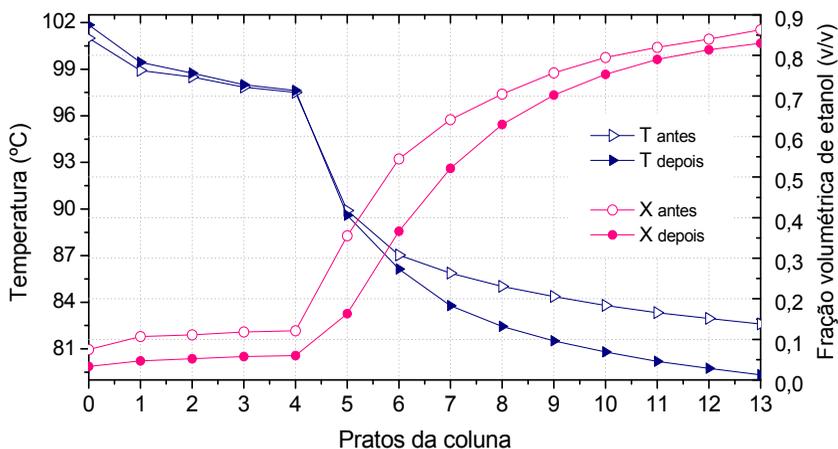


Figura 4.27. Comparação entre os perfis de temperatura e composição dos pratos da coluna, antes e depois da aplicação da perturbação de degrau negativo (-50%) na fração de etanol na alimentação.

Observa-se a redução do componente mais volátil em todos os estágios da coluna após a aplicação da perturbação. Já em relação ao perfil de temperatura, verifica-se que depois da perturbação, as temperaturas da seção de esgotamento permanecem em um patamar levemente superior, no entanto na seção de retificação ocorre inversão do perfil, com temperaturas inferiores após a perturbação, decorrente da energia adicionada ao processo ser relativamente baixa para o novo estado, impossibilitando gerar uma vaporização adequada da mistura nesta seção.

#### 4.2.2.2 Perturbação degrau positivo (+50%) na composição de etanol da corrente de alimentação – Malha Aberta

A segunda perturbação em malha aberta foi uma perturbação degrau positivo, onde a fração volumétrica de etanol na alimentação foi aumentada de 0,1 para 0,15%.

Com este tipo de perturbação, a coluna passa a operar de maneira a aumentar a composição do etanol no topo. Na Figura 4.28, observa-se que nas duas seções da coluna (esgotamento e retificação) as temperaturas diminuem seu valor, em função do incremento do componente mais volátil. Verifica-se também a resposta mais rápida da seção de esgotamento, devido a sua característica hidráulica. Em virtude da elevação da quantidade de etanol na alimentação, o ponto de bolha em cada estágio é alterado e migra para um valor inferior ao anterior à perturbação, acarretando a diminuição das temperaturas, conforme observado.

No interior da coluna o que ocorre é o acréscimo da composição de etanol na fase vapor, conforme Figura 4.29. Com esse incremento ocorre também o aumento da energia de vaporização. Desta forma, a energia necessária para evaporar a quantidade extra de etanol faz com que a temperatura migre para um novo estado estacionário inferior ao anterior. Neste caso, a energia fixa adicionada pelo refervedor e resistência do prato 2, mesmo sendo considerada em excesso em relação ao estado anterior, não é suficiente para suprir o aumento da energia de evaporação e evaporar a quantidade extra de etanol adicionada ao sistema.

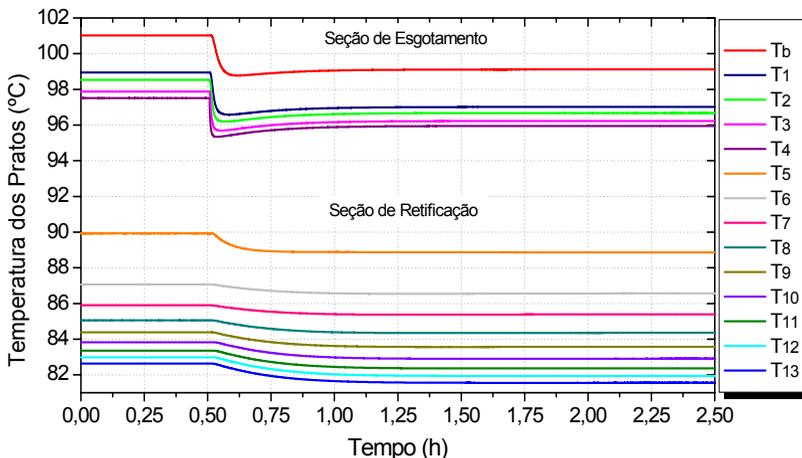


Figura 4.28. Comportamento das temperaturas em todos os pratos da coluna, com aplicação do degrau positivo (+50%) na composição volumétrica da alimentação.

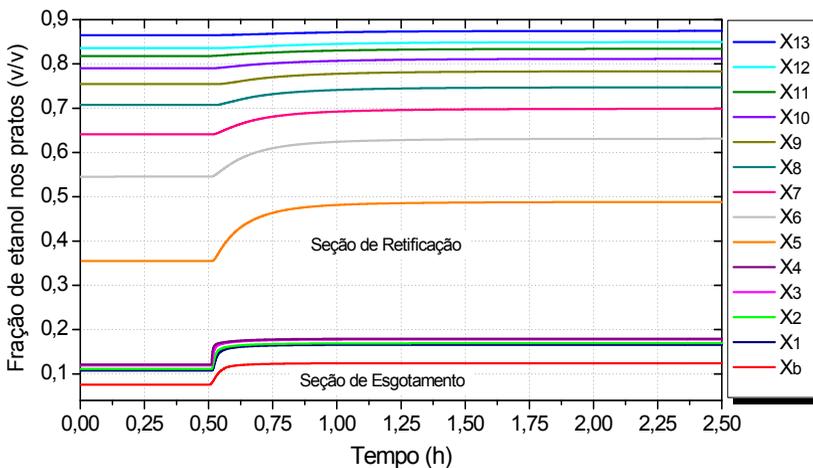


Figura 4.29. Comportamento das frações de etanol em todos os pratos da coluna, com aplicação do degrau positivo (+50%) na composição volumétrica da alimentação.

Observa-se que o incremento de etanol na alimentação propicia o incremento deste mesmo componente em todos os estágios da coluna, com um fator de amortecimento da base para o topo da coluna. O efeito é mais rápido na seção de esgotamento em virtude da propagação na

fase líquida, a qual é maior nesta região da coluna. Observa-se pelo comportamento dinâmico do processo, que o sistema atingiu o novo estado estacionário em torno de 1,5 h.

A Figura 4.30 ilustra os perfis de temperatura e composição antes da perturbação e após a aplicação do degrau negativo.

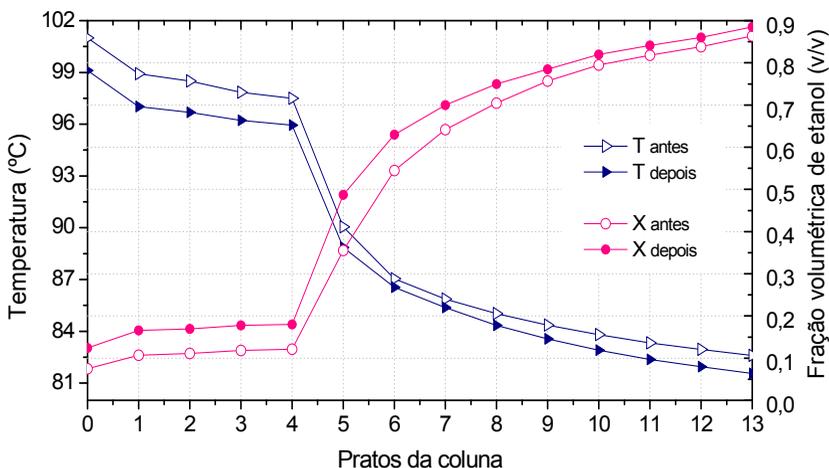


Figura 4.30. Comparação entre os perfis de temperatura e composição dos pratos da coluna, *antes* e *depois* da aplicação da perturbação de degrau positivo (+50%) na fração de etanol na alimentação.

Pela análise da Figura 4.30 é possível observar detalhadamente o efeito provocado pela perturbação em todos os estágios, principalmente o efeito nos dois extremos da coluna. Enquanto que na base da coluna ocorre uma redução de 2,5°C, no topo o efeito foi um pouco menor com redução de aproximadamente 1,5°C.

#### 4.2.2.3 Perturbação de degrau positivo (+100%) no calor cedido à resistência do prato 2 – Malha Aberta

Em virtude do efeito proporcionado pela ação da resistência do prato 2 sobre o processo, foi realizado um estudo para avaliar a dinâmica da ação distribuída em malha aberta. Para tal, foi aplicado um degrau positivo no calor cedido ao prato 2. Inicialmente o processo encontrava-se em estado estacionário ( $X_a=0,1$  [v/v];  $T_a=90$ [°C] e  $Q_a=300$  [L.h<sup>-1</sup>]) com a resistência elétrica do prato 2, em modo manual, com 50% de sua capacidade, ou seja 1,75kW.

No instante de tempo  $t=0,5$  h foi aplicado um degrau positivo no fornecimento de calor do prato 2, onde o calor cedido à resistência foi aumentado de 1,75 kW para 3,5 kW (100% da capacidade). O efeito pode ser visualizado na Figura 4.31. Nesta situação o processo migrou para um ponto de operação diferente do anterior, conforme destacado pelos perfis de temperatura.

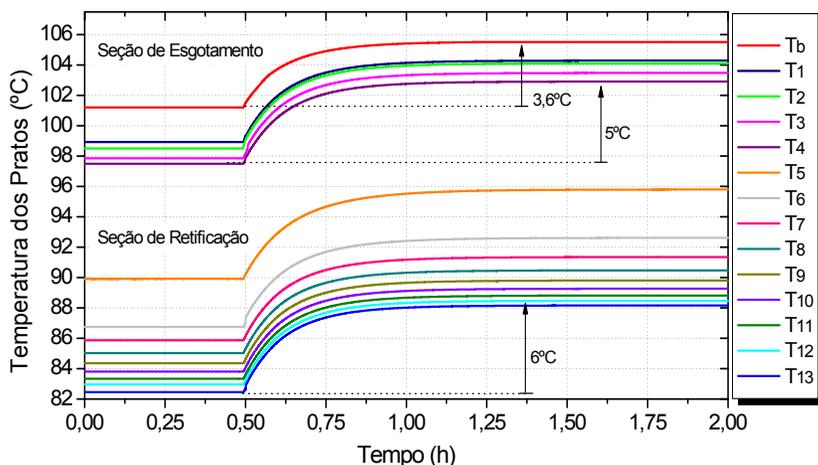


Figura 4.31. Comportamento do perfil de temperatura em todos os pratos da coluna, após aplicação do degrau positivo (+100%) com incremento no fornecimento do calor cedido ao prato 2.

Percebe-se que o impacto da perturbação, causado pelo incremento de energia ao prato 2, é diferenciado nas duas seções da coluna, sendo este menor na base (em torno de 3,6 °C) e superior no topo (6 °C), conforme destacado na Figura 4.31. Este comportamento indica que a introdução do calor no interior da coluna se propaga em relação a todos os pratos, porém com um ganho ligeiramente maior para os pratos superiores. Esta diferença de magnitude pode ser atribuída a propagação do efeito prato a prato pela fase vapor. O ganho é aproximadamente 2 vezes superior no último prato, quando comparado com a base da coluna. Novamente este comportamento foi observado por Marangoni (2005) ao avaliar a dinâmica do processo desta mesma unidade de destilação com dados experimentais, aplicando perturbações no calor do refeedor e no fornecimento de energia à resistência. Vale salientar que naquela ocasião, algumas condições operacionais e alguns sensores da planta eram diferentes, em relação à atual unidade.

#### 4.2.2.4 Avaliação da dinâmica em malha aberta após as perturbações aplicadas

Este estudo dinâmico foi de extrema importância para verificação das características do processo. Com a aplicação destas perturbações de magnitudes idênticas e conhecidas, verifica-se que os impactos são mais acentuados próximos ao prato onde ocorre a alimentação e se propagam com um fator de amortecimento para a seção de retificação, gerando uma pequena redução e aumento da fração de etanol no topo, para a perturbação degrau negativa e positiva, respectivamente. Isto mostra claramente que os pratos superiores são afetados pelas perturbações aplicadas com um atraso maior, decorrente do tempo necessário para sua propagação. Já o efeito causado na seção de esgotamento é maior nos dois casos, devido à propagação mais rápida na fase líquida. Estes comportamentos já eram esperados e mostram que o modelo em malha aberta é adequado e representativo.

Observa-se também que as temperaturas possuem grande influência na composição dos pratos. Pela análise de sensibilidade serão determinados os estágios mais sensíveis da coluna e que consequentemente afetam a composição de forma mais significativa.

Quando aplicado o degrau positivo na composição de etanol na alimentação, foi possível observar o efeito provocado pela perturbação nos dois extremos da coluna. Sendo que na base e no topo da coluna ocorrem reduções de aproximadamente 2,5°C e 1,5°C, respectivamente. Escolheu-se esta perturbação para fazer analogia aos resultados obtidos por Marangoni (2005), a qual aplicou uma perturbação degrau de +40% na fração volumétrica de uma coluna de destilação, operando com mistura etanol/água. Estes mesmos efeitos foram observados pela autora.

As aplicações das perturbações, tanto negativas quanto positivas, alteraram o modo de operação da unidade, fazendo com que ela migrasse para estados estacionários diferente dos anteriores, uma vez que as malhas de controle não estão atuando de modo a rejeitar as perturbações. O tempo para atingir esse novo ponto de operação, foi similar para ambos os casos, em torno de 1,5 horas. Este valor está coerente com o encontrado na literatura. Fieg (2002) obteve uma transição de 1,5 a 2,5 h para uma perturbação efetuada na composição de alimentação de uma coluna multicomponente e Marangoni (2005), operando a mesma unidade experimental empregada para validar as simulações deste trabalho, encontrou novos valores de composição cerca de 1,5 h após a perturbação.

Em relação ao comportamento do processo quando aplicada uma perturbação no fornecimento de energia ao prato 2, caracterizando o aquecimento distribuído, grande diferencial deste estudo, destaca-se que o efeito é mais pronunciado nos pratos superiores da coluna, em relação a base. Este fato auxilia no entendimento dos efeitos observados nos perfis de temperatura e composição de todos os pratos da coluna, causados pelas perturbações positivas e negativas aplicadas na composição de alimentação.

Os resultados apresentados definem o comportamento da coluna de destilação operando em malha aberta. Suas respostas são coerentes com comportamentos encontrados na literatura para situações similares. Desta forma, o modelo pode ser considerado adequado para gerar os dados necessários para desenvolvimento do sensor por software. Conforme constatado pelas análises gráficas, sua dinâmica é lenta, com presença de atrasos nas respostas e os efeitos observados estão intimamente ligados ao fornecimento de energia em um ponto intermediário da coluna. Estas características, associadas ao longo tempo necessário para a obtenção do novo estado estacionário, quando o processo é perturbado, consistem no fator motivador para a implementação do sensor por software usado em conjunto com abordagem distribuída, o que poderá nos fornecer respostas imediatas do comportamento do sistema, para tomada de decisões de forma mais ágil.

#### 4.2.2.5 Determinação do intervalo máximo entre as perturbações no simulador na etapa de obtenção dos dados de treinamento – Malha Aberta

Na Figura 4.32 é ilustrado o efeito causado quando alteradas as condições da corrente de alimentação do processo. Neste caso, perturbou-se simultaneamente as três variáveis desta corrente. Foram aplicados os valores dos extremos da faixa de operação, de tal forma que o impacto causado seja o maior possível em relação à variável de saída. Ou seja, na Figura 4.32a e Figura 4.32b tem-se representadas às condições que mais afetam positiva e negativamente a composição de topo da coluna, respectivamente. Verifica-se o triplo efeito das variáveis da corrente de alimentação: temperatura, vazão volumétrica e composição volumétrica de etanol sobre a composição de topo. Foi respeitada a faixa de operação que foi possível rodar o simulador HYSYS<sup>®</sup> em malha aberta, conforme detalhado no *Capítulo III, seção 3.3.1.5*, visto que ao extrapolar estes limites, o software não converge. A simbologia ilustrada nas legendas da Figura 4.32a,b ( $T_a$  [°C];  $Q_a$  [L.h<sup>-1</sup>];

$X_a$  [v/v]) expressam as perturbações aplicadas na corrente de alimentação do processo, partindo sempre do ponto de operação nominal ( $T_{a90}$ ;  $Q_{a300}$ ;  $X_{a0,10}$ ).

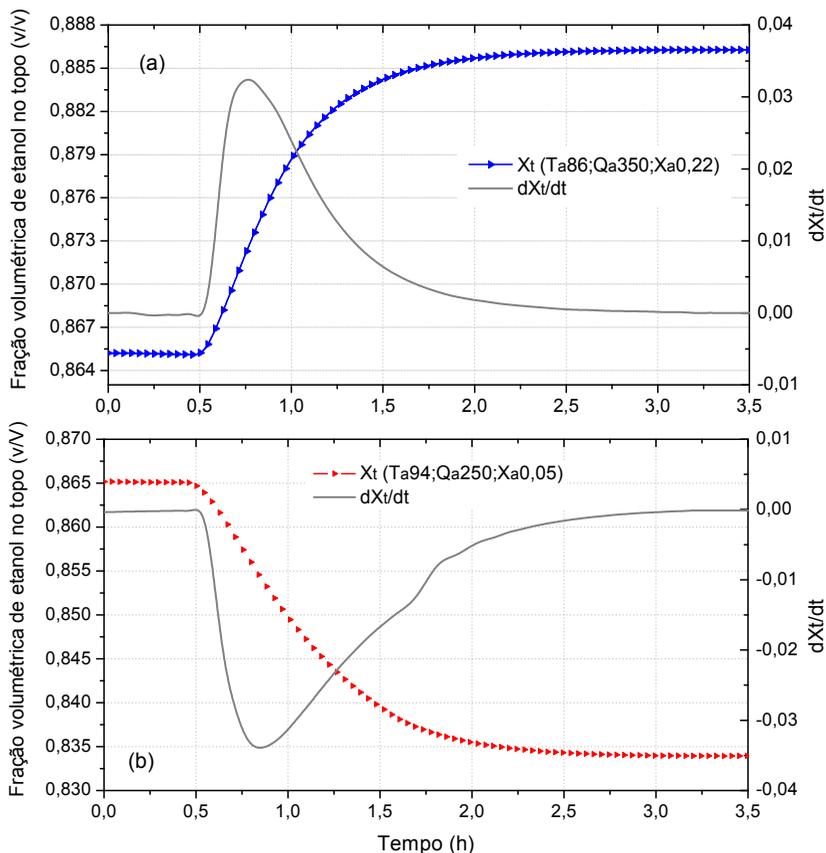


Figura 4.32. Perfil de fração volumétrica de etanol no topo da coluna com suas respectivas derivadas (a) fração máxima atingida; (b) fração mínima atingida.

Pela análise da Figura 4.32 e auxílio das derivadas das curvas, verifica-se que nas situações mais severas que o sistema é submetido ele demora em torno de 3,5 h para atingir o novo estado estacionário, após aplicadas as perturbações nas variáveis da alimentação. Desta forma, este valor foi o intervalo de tempo máximo entre as aplicações das

perturbações. Para perturbações mais leves o tempo foi proporcionalmente abaixo deste intervalo.

Este conhecimento do comportamento dinâmico da coluna irá nos auxiliar na determinação dos parâmetros do processo e tempo de amostragem para aquisição dos dados que, posteriormente, serão empregados no treinamento da rede neural e desenvolvimento do sensor virtual.

#### 4.2.2.6 Determinação do tempo de amostragem

Deve-se certificar que o conjunto de dados que serão usados para treinamento do sensor contenham dados representativos do comportamento dos perfis das curvas das variáveis selecionadas como entradas e saídas do sensor, com quantidades suficientes de dados para gerar um padrão de treinamento, além de abranger toda a faixa que se deseje que o sensor opere.

Um dos fatores de extrema relevância, importante para contemplar as situações mencionadas anteriormente, é o tempo de amostragem durante a aquisição dos dados simulados. Este deve ser suficientemente pequeno, capaz de acompanhar a evolução das variáveis ao longo do tempo sem perder nenhum efeito dinâmico, e ao mesmo tempo gerar menor volume de dados possíveis, de forma a não obter dados redundantes para o sensor.

Pela análise feita anteriormente, verificou-se que a composição de saída possui uma dinâmica com resposta lenta e com atraso, devido a propagação da perturbação por todos os estágios da coluna. Desta forma, a variável de saída, composição volumétrica de etanol no topo ( $X_t$ ), não é a melhor opção para determinação do tempo de amostragem. Pela análise do comportamento dinâmico do processo nas Figuras 4.25 e 4.28, verifica-se que a temperatura  $T_4$  (prato da alimentação) é a que possui dinâmica mais rápida, sendo o prato que recebe o efeito da perturbação diretamente. Por este motivo, este prato apresenta a menor constante de tempo, dentre as variáveis avaliadas, e sua temperatura será usada para determinação do tempo de amostragem.

Foram aplicadas perturbações de magnitudes -5% a +5% nas variáveis manipuladas  $X_a$ ,  $T_a$ ,  $Q_a$  e verificada a resposta de saída da variável de interesse, temperatura do quarto prato ( $T_4$ ), conforme Figura 4.33. As perturbações negativas e positivas nas variáveis foram escolhidas de modo que a temperatura do quarto prato fosse afetada positivamente e, a magnitude, de modo que não fosse extrapolada a faixa de operação. Das três variáveis manipuladas, verifica-se que a

temperatura e a composição de alimentação são as variáveis que afetam mais rapidamente a temperatura do prato 4, em comparação com a vazão da alimentação. Observa-se também que ambas possuem constantes de tempo muito similares. Foi escolhida a curva da temperatura da alimentação para efetuarmos o cálculo aproximado da menor constante de tempo do processo, por facilitar a aplicação da metodologia empregada para identificação.

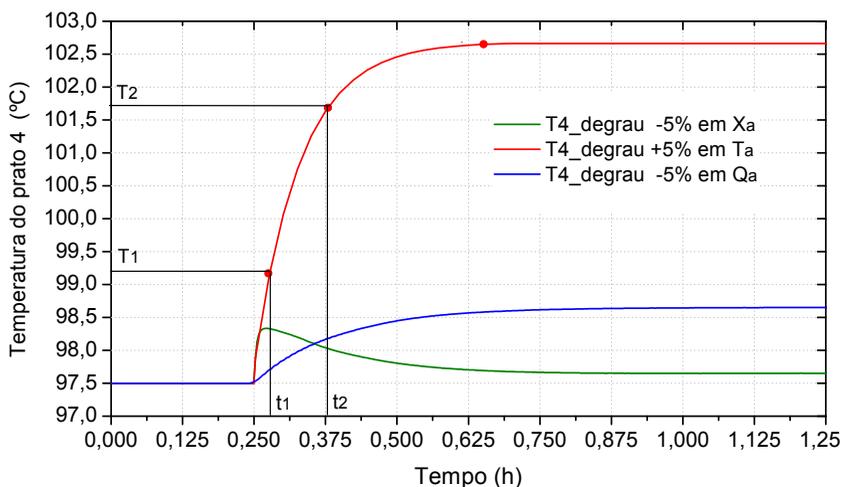


Figura 4.33. Resposta da temperatura do quarto prato da coluna após aplicação de perturbações degraus na composição, vazão e temperatura da alimentação.

Empregando-se o método de identificação proposto por Sundaresan e Krishnaswamy (1977) que utiliza dois tempos ( $t_1$  e  $t_2$ ) estimados a partir da curva de resposta à perturbação, correspondendo a 35,3% e 85,3% do tempo de resposta, respectivamente, a constante de tempo pode ser estimada pela Equação 4.1:

$$\tau = 0.67(t_2 - t_1) \quad (\text{Eq. 4.1})$$

Para encontrarmos os tempos  $t_1$  e  $t_2$ , deve-se primeiramente obter as temperaturas  $T_1$  e  $T_2$ , respectivas. O resultado da amplitude da variação da temperatura do prato 4 ( $\Delta T_4 = 5,1^\circ\text{C}$ ) é multiplicado por 35,3% e 85,3% do tempo de resposta. Logo temos (Equação 4.2) e (Equação 4.3):

$$T_1 = 0,353.5,1 = 1,80^\circ C + 97,50^\circ C = 99,30^\circ C \quad (\text{Eq. 4.2})$$

$$T_2 = 0,853.5 = 4,35^\circ C + 97,5^\circ C = 101,85^\circ C \quad (\text{Eq. 4.3})$$

De posse desses valores calculados, com o auxílio do gráfico, projeta-se uma reta na curva de resposta obtida e encontram-se os valores respectivos a esses tempos:

$$t_1 = 0,28 \text{ h} \quad e \quad t_2 = 0,375 \text{ h}$$

A constante de tempo pode então ser estimada pela equação Equação 5.1:

$$\tau = 0,67(t_2 - t_1) = 0,67(0,375 - 0,28)$$

$$\tau = 0,0636 \text{ h}$$

Desta forma, a constante de tempo para composição de saída foi estimada em 0,0636 h, equivalente a aproximadamente 230 s. Para definição do tempo de amostragem, é usual empregar  $\tau/10$  ou  $\tau/20$  para garantir que nenhum efeito seja desprezado. A segunda opção foi escolhida. Desta forma, o tempo de amostragem que seria 11,5 s foi arredondado e definido como sendo 10 s.

#### 4.2.2.7 Análise de sensibilidade

De acordo com Secchi *et al.*, (2006) a análise de sensibilidade fornece informações sobre as respostas do processo frente a modificações em variáveis manipuladas.

Para a identificação dos estágios mais sensíveis e, conseqüentemente, escolha das temperaturas dos mesmos para serem usadas como entradas da rede neural, utilizaram-se os estudos realizados por Marangoni (2005). Como a autora realizou sua pesquisa na mesma unidade experimental em questão, entende-se que seus dados são válidos para este estudo e serão utilizados como base, feitas as devidas ressalvas. Assim, a definição da escolha dos pratos mais sensíveis e uso de suas respectivas temperaturas como entrada para rede foi realizada a partir desta análise, associadas às características da planta. Em seu trabalho, Marangoni (2005) aplicou os métodos de análise de sensibilidade sugeridos por Kalid (1999). O objetivo foi determinar qual o prato mais sensível para aplicação do controle distribuído. Foram

utilizados três dos cinco métodos sugeridos pelo autor: pratos sucessivos, simetria de sensibilidade e máxima sensibilidade.

No primeiro método, as diferenças entre as temperaturas de dois pratos sucessivos foram calculados ao longo da coluna e o prato mais sensível foi o que apresentou a maior diferença em relação aos pratos adjacentes. No segundo método, um perfil de temperatura foi obtido para um determinado valor da variável manipulada (neste caso, vazão de refluxo e o calor do refeedor). O prato mais sensível fornece uma resposta simétrica para variações positivas e negativas iguais. Finalmente, o terceiro método analisa o prato com o maior valor da derivada da temperatura ou variável modificada em relação aos estágios da coluna para a variação positiva e negativa quando o processo é perturbado.

Os resultados obtidos com o primeiro método (pratos sucessivos) usando três composições distintas de etanol na corrente de alimentação demonstraram a possibilidade de utilizar os pratos 1, 2, 3, 5 e 7. Como o quinto e sétimo pratos estão localizados na seção de retificação, foram descartados pela autora. Baseado em estudos diabáticos, na qual a proposta foi fundamentada, esta seção é mais indicada para remover calor do que para adicioná-lo. O prato 1 foi descartado pois este prato recebe a corrente de vapor do refeedor. Para definição do estágio mais sensível, uma vez que o método 1 não foi conclusivo, o método de simetria de sensibilidade e o máxima sensibilidade (métodos 2 e 3) foram utilizados. O método de máxima sensibilidade indicou novamente os pratos 5 e 7, mas o método de simetria de sensibilidade indicou os pratos 2 e 3 como os mais sensíveis da coluna. Desta forma, a autora utilizou os pratos 2 e 3 para *emprego do controle distribuído*, que era seu objetivo. Maiores detalhes dos métodos podem ser obtidos em Marangoni (2005).

Nesta etapa do trabalho, um dos objetivos é verificar os pratos mais sensíveis, indiferente de estarem localizados na seção de retificação ou esgotamento. Desta forma, algumas considerações para as escolhas ou descartes de alguns pratos feitos pela autora, não são aplicadas neste estudo. Ou seja, levaremos em consideração os pratos 1, 5 e 7 dos métodos de pratos sucessivos e de máxima sensibilidade; e os pratos 2 e 3 do método de simetria de sensibilidade. Desta forma, as temperaturas destes podem ser utilizadas como entradas para a rede neural artificial.

### 4.3 MODELO EMPÍRICO DINÂMICO BASEADO EM REDES NEURAIS – MALHA ABERTA

Nesta etapa será apresentado um modelo empírico não-linear baseado em redes neurais desenvolvido para a coluna de destilação, cuja finalidade é ser empregado como um sensor por software, fornecendo informação para implementação futura de técnicas de controle multivariáveis avançadas. Para isso, é necessário o conhecimento da robustez do modelo e de suas limitações, de forma a não comprometer a qualidade das observações.

O potencial de aplicação de redes neurais em processos industriais é vasto. A habilidade das redes neurais de capturar e modelar processos severamente não-lineares, dinâmicos ou ainda acoplados faz delas ferramentas poderosas em controle baseado em modelos e monitoração.

#### 4.3.1 Determinação da função de transferência mais adaptada para o conjunto de dados para redes neurais Multi-Layer Perceptron (MPL)

Inicialmente foram testadas 107 redes neurais artificiais (RNA's) *Feedforward* do tipo *Multi-Layer Perceptron* – MPL, empregando o algoritmo de treinamento *Levenberg-Marquardt* (LM), que é uma variação do método de Newton, o qual foi projetado para minimizar funções que são somas quadráticas de outras funções não-lineares (HAGAN *et al.*, 1994). Este algoritmo de treinamento é considerado rápido e eficaz, para redes de tamanho moderado, e sua aplicação é adequada para treinamento de redes neurais onde o índice de desempenho é o erro quadrático médio.

O objetivo principal desta etapa é determinar a função de transferência que obtém o melhor desempenho com o conjunto de dados simulados em malha aberta, empregando o algoritmo *Levenberg-Marquardt* (LM). Tendo em vista que a qualidade da inferência é dada pela capacidade superior de predição de determinada rede neural para a propriedade chave, torna-se imprescindível a comparação de diferentes inferências visando à determinação daquela de maior precisão. Justamente por este motivo, várias redes neurais foram avaliadas.

Foram testadas várias arquiteturas de redes, com distintas funções de transferência, variando-se o número de camadas intermediárias e o número de neurônios em cada camada, conforme descrito no *Capítulo III*, seções 3.3.2.3.3 e 3.3.2.3.4. Conforme citado

anteriormente, não existem regras específicas para encontrar a arquitetura de rede que fornecerá o melhor resultado, decorrente do grande número de problemas específicos que podem ocorrer durante a etapa de treinamento, relativos ao conjunto de dados disponíveis.

Na etapa do treinamento a rede foi alimentada com 78.833 pontos, formando uma matriz de entrada  $X_{tre}$   $\langle 78.833 \times 8 \rangle$  e uma matriz de saída  $Y_{tre}$   $\langle 78.833 \times 1 \rangle$ . Já os dados obtidos com o emprego do simulador, para o conjunto de dados de validação (em torno de 25% do tamanho do conjunto de dados do treinamento) formaram uma matriz de entrada  $X_{val}$   $\langle 18.610 \times 8 \rangle$  e de saída  $Y_{val}$   $\langle 18.610 \times 1 \rangle$ . A Figura 4.34 representa os histogramas mostrando a distribuição dos conjuntos de dados usados para treinamento e para validação cruzada.

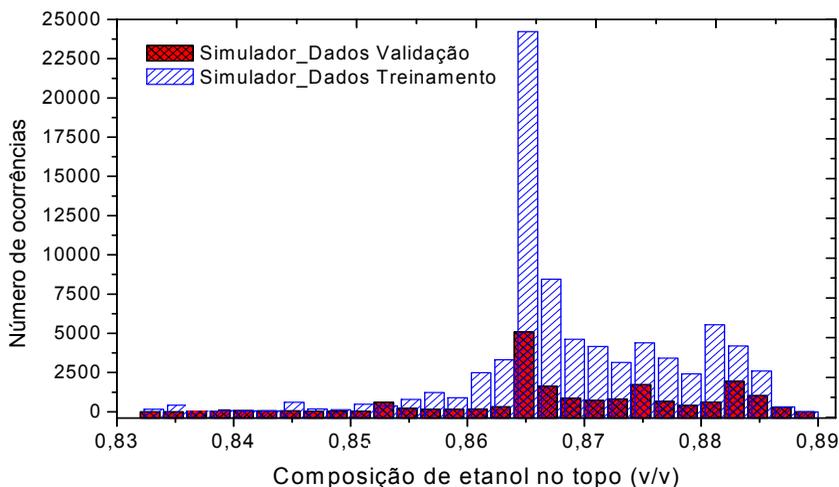


Figura 4.34. Histogramas da composição de etanol no topo da coluna. Dados gerados pelo simulador, etapa do treinamento e validação.

A grande quantidade de dados de composição com valor 0,865 se justifica pois equivale ao valor desejado do produto de topo em estado estacionário, no ponto de operação nominal da coluna. Após cada perturbação, o processo atinge um novo estado estacionário e permanece nele. Porém, logo em seguida, retira-se a perturbação e o processo retorna ao ponto de operação. Este procedimento foi repetido para todas as perturbações aplicadas. Observa-se ainda que os dados de validação contemplaram toda a faixa na qual as redes foram treinadas, o que confere maior confiabilidade aos resultados obtidos na etapa da validação.

### 4.3.1.1 Medidas de desempenho das redes MPL avaliadas – Etapa do treinamento

Os resultados obtidos na etapa de treinamento das redes neurais MPL empregando o algoritmo de treinamento *Levenberg-Marquardt*, com redes variando de 1 a 3 camadas intermediárias com 3 a 30 neurônios em cada camada, são ilustrados nas Figuras 4.35, 4.36 e 4.37. Nestas figuras verifica-se o desempenho das arquiteturas testadas com diferentes funções de transferências aplicadas nas camadas intermediárias e de saída. Foram avaliadas combinações entre as funções *tangente hiperbólica*, *logarítmica* e *linear*, sendo que a função linear foi aplicada apenas na camada de saída das redes, usando-se variantes com padronizações distintas: *purelin* (-1 a +1) e *poslin* (0 a +1). O desempenho destas redes foi avaliado pela análise do erro médio quadrático (MSE).

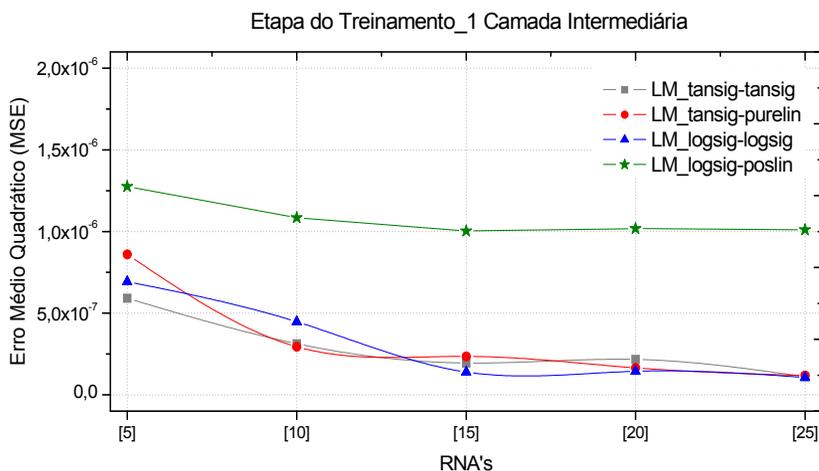


Figura 4.35. Comparação do desempenho das funções de transferência para redes com uma camada intermediária e diferentes números de neurônios, na etapa do treinamento.

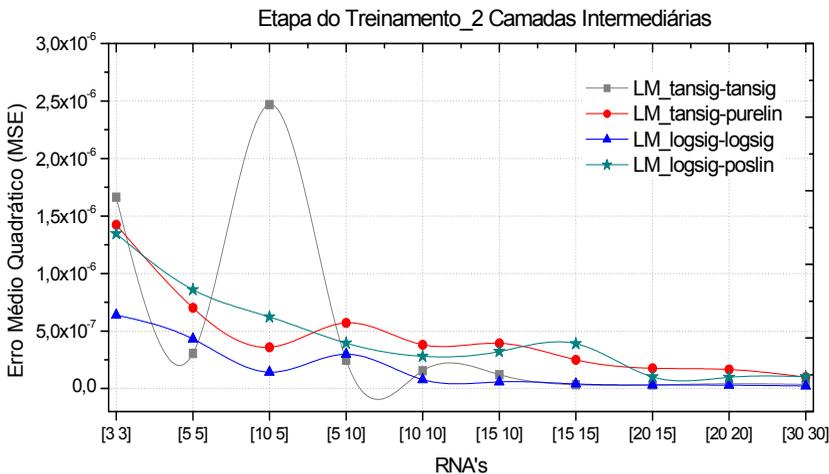


Figura 4.36. Comparação do desempenho das funções de transferência para redes com duas camadas intermediárias e diferentes números de neurônios, na etapa do treinamento.

Para a RNA de 2 camadas com a função logarítmica, observa-se que a partir de 10 neurônios nas camadas intermediárias não houve um melhora significativa do desempenho da rede subsequentes.

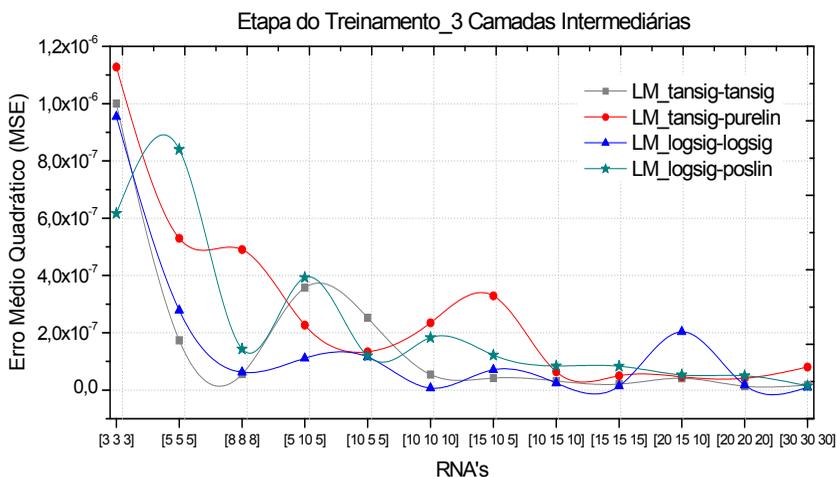


Figura 4.37. Comparação do desempenho das funções de transferência para redes com três camadas intermediárias e diferentes números de neurônios, na etapa do treinamento.

Percebe-se que as redes neurais artificiais (RNA's) de 3 camadas, com a função logarítmica aplicada nas camadas intermediárias, se mantiveram em uma região de erro mínimo para maioria das redes testadas acima de 5 neurônios em cada camada. Esta função foi a que obteve melhores resultados comparativamente com as demais. Observa-se ainda que a função logarítmica apresenta um comportamento oscilatório, com erros maiores para redes com números distintos de neurônios nas camadas ocultas e menores erros quando empregado o mesmo número de neurônios nas camadas ocultas, com destaque para rede de 10 neurônios em cada camada, a qual atingiu um erro mínimo com um número moderado de neurônios. Este fato pode ser um indicativo desta rede ter boa capacidade de generalização. Geralmente, quanto maior a quantidade de neurônios nas camadas intermediárias, maior a necessidade de ajustes de pesos e parâmetros, o que pode auxiliar na etapa do treinamento, porém levar a resultados indesejáveis na etapa de validação, devido à grande quantidade de graus de liberdade e/ou problemas de sobre-treinamento.

Foram selecionadas as 4 melhores redes de cada uma das 3 arquiteturas testadas: com uma, duas e três camadas intermediárias, todas utilizando o algoritmo de treinamento *Levenberg-Marquardt (LM)*. As 12 redes que apresentaram melhor desempenho na etapa do treinamento, com seus respectivos erros médios quadráticos, além de informações sobre a arquitetura neural e o número de interações realizadas durante o treinamento, são descritas na Tabela 4.1.

Na Figura 4.38 é ilustrada a predição das 3 redes neurais (em destaque na Tabela 4.1) que obtiveram o menor erro (MSE) durante o treinamento e comparada com a saída das simulações do software HYSYS®. Das redes avaliadas, as redes com três camadas intermediárias se destacaram, usando funções de transferência logarítmica nas três camadas e também logarítmica e linear na camada de saída, apresentando erros com grandeza em torno de  $10^{-9}$ .

Pela análise da Figura 4.38, observa-se sobreposição dos valores fornecidos pelo simulador e pelas redes (preditos) para composição de produto de topo da coluna. Pela análise deste resultado verifica-se que, durante a etapa de treinamento, a rede foi capaz de criar padrões de aprendizado, predizendo os valores de composição de topo da coluna de forma eficaz.

Tabela 4.1. Seleção das 12 redes com melhor desempenho na etapa do treinamento empregando o algoritmo de treinamento *Levenberg-Marquard*, com distintas funções de transferência.

<i>Nº de camadas intermediárias</i>	<i>Nº de neurônios</i>	<i>Funções de transferência (intermed-saída)</i>	<i>Padrão-zação</i>	<i>Nº de épocas</i>	<i>*MSE_tre</i>	<i>**MSE_val</i>	<i>**MAPE_val</i>
1	[15]	<i>logsig-logsig</i>	(+0,1; +0,9)	169	1,3959E-7	4,2268E-7	0,04674
1	[25]	<i>logsig-logsig</i>	(+0,1; +0,9)	330	1,0642E-7	1,6420E-6	0,07614
1	[25]	<i>tansig-purelin</i>	(-0,9; +0,9)	402	1,1867E-7	5,9865E-7	0,04595
1	[25]	<i>tansig-tansig</i>	(-0,9; +0,9)	671	1,1052E-7	7,8410E-7	0,05860
2	[10 10]	<i>logsig-logsig</i>	(+0,1; +0,9)	595	7,8084E-8	3,3901E-7	0,03567
2	[15 15]	<i>tansig-tansig</i>	(-0,9; +0,9)	376	3,3663E-8	5,9947E-7	0,04710
2	[20 15]	<i>logsig-logsig</i>	(+0,1; +0,9)	302	3,2122E-8	7,1564E-7	0,05597
2	[30 30]	<i>tansig-tansig</i>	(-0,9; +0,9)	603	3,6093E-8	2,5176E-6	0,09273
3	[8 8 8]	<i>logsig-logsig</i>	(+0,1; +0,9)	350	6,2415E-8	3,8030E-7	0,03785
3	[10 10 10]	<i>logsig-logsig</i>	(+0,1; +0,9)	337	7,0741E-9	2,3724E-7	0,03117
3	[15 15 15]	<i>logsig-logsig</i>	(+0,1; +0,9)	743	1,4411E-8	2,2589E-6	0,08660
3	[30 30 30]	<i>logsig-poslin</i>	(+0,1; +0,9)	694	1,4692E-8	3,5002E-6	0,10728

\**MSE\_tre* = Erro médio quadrático na etapa do treinamento.

\*\**MSE\_val* e *MAPE\_val* = Erro médio quadrático e erro percentual médio absoluto na etapa da validação.

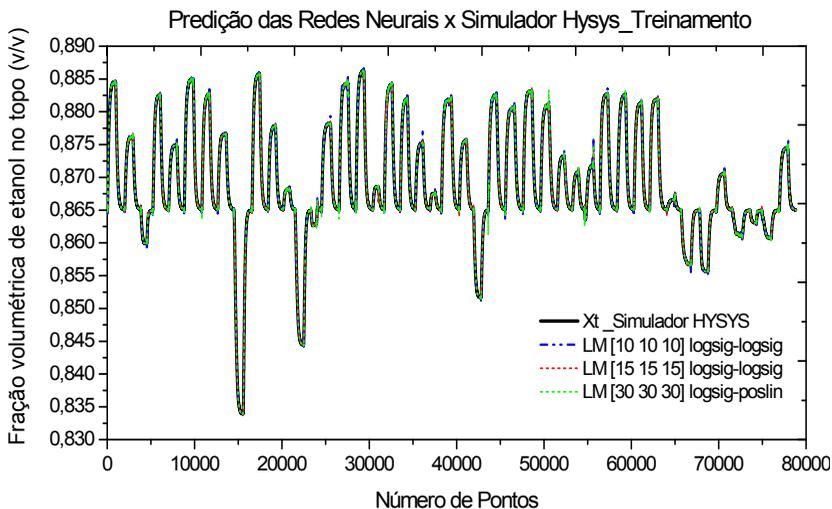


Figura 4.38. Comparação entre a saída simulada e a predição das melhores redes testadas, com o algoritmo *Levenberg-Marquardt* (LM), na etapa do treinamento.

Na Figura 4.39 observa-se uma ampliação de uma seção dos resultados da saída do simulador *versus* a predição das redes

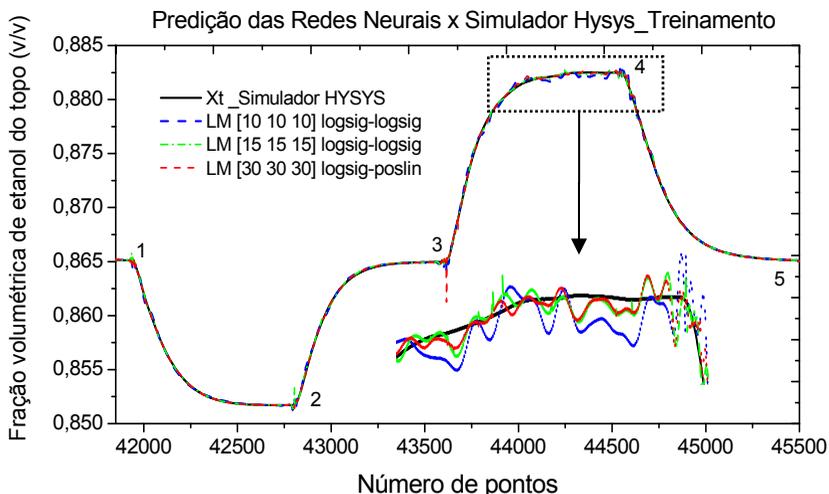


Figura 4.39. Ampliação de uma seção da predição das 3 melhores redes testadas comparadas com saída simulada, com o algoritmo *Levenberg-Marquardt* (LM), na etapa do treinamento.

Os números antes de cada inflexão das curvas, observados na Figura 4.39 simbolizam as perturbações aplicadas e retiradas na etapa da geração dos dados com o simulador. Seus respectivos estados estacionários seguem descritos: (1) Ponto de operação nominal - Estado Estacionário EE1 ( $T_a 90; Q_a 300; X_a 0,10$ ) no qual foi aplicado um degrau negativo na fração volumétrica de etanol da alimentação de 0,10 para 0,05. (2) Estado Estacionário EE2 ( $T_a 90; Q_a 300; X_a 0,05$ ). Após atingida a estabilidade, foi retirada a perturbação e o sistema retorna ao ponto de operação. (3) Estado Estacionário EE1, aplicado degrau positivo novamente na fração volumétrica de etanol de 0,10 para 0,20. (4) Estado Estacionário EE3 ( $T_a 90; Q_a 300; X_a 0,20$ ) após atingida a estabilidade retirou-se a perturbação. (5) Estado Estacionário EE1.

Onde:  $T_a$  – temperatura da alimentação [ $^{\circ}\text{C}$ ];  
 $Q_a$  – vazão da alimentação [ $\text{L}\cdot\text{h}^{-1}$ ];  
 $X_a$  – composição da alimentação [ $\text{v/v}$ ].

#### 4.3.1.2 Medidas de desempenho das redes MPL avaliadas – Etapa da validação cruzada

Na etapa da validação cruzada, objetiva-se encontrar redes que apresentam melhor capacidade de generalização, ao serem alimentadas com dados distintos dos quais elas foram treinadas (porém sem extrapolar os limites do treinamento) e que possuam capacidade de inferir a composição do produto de topo com o menor erro possível. Para verificar o desempenho das redes treinadas, elaborou-se o gráfico de barras para melhor visualização dos erros (Figura 4.40). Como critério de comparação foi usado o erro médio quadrático (MSE) e para auxiliar na análise, o erro percentual médio absoluto (MAPE) também foi empregado, conforme Tabela 4.1. Foram avaliadas as 12 redes pré-selecionadas anteriormente.

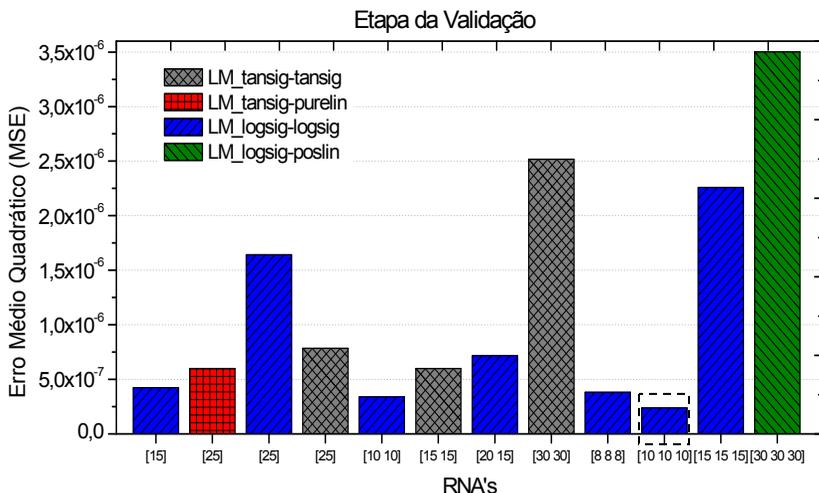


Figura 4.40. Comparação do desempenho das melhores redes, com o algoritmo *Levenberg-Marquardt*, na etapa validação cruzada.

Foi verificado na etapa do treinamento que, à medida que se aumenta o número de neurônios das camadas intermediárias, o erro diminui. Porém, o erro estabiliza em um patamar, no qual não se justifica mais o incremento de mais neurônios.

Além disso, pela análise da Figura 4.40, verifica-se que na etapa da validação, à medida que se aumenta o número neurônios nas camadas intermediárias, o erro médio quadrático aumenta, mesmo para redes que obtiveram um bom resultado na etapa do treinamento. Este fato comprova que redes com arquiteturas muito robustas (grandes quantidades de neurônios) se adaptam bem aos dados de treinamento, em virtude do aumento de parâmetros ajustáveis, porém não possuem grande capacidade de generalização, ao serem apresentados dados novos a rede.

Os resultados das três redes que obtiveram o melhor e o pior desempenho na validação são descritos na Tabela 4.2 e ilustrados na Figura 4.41. A apresentação dos resultados das 3 piores redes será feito para demonstrar que, mesmo as redes com pior desempenho, obtiveram predições razoáveis. Este resultado foi obtido através da validação cruzada, onde foram apresentados à rede um novo conjunto de dados com condições distintas das que ela foi treinada, conforme detalhado no *Capítulo III, seção 3.3.2.3.5 Tabela 3.7*.

Tabela 4.2. Seleção das 3 redes com melhor e pior desempenho na etapa da validação.

	Algoritmo	Nº de camadas	Nº de neurônios	Funções de transferência	*MSE_val	*MAPE_val
Melhores	LM	1	[15]	logsig-logsig	4,2268E-7	0,0467
	LM	2	[10 10]	logsig-logsig	3,3901E-7	0,0356
	LM	3	[10 10 10]	logsig-logsig	2,3724E-7	0,0311
Piores	LM	1	[25]	logsig-logsig	1,6420E-6	0,0761
	LM	2	[30 30]	tansig-tansig	2,5176E-6	0,1027
	LM	3	[30 30 30]	logsig-poslin	4,6702E-6	0,1172

\*MSE\_val e MAPE\_val = Erro médio quadrático e erro percentual médio absoluto na etapa da validação.

Pela análise dos dados contidos na Tabela 4.2 e na Figura 4.41, verifica-se que a rede neural de 3 camadas, composta de 10 neurônios em cada camada, conseguiu uma predição dos valores de composição de produto de topo mais próxima da resposta do simulador, apresentado os menores erros, tanto para o erro médio quadrático (MSE), quanto para o erro percentual médio absoluto (MAPE).

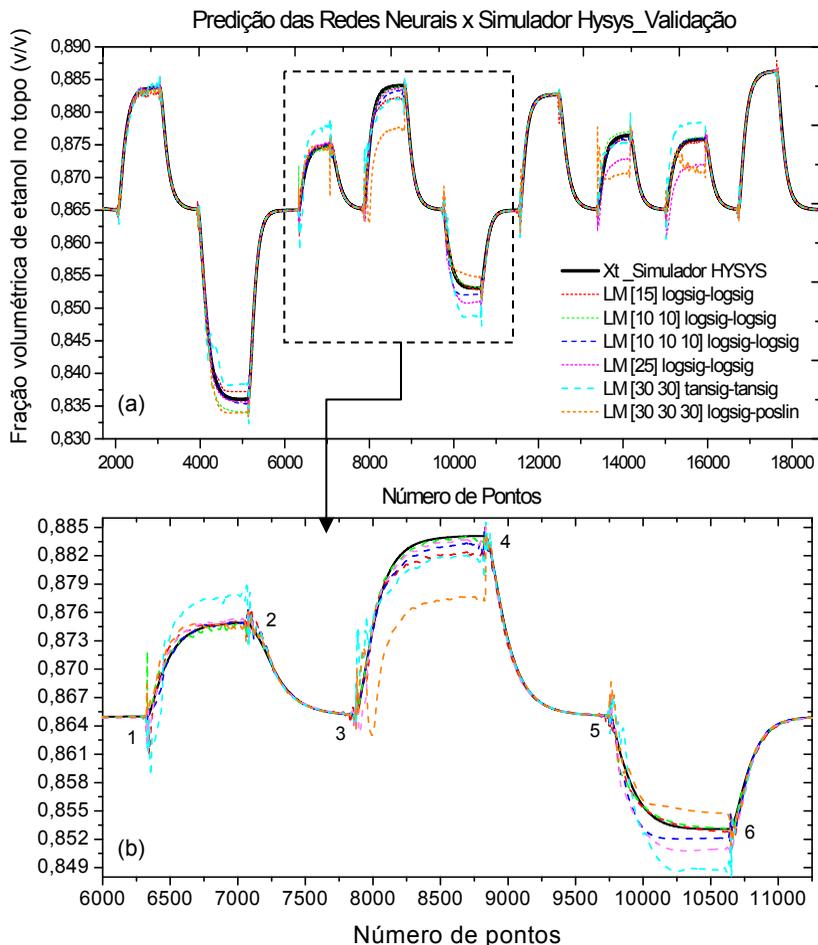


Figura 4.41. (a) Comparação entre a saída simulada e a predição das 3 melhores e piores redes testadas, com o algoritmo *Levenberg-Marquardt* (LM), na etapa da validação; (b) ampliação de uma seção da etapa de validação, com descrição detalhada das perturbações aplicadas.

Segue a descrição codificada de cada estado estacionário atingido após aplicação da perturbação: (1) EE1 ( $T_{a90}; Q_{a300}; X_{a0}, 10$ ). (2) EE2 ( $T_{a90}; Q_{a330}; X_{a0}, 17$ ). (3) EE1. (4) EE3 ( $T_{a88}; Q_{a240}; X_{a0}, 21$ ). (5) EE1. (6) EE4 ( $T_{a92}; Q_{a345}; X_{a0}, 06$ ). (7) EE1.

Finalizada esta etapa, conclui-se que, quando utilizado o algoritmo de treinamento *Levenberg-Marquardt* – *LM* para o conjunto de dados em que a rede foi treinada, a função *logarítmica* aplicada nas camadas intermediárias e na camada de saída foi a que mais se adaptou aos dados e que obteve o melhor desempenho comparativamente com as combinações das funções *tangente hiperbólica* e *linear*. Desta forma, na etapa seguinte, onde se pretende avaliar outros algoritmos de treinamento, a função *logarítmica* será aplicada novamente nas camadas intermediárias e de saída. Porém, sabe-se que com a aplicação desta função de transferência nos demais algoritmos de treinamento, estes não necessariamente fornecerão o melhor desempenho, fato que só poderia ser determinado e afirmado através de um estudo experimental completo e exaustivo. Foram realizados alguns testes preliminares com novos algoritmos de treinamento, confrontando outras funções com a *logarítmica*, e os resultados preliminares indicaram o melhor desempenho novamente desta função.

Na etapa seguinte serão avaliados outros algoritmos de treinamento (descritos com detalhes no *Capítulo III, seção 3.3.2.3.2*), nos quais será aplicada a mesma função de transferência (*logarítmica*) para auxiliar na comparação dos resultados obtidos. O objetivo final é obter uma rede com boa predição de composição de topo, para ser aplicada como *soft sensor* na unidade piloto de destilação, e sua predição poder ser empregada para desenvolvimento de sistemas de controles e otimização futura da unidade piloto.

#### 4.3.2 Determinação do algoritmo de treinamento mais adaptado para o conjunto de dados para redes neurais Multi-Layer Perceptron (MPL)

Para redes do tipo *backpropagation* vários tipos de algoritmos de treinamentos são disponibilizados e diferem entre si pela estratégia de otimização empregada ao realizar ajustes dos pesos. Cada um possui requisitos diferenciados em termos de memória computacional exigida para processamento dos cálculos e tempo de convergência.

Os algoritmos de treinamento que serão avaliados neste trabalho são: *Taxa de Aprendizado Adaptativa* (GDX); *Resilient Backpropagation* (RP); *Gradiente Conjugado em Escala* (SCG); *Levenberg-Marquardt* (LM); *Regularização de Bayesian* (BR) e *One Step Secant Method* (OSS). No *Capítulo II*, relativo à revisão bibliográfica, e no *Capítulo III*, onde é descrito a metodologia

empregada neste trabalho, são encontrados maiores detalhes dos algoritmos e o motivo que nortearam suas escolhas, bem como os métodos de otimização empregados por cada um deles.

O estudo com as avaliações do comportamento de distintos algoritmos de treinamento foi realizado objetivando, além de encontrar o algoritmo que mais se adapte e forneça as melhores respostas ao conjunto de dados, contribuir cientificamente para as pesquisas na área. São poucas pesquisas que relatam em quais situações cada algoritmo é recomendado, e que apresentam um comparativo do desempenho entre os algoritmos de treinamento empregando diferentes métodos.

Para realizar estes testes e verificar o desempenho dos algoritmos de treinamento foram avaliadas um total de 135 redes neurais do tipo MPL, sendo 25 com uma camada, 50 com duas camadas e 60 redes com três camadas intermediárias, todas com número de neurônios variando de 3 a 30 em cada camada.

#### 4.3.2.1 Medidas de desempenho das redes MPL avaliadas – Etapa do Treinamento

Na Figura 4.42 e Figura 4.43, são apresentados comparativos do comportamento do erro médio quadrático (MSE) das redes avaliadas com os algoritmos de treinamento analisados, com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias.

Pela análise dos resultados da Figuras 4.42 e 4.43 verifica-se claramente que os algoritmos de treinamento *Levenberg-Marquardt* (LM) e *Regularização de Bayesian* (BR) que empregam o *Método quasi-Newton* se destacam dentre os algoritmos avaliados, resultando em treinamentos com erros médios quadráticos muito inferiores aos demais. Ao comparar o comportamento destes dois algoritmos nas diferentes arquiteturas com 1, 2 e 3 camadas intermediárias, verifica-se que existe uma tendência de redução gradativa do erro de treinamento com aumento do número de camadas e de neurônios em cada camada. Para redes com 1 camada intermediária percebe-se que não existe ganho significativo de desempenho com mais de 15 neurônios. Já para redes com 2 e 3 camadas intermediárias, o mesmo é observado a partir de redes com mais de 10 neurônios em cada camada. De forma geral, concluiu-se que o algoritmo BR obteve respostas que apresentam o menor erro médio quadrático, dentre todos os algoritmos avaliados.

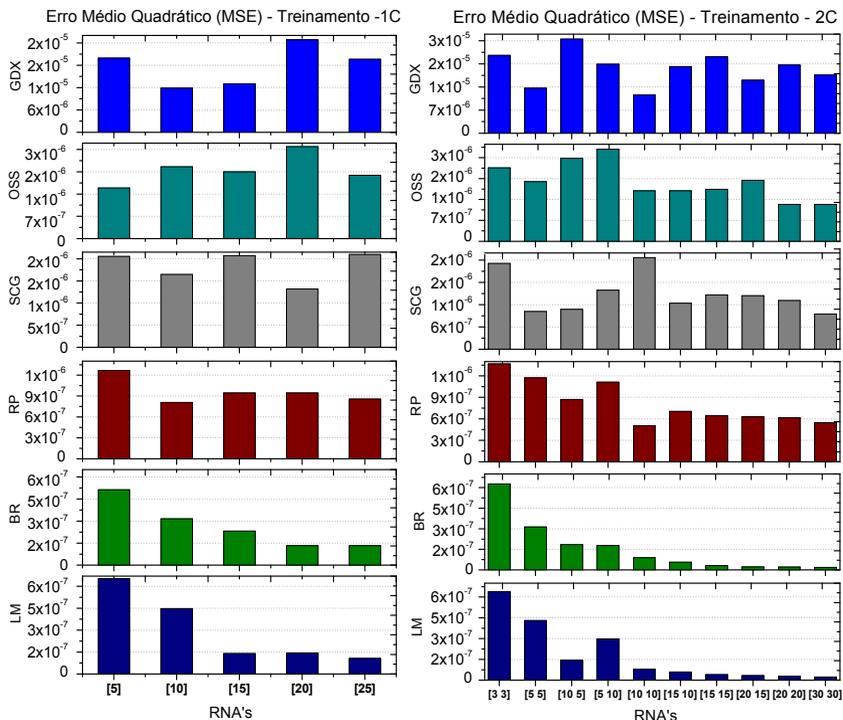


Figura 4.42. Comparativo do comportamento do erro médio quadrático (MSE) dos algoritmos de treinamento analisados individualmente, com o incremento do número de neurônios para redes com 1 e 2 camadas intermediárias.

Muitos pesquisadores (LU *et al.*, 2004; MOREJON e PRINCIPE, 2004; KERMANI *et al.*, 2005; SINGH *et al.*, 2007; MUKHERJEE e SRIKANTA, 2012) desenvolveram estimadores envolvendo redes neurais baseados no algoritmo *Levenberg-Marquardt* (LM). Esse fato pode ser atribuído a sua característica de realizar treinamentos rápidos, porém seu uso é aconselhável para redes de tamanho moderado. Possui a função de redução de memória que pode ser utilizada quando possuir muitos dados para o treinamento.

O método LM é um algoritmo tipo *quasi-Newton* anterior às redes neurais, que utiliza a informação das segundas derivadas em contraste com o gradiente descendente, que usa apenas a primeira. Por este motivo o algoritmo LM tende a ser mais eficiente que o gradiente descendente, fato constatado pela análise dos resultados. Contudo, há a necessidade de se armazenar uma aproximação da matriz Hessiana, que

contêm as segundas derivadas, o que resulta num uso significativamente maior da memória do computador, podendo comprometer sua eficiência no treinamento das redes demasiadamente grandes.

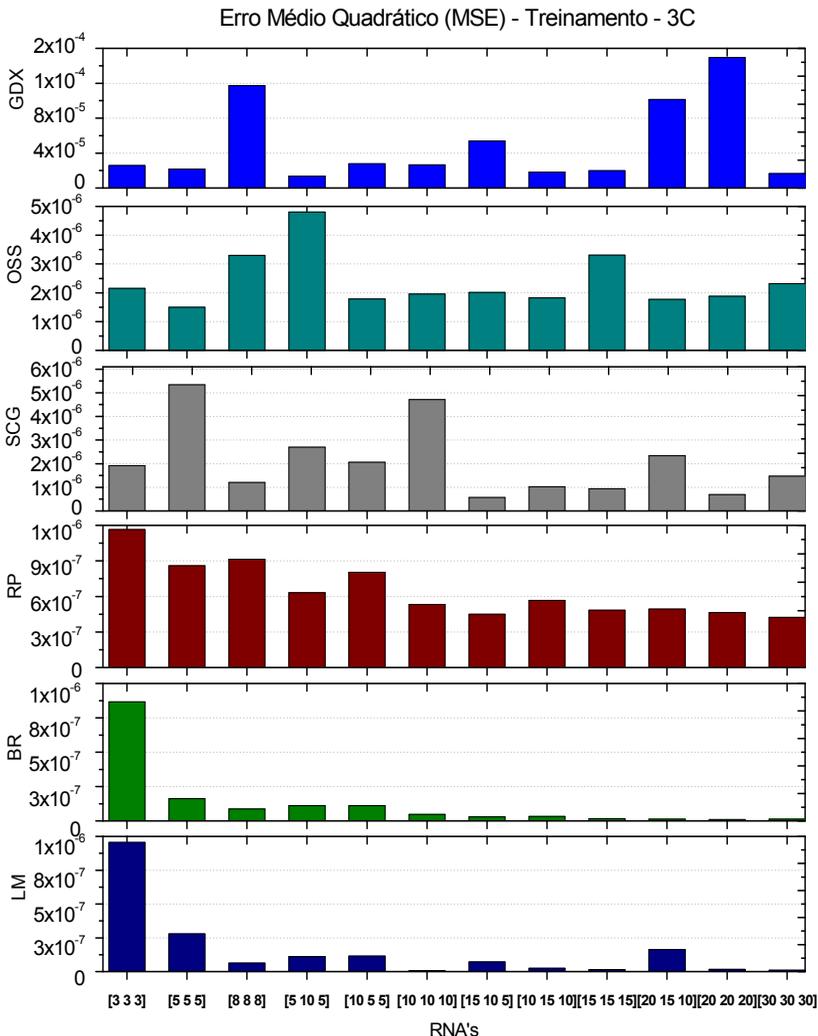


Figura 4.43. Comparativo do comportamento do erro médio quadrático (MSE) dos algoritmos de treinamento analisados individualmente, com o incremento do número de neurônios para redes com 3 camadas intermediárias.

Em relação ao algoritmo *Resilient Backpropagation* (RP), alguns autores, como Anastasiadis *et al.*, (2004) e Riedmiller e Braun (1993), o definem como sendo uma melhoria sobre a técnica de aprendizagem *backpropagation*. Neste, a alteração dos pesos sinápticos é baseada em informações de gradientes locais. Segundo Soares (2008), que empregou este algoritmo para predição de séries temporais econômicas, é importante citar que o ajuste da taxa de aprendizado e dos pesos sinápticos dependem dos termos do gradiente e não da sua magnitude, como no *backpropagation*. Com a utilização deste algoritmo o tamanho da atualização dos pesos depende apenas da seqüência de troca de sinais das derivadas parciais. Este fato proporciona que ele realize um grande número de interações. Por ser considerado um algoritmo simples com rápida convergência e pouca memória requisitada, obteve um bom desempenho (com erros na grandeza de  $10^{-7}$ ). Segue a tendência de redução do erro com aumento do número de camadas e neurônios, porém ainda apresenta erros médios quadráticos superiores aos algoritmos LM e BR.

Os piores resultados comparativamente ficaram com os algoritmos *Taxa de Aprendizado Adaptativa* (GDX), *One Step Secant Method* (OSS) e *Gradiente Conjugado em Escala* (SCG), os quais forneceram erros médios quadráticos com grandezas variando de  $10^{-4}$  a  $10^{-6}$ , sendo que não possuem uma tendência clara de melhora de desempenho com aumento ou redução de camadas e neurônios. Porém, observa-se que de forma geral, as redes de tamanho intermediário obtiveram erros menores.

O uso do algoritmo *Taxa de Aprendizado Adaptativa* (GDX) é aconselhável para ser empregado em treinamentos por batelada. Possui convergência mais rápida do que os outros algoritmos que empregam a técnica de gradiente descendente, como, por exemplo, o GD (*Gradiente descendente sem momento*) que apresenta resposta lenta e pode ser usado em treinamentos de modo incremental.

O método de *Gradiente Conjugado* é um método alternativo aos métodos de *quasi-Newton* e *Gradiente descendente*. Possui requisitos de memória similares ao método de *quasi-Newton*. Neste grupo, pode-se citar o algoritmo SCG – *Gradiente Conjugado em Escala*, utilizado neste estudo, que possui como característica adaptar-se com facilidade a uma grande variedade de problemas. Porém, para o conjunto de dados específicos ao qual foi apresentando, não obteve um bom desempenho comparativamente aos demais, sendo superior apenas aos algoritmos GDX e OSS.

Das redes treinadas foram selecionadas 2 redes de cada camada testada do algoritmo BR que obteve o melhor desempenho. Dos demais algoritmos LM, RP, SCG e OSS selecionou-se 1 rede de cada camada, totalizando 18 redes. Desta análise será definida a melhor rede para cada método e algoritmo avaliado.

Analisando-se os dados da Tabela 4.3, observa-se os baixos valores dos erros do treinamento e validação, o que mostra que as duas etapas apresentaram boa convergência para os dados, demonstrando que o treinamento foi bem sucedido. Além disso, a metodologia e os critérios de parada aplicados foram eficazes e as redes aprenderam os padrões necessários sem especialização no conjunto de treinamento.

Um detalhe da Tabela 4.3 a ser destacado é o grande número de épocas do algoritmo *Resilient Backpropagation* (RP), em todas as camadas testadas, quando comparado com os demais algoritmos avaliados. Este fato vem de encontro à característica desse algoritmo comentada anteriormente, de ser um algoritmo rápido e que necessita de pouca memória e recurso computacional, em virtude da atualização dos pesos que dependem apenas da seqüência de troca de sinais das derivadas parciais. Este fato proporciona que ele realize um grande número de interações em pouco tempo durante seu treinamento. Este conjunto de interações são apresentadas pelo software e são denominadas de épocas.

A Figura 4.44, ilustra o comportamento final da dispersão dos dados estimados e esperados para todos algoritmos de treinamento avaliados. As curvas são relativas aos melhores resultados obtidos com determinada arquitetura de rede na etapa de treinamento. É importante destacar que o conjunto de dados foi o mesmo para todos os algoritmos avaliados. O diferencial entre os resultados pode ser atribuído ao método de otimização empregado por cada algoritmo. Observa-se que os algoritmos BR e LM obtiveram os melhores resultados, em comparações com os demais algoritmos, os quais obtiveram desempenho abaixo do esperado.

Tabela 4.3. Seleção das 18 redes com melhor desempenho na etapa do treinamento, com os seus respectivos erros.

Algoritmo	Nº de camadas	Nº de neurônios	Nº de épocas	*MSE_tre*	**MSE_val	**MAPE_val
BR	1	[20]	582	1,3261E-7	5,1554E-7	0,04363
BR	1	[25]	571	1,3250E-7	3,3635E-7	0,04144
LM	1	[25]	330	1,1642E-7	1,6420E-6	0,08114
RP	1	[10]	5892	8,0733E-7	9,9194E-7	0,07499
SCG	1	[20]	479	1,3160E-6	1,6087E-6	0,09352
OSS	1	[5]	1085	1,5922E-6	2,0365E-6	0,11116
<b>BR</b>	<b>2</b>	<b>[15 10]</b>	<b>397</b>	<b>5,5947E-8</b>	<b>2,0889E-7</b>	<b>0,02963</b>
BR	2	[20 20]	523	2,0275E-8	8,7008E-7	0,04931
LM	2	[15 10]	316	5,8620E-8	6,4710E-7	0,05284
RP	2	[10 10]	2983	5,0422E-7	6,8289E-7	0,05898
SCG	2	[5 5]	933	1,0180E-6	1,1192E-6	0,07994
OSS	2	[20 20]	1102	1,2339E-6	1,6864E-6	0,10116
BR	3	[10 10 10]	446	4,7203E-8	3,8431E-7	0,03293
BR	3	[30 30 30]	726	1,3763E-8	1,8444E-6	0,07967
LM	3	[10 10 10]	337	3,5567E-8	2,3724E-7	0,03317
RP	3	[15 10 5]	3754	4,5017E-7	2,7684E-6	0,11451
SCG	3	[15 10 5]	1103	5,7084E-7	6,8263E-7	0,06185
OSS	3	[5 5 5]	1054	1,4961E-6	2,0389E-6	0,10529

\*MSE\_tre = Erro médio quadrático na etapa do treinamento.

\*\*MSE\_val e MAPE\_val = Erro médio quadrático e erro percentual médio absoluto na etapa da validação.

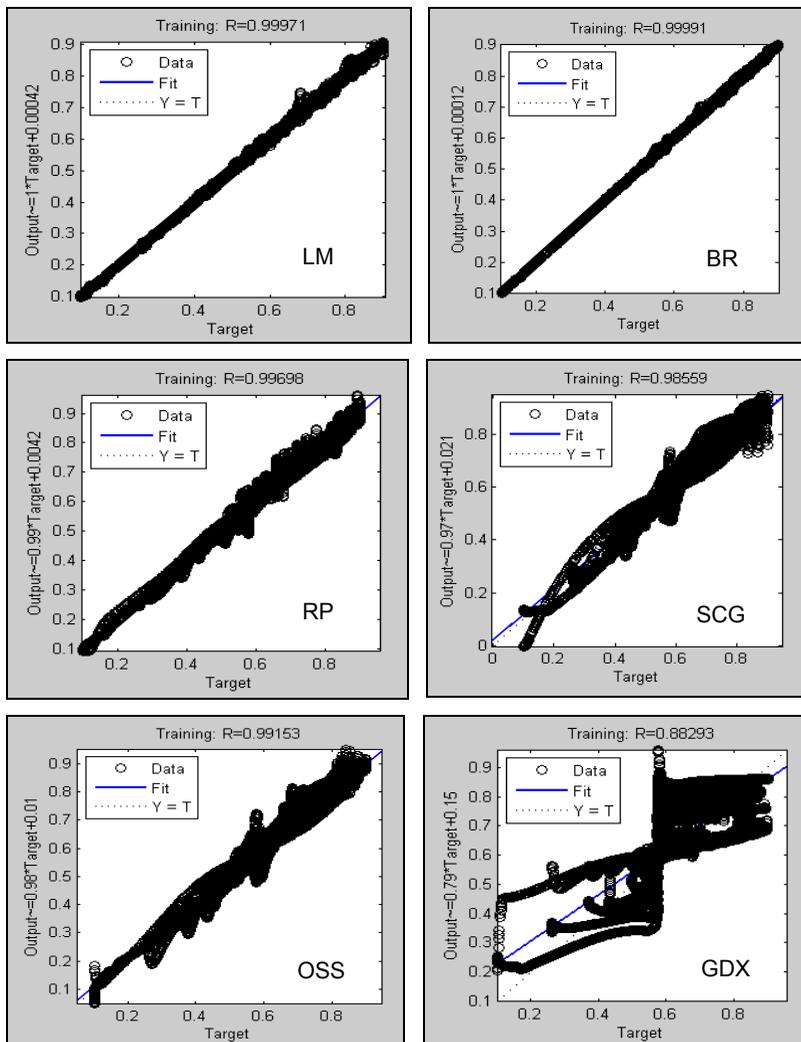


Figura 4.44. Dispersão dos dados estimados pela rede e esperados, melhores resultados para cada algoritmo avaliado na etapa de treinamento.

Todos os algoritmos empregam técnicas diferenciadas na determinação dos padrões de treinamento e diferem entre si basicamente pela estratégia de otimização utilizada para realizar os ajustes nos pesos da rede neural.

### 4.3.2.2 Medidas de desempenho das redes MPL avaliadas – Etapa da validação cruzada

Para o algoritmo GDJ, que também fez parte dos testes, o desempenho ficou muito abaixo da média das demais redes (conforme Figuras 4.42 e 4.43). Por esta razão, não foi inserido na Figura 4.45, devido à disparidade da escala. Para este algoritmo a rede de 2 camadas intermediárias com 10 neurônios em cada camada foi o melhor resultado obtido, com erro médio quadrático de 1,1565E-5 na etapa de treinamento e 1,6751E-5 na etapa da validação, com erro médio percentual médio absoluto (MAPE) de 0,35485.

A Figura 4.45 ilustra a comparação do desempenho das 18 melhores redes (selecionadas de acordo com o resultado de seu treinamento) com os algoritmos avaliados na etapa validação cruzada.

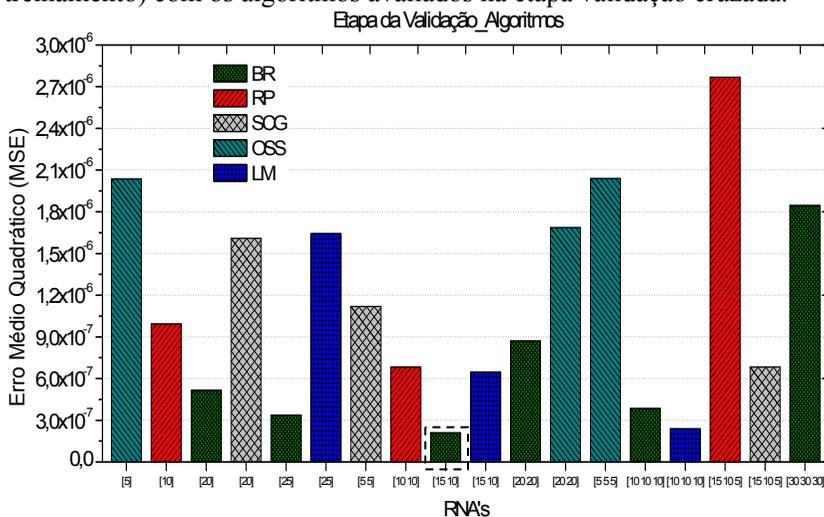


Figura 4.45. Comparação do desempenho das melhores redes com os algoritmos avaliados, na etapa validação cruzada.

Pela análise dos dados contidos na Figura 4.45 conclui-se que a rede com o algoritmo de treinamento BR com 2 camadas intermediárias com 15 e 10 neurônios em cada obteve o melhor desempenho, atingindo uma região com baixo erro médio quadrático. A rede de 3 camadas com 10 neurônios em cada empregando o algoritmo LM obteve o segundo melhor desempenho. Para o algoritmo RP a melhor rede foi de 2 camadas intermediárias com 10 neurônios em cada. Para o algoritmo SCG a melhor foi a rede de 3 camadas com 15, 10 e 5 neurônios,

respectivamente. O algoritmo de treinamento OSS obteve o segundo pior desempenho na validação (confirmando os resultados do treinamento). Para este algoritmo a melhor rede obtida foi de 2 camadas com 20 neurônios em cada camada intermediária.

Na Figura 4.46 é apresentada a comparação entre a saída simulada e a predição das redes que obtiveram melhor desempenho para os dados de validação cruzada, com seus respectivos algoritmos de treinamento.

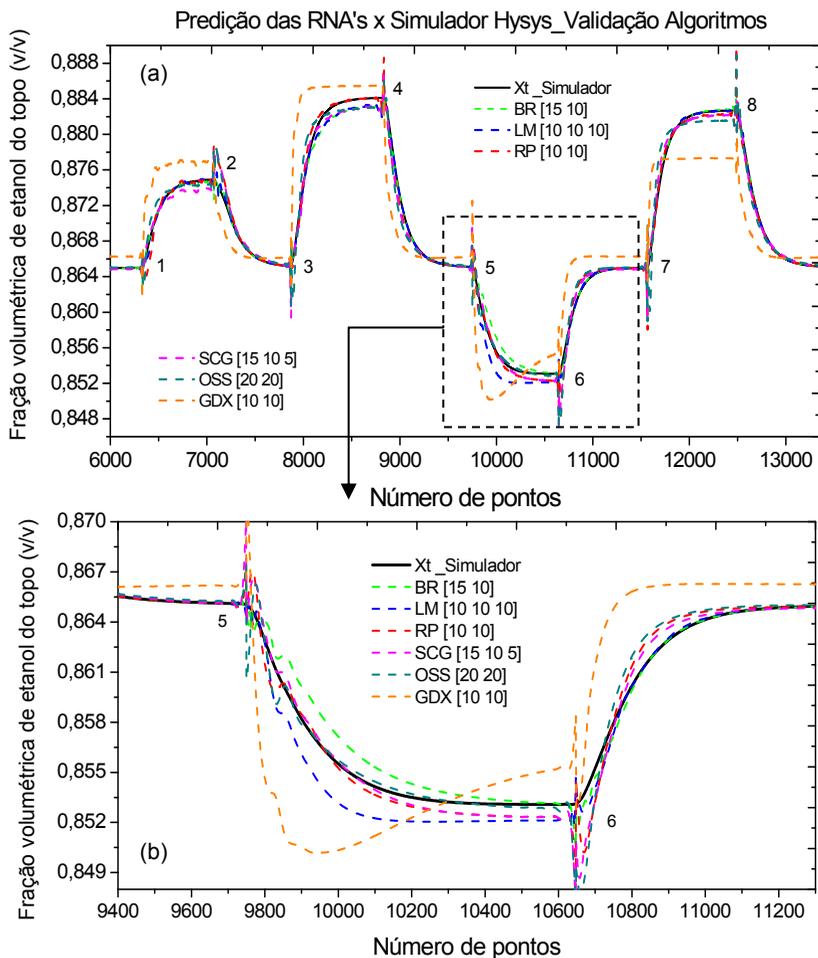


Figura 4.46. (a) Comparação entre a saída simulada e a predição das melhores redes com seus respectivos algoritmos de treinamento, na etapa da validação; (b) ampliação de uma perturbação.

Pela análise das curvas da Figura 4.46, e auxiliado pelos valores da Tabela 4.3 onde são apresentados os erros médios quadráticos ( $MSE\_val$ ) e erros percentuais médios absoluto ( $MAPE\_val$ ), conclui-se que a rede com o algoritmo de treinamento BR com 2 camadas intermediárias, com 15 e 10 neurônios em cada, obteve a melhor resposta para o conjunto de dados de validação.

É importante salientar que dentre todos os algoritmos testados, nenhum deles pode ser considerado como melhor que o outro. Pode-se afirmar que um deles, neste caso o BR, foi o mais adaptado para um determinado conjunto de dados com características específicas e obteve boa convergência.

O algoritmo *Regularização de Bayesian* (BR) será o algoritmo de treinamento empregado na próxima etapa, onde serão avaliadas redes recorrentes de Elman. Este algoritmo representa uma modificação do algoritmo *Levenberg-Marquardt* sendo empregado para gerar redes com melhor generalização e, além disso, reduz a dificuldade de se determinar a arquitetura de rede otimizada. Estes fatores foram determinantes e contribuíram para sua escolha, aliado ao bom desempenho obtido por este algoritmo com o conjunto de dados simulados. Este foi um critério adotado para diminuir o número de testes com configurações de redes recorrentes a serem avaliadas posteriormente.

Conforme mencionado, na próxima etapa, serão testadas algumas arquiteturas e configurações de redes recorrentes de Elman, para comparar seu desempenho frente às melhores redes *feedforward* - MPL (sem realimentação) obtidas nesta seção. Pelo fato das redes recorrentes reconhecerem padrões variantes no tempo e possuírem capacidade de armazenar informações passadas, espera-se obter redes recorrentes com bom desempenho para o processo que se deseja modelar e aplicar o sensor virtual.

#### 4.3.3 Redes recorrentes de Elman

Para os testes com as redes recorrentes de Elman foram avaliadas 54 redes no total, aplicando o algoritmo de treinamento *Regularização de Bayesian* (BR) com 1, 2 e 3 camadas intermediárias, variando o número de neurônios. Foram empregadas duas combinações diferentes de funções de transferência: a primeira com a função *logarítmica* em todas as camadas inclusive, na saída e, a segunda, com função *tangente hiperbólica* nas camadas intermediárias e função *linear* na saída, sendo esta última a configuração recomendada para este tipo de rede, conforme informações contidas no software MATLAB®.

### 4.3.3.1 Medidas de desempenho das redes recorrentes de Elman avaliadas – Etapas do treinamento e validação cruzada

Para todas as redes foi calculado o erro médio quadrático (MSE) na etapa do treinamento e da validação e o erro percentual médio absoluto (MAPE) apenas na etapa da validação. De posse dos resultados, foram pré-selecionadas as 28 redes com melhor desempenho. Os resultados das medidas de desempenho destas redes são descritas na Tabela 4.4 e ilustradas nas Figuras 4.47 e 4.48.

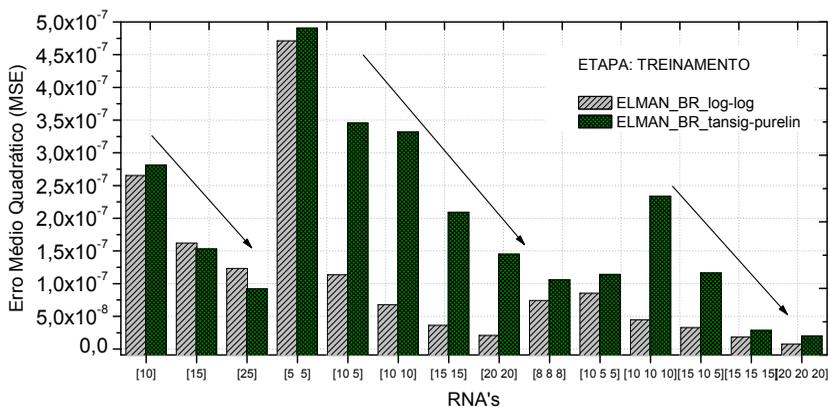


Figura 4.47. Comparativo do comportamento do erro médio quadrático (MSE) das redes de Elman, com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias, na etapa do treinamento.

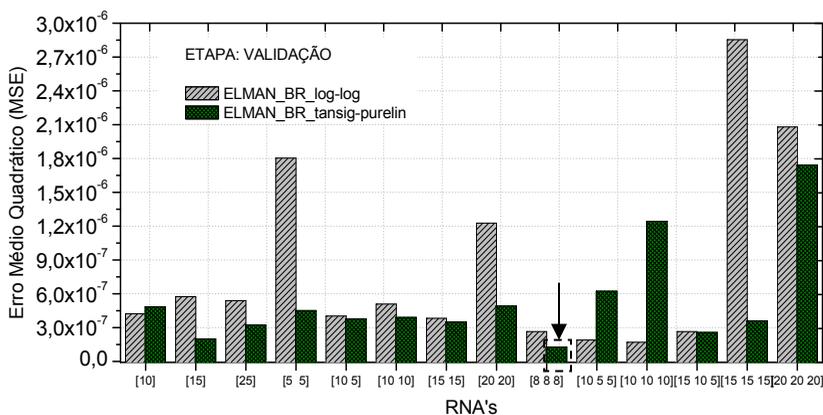


Figura 4.48. Comparativo do desempenho das melhores redes de Elman avaliadas, na etapa de validação cruzada.

Tabela 4.4. Seleção das 28 redes de Elman com melhor desempenho na etapa do treinamento, com os as funções de transferências avaliadas e seus respectivos erros.

<i>Nº de camadas</i>	<i>Funções transf. (intermed/saida)</i>	<i>Nº de neurônios</i>	<i>Nº de épocas</i>	<i>*MSE_tre</i>	<i>**MSE_val</i>	<i>**MAPE_val</i>
1	<i>logsig-logsig</i>	[10]	634	2,6542E-7	4,2270E-7	0,04889
1	<i>tansig-purelin</i>	[10]	414	2,8143E-7	4,8540E-7	0,04873
1	<i>logsig-logsig</i>	[15]	272	1,6222E-7	5,7345E-7	0,04770
1	<i>tansig-purelin</i>	[15]	509	1,5331E-7	1,9946E-7	0,03044
1	<i>logsig-logsig</i>	[25]	716	1,2334E-7	5,2355E-7	0,04486
1	<i>tansig-purelin</i>	[25]	816	9,2311E-8	3,2237E-7	0,03666
2	<i>logsig-logsig</i>	[5 5]	450	4,7118E-7	1,8036E-6	0,10457
2	<i>tansig-purelin</i>	[5 5]	423	4,9075E-7	4,5144E-7	0,05312
2	<i>logsig-logsig</i>	[10 5]	198	1,1350E-7	4,0191E-7	0,03616
2	<i>tansig-purelin</i>	[10 5]	141	3,3856E-7	3,7694E-7	0,04538
2	<i>logsig-logsig</i>	[10 10]	584	6,8096E-8	5,0917E-7	0,04166
2	<i>tansig-purelin</i>	[10 10]	636	3,5095E-7	3,9079E-7	0,04396
2	<i>logsig-logsig</i>	[15 15]	857	3,6300E-8	3,8159E-7	0,04099
2	<i>tansig-purelin</i>	[15 15]	919	2,0957E-7	3,5044E-7	0,03638

\*MSE\_tre = Erro médio quadrático na etapa do treinamento.

\*\*MSE\_val e MAPE\_val = Erro médio quadrático e erro percentual médio absoluto na etapa da validação. Continuação da tabela na próxima página.

Continuação.

<i>Nº de camadas</i>	<i>Funções transf. (intermed/saída)</i>	<i>Nº de neurônios</i>	<i>Nº de épocas</i>	<i>*MSE_tre</i>	<i>**MSE_val</i>	<i>**MAPE_val</i>
2	<i>logsig-logsig</i>	[20 20]	785	2,0931E-8	1,2274E-6	0,05412
2	<i>tansig-purelin</i>	[20 20]	1511	1,4519E-7	4,9103E-7	0,04897
3	<i>logsig-logsig</i>	[8 8 8]	623	7,4143E-8	2,6487E-7	0,03346
3	<b><i>tansig-purelin</i></b>	<b>[8 8 8]</b>	<b>250</b>	<b>1,0578E-7</b>	<b>1,2680E-7</b>	<b>0,02695</b>
3	<i>logsig-logsig</i>	[10 5 5]	480	8,5284E-8	1,9008E-7	0,02845
3	<i>tansig-purelin</i>	[10 5 5]	321	1,1422E-7	6,2348E-7	0,04930
3	<i>logsig-logsig</i>	[10 10 10]	610	4,4963E-8	1,6921E-7	0,02866
3	<i>tansig-purelin</i>	[10 10 10]	209	2,3407E-7	1,2441E-6	0,07945
3	<i>logsig-logsig</i>	[15 10 5]	863	3,3024E-8	2,6422E-7	0,02983
3	<i>tansig-purelin</i>	[15 10 5]	401	1,1645E-7	2,5807E-7	0,03397
3	<i>logsig-logsig</i>	[15 15 15]	900	1,8400E-8	2,8539E-6	0,10494
3	<i>tansig-purelin</i>	[15 15 15]	920	2,8642E-8	3,6081E-7	0,03250
3	<i>logsig-logsig</i>	[20 20 20]	1310	9,5625E-9	2,0796E-6	0,07253
3	<i>tansig-purelin</i>	[20 20 20]	837	2,0413E-8	1,7419E-6	0,08098

\*MSE\_tre = Erro médio quadrático na etapa do treinamento.

\*\*MSE\_val e MAPE\_val = Erro médio quadrático e erro percentual médio absoluto na etapa da validação.

Novamente percebe-se pela análise da Figura 4.47 que na etapa do treinamento, dentro da mesma camada, o erro vai diminuindo à medida que se aumenta o número de neurônios. Este comportamento é ilustrado pelas setas contidas no gráfico analisado.

Ainda analisando-se os dados contidos na Figura 4.47, verifica-se que as redes com as funções de transferência *logarítmica* aplicadas nas camadas intermediárias e na saída obtiveram desempenho superior na etapa do treinamento, frente às redes com funções *tangente hiperbólica* com saída *linear*. Porém, durante a validação (Figura 4.48), essa tendência não foi mantida. As redes com funções de transferência *tangente hiperbólica* com saída *linear* obtiveram maior capacidade de generalização aos novos dados, apresentando erros menores, sendo esta última recomendada pelo manual do *neural network toolbox* do MATLAB® para redes de Elman. Desta forma, o desempenho das redes recorrentes deve ser avaliado paralelamente, isto é, conjuntamente nas etapas de treinamento e validação.

Na Figura 4.49 são ilustradas as 4 melhores redes de Elman. Observa-se sobreposição dos valores fornecidos pelo simulador e pelas redes (preditos) para composição de produto de topo da coluna. Destaque para a rede de 3 camadas com 8 neurônios em cada camada que usa a função de transferência *tangente sigmóide* nas camadas intermediárias e com saída *linear*.

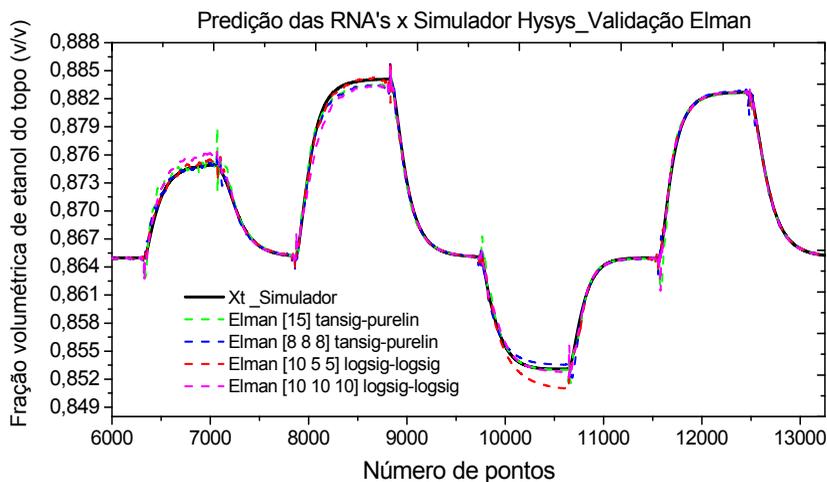


Figura 4.49. Comparação entre a saída simulada e a predição das 4 melhores redes de Elman, na etapa da validação.

Percebe-se capacidade superior de predição da composição de etanol no topo da coluna com o emprego das redes recorrentes quando se confronta com os resultados das redes MPL, com diversos algoritmos de treinamento (Figura 4.46). Este fato pode estar relacionado com as unidades de contexto que as redes de Elman possuem que servem para memorizar as ativações anteriores das camadas intermediárias. Nestas unidades ficam armazenadas informações atrasadas no tempo, as quais constituem a memória da rede, fato que atribui grande diferencial destas redes em comparação com as MPL's.

#### 4.3.4 Comparativo entre redes Multi-Layer Perceptron (MPL) e Redes Recorrentes de Elman

Nesta seção é realizada uma comparação entre o desempenho das redes *Multi-Layer Perceptron* (MPL) que empregam *backpropagation* com as redes recorrentes de Elman com realimentação. Foram confrontadas as duas melhores redes obtidas (uma rede MPL e outra rede recorrente de Elman) dentre todas as redes avaliadas. O resultado pode ser observado na Figura 4.50. Pela sua análise pode-se concluir que a rede recorrente de Elman com 3 camadas e 8 neurônios em cada camada, que usa a função de transferência *tangente sigmóide* nas camadas intermediárias com saída *linear* empregando o algoritmo de treinamento *BR* é a melhor rede para o conjunto de dados usados na validação. A rede *feedforward* - MPL de 2 camadas com algoritmo *BR* também obteve excelente resultado. Ambas poderiam ser empregadas no sensor para predição da composição de etanol no topo, mesmo tendo características distintas. Percebe-se que as redes apresentam erros admissíveis para as diversas circunstâncias pela qual o processo pode migrar durante sua operação.

Porém, a vantagem das redes recorrentes, frente às redes *feedforward-MPL*, é que elas incorporam a dinâmica do processo dentro da rede, onde ocorre a realimentação interna (a cada interação) das saídas de cada camada intermediária para as unidades de contexto e, por conseguinte, estas informações são repassadas a primeira camada intermediária da rede, propagando-se para as demais durante seu treinamento. Esta realimentação dos valores internos em um passo seguinte torna esta rede capaz de armazenar informações do passo anterior e utilizá-las em um momento futuro. Este fato confere a elas capacidade superior de predição de valores futuros, sendo capaz de inferir e reconhecer padrões temporais e espaciais, principalmente durante períodos transientes.

Já as redes MPL incorporam a dinâmica do processo fora da rede, pelos próprios dados dinâmicos que a rede é alimentada. É um tipo de rede sem realimentação, onde a informação é distribuída pela rede em um só sentido entre as camadas. No seu treinamento, as derivadas do erro apresentado pela rede são utilizadas na correção dos pesos internos, de maneira retroativa. Apresentam um bom desempenho, conforme observado. Porém, não possuem tanta capacidade preditiva em situações extremas, mas obteve um ótimo desempenho para o conjunto de dados específicos que foi treinada e validada.

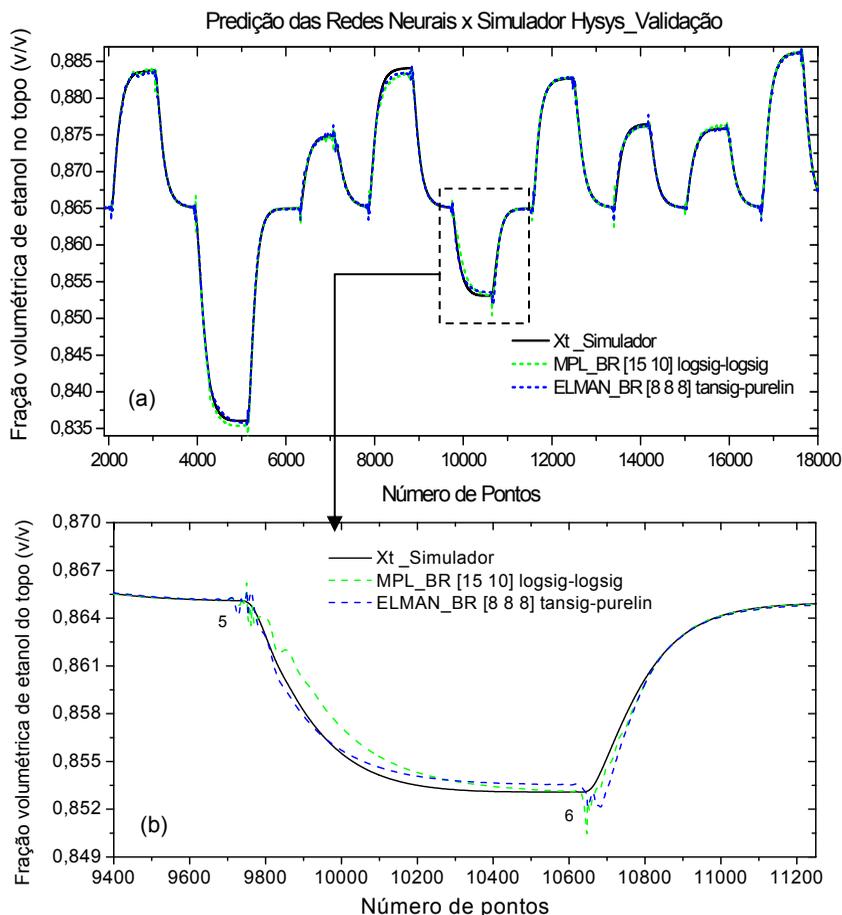


Figura 4.50. (a) Comparação entre a saída simulada e a predição das 2 melhores redes gerais MPL x ELMAN, na etapa da validação; (b) ampliação de uma perturbação.

Pela análise dos resultados da Figura 4.50 percebe-se que a predição da rede acompanha muito bem as tendências das curvas dos dados simulados, com erros bastantes homogêneos mesmo para um grande número de pontos (18.610 pontos, equivalente a 51,7 horas de operação da unidade simulada). Apresentam erros um pouco maiores nos picos quando a tendência das curvas é alterada devido a aplicação de alguma perturbação mas, logo em seguida, acompanha a tendência. Este fato comprova que o treinamento (efetuado com 78.833 pontos, equivalente a 220 horas de operação da coluna simulada) foi bastante abrangente.

Uma ilustração das arquiteturas finais das 2 redes que obtiveram melhor desempenho nesta etapa, para o conjunto de dados da validação, pode ser observado na Figura 4.51.

Ao efetuar-se a comparação da arquitetura das redes MPL e Elman, na rede de Elman ocorre uma redefinição da camada de entrada, que passa a ser dividida em duas partes. A primeira parte corresponde ao vetor de entrada propriamente dito. A segunda parte, chamada de unidades de contexto, consiste na cópia das saídas dos neurônios da camada escondida no instante  $n-1$ . A camada de contexto da rede de Elman é responsável por armazenar a informação temporal relativa às entradas anteriores, o que a diferencia da rede MPL.

Neste trabalho, as ativações de cada uma das camadas intermediárias são realimentadas para as unidades de contexto. No caso específico da rede descrita na Figura 4.51b, as ativações das três camadas intermediárias foram realimentadas.

Com relação à quantidade de unidades de contexto, esta depende diretamente de quantas camadas serão feitas as realimentações. Na rede de Elman selecionada, com 3 camadas intermediárias, conforme Figura 4.51b, temos uma camada com 24 unidades de contexto. Esta camada contém informações atrasadas no tempo que constituem a memória da rede e complementam as 8 entradas dinâmicas da rede.

As unidades de contexto são usadas para memorizar as ativações anteriores das unidades intermediárias e podem ser consideradas como atraso no tempo em um passo. As conexões *feedforward* são modificáveis, e as conexões recorrentes podem ser fixas, motivo pelo qual a rede de Elman é classificada como parcialmente recorrente. Do ponto de vista das conexões, a rede de Elman possui apenas recorrências globais.

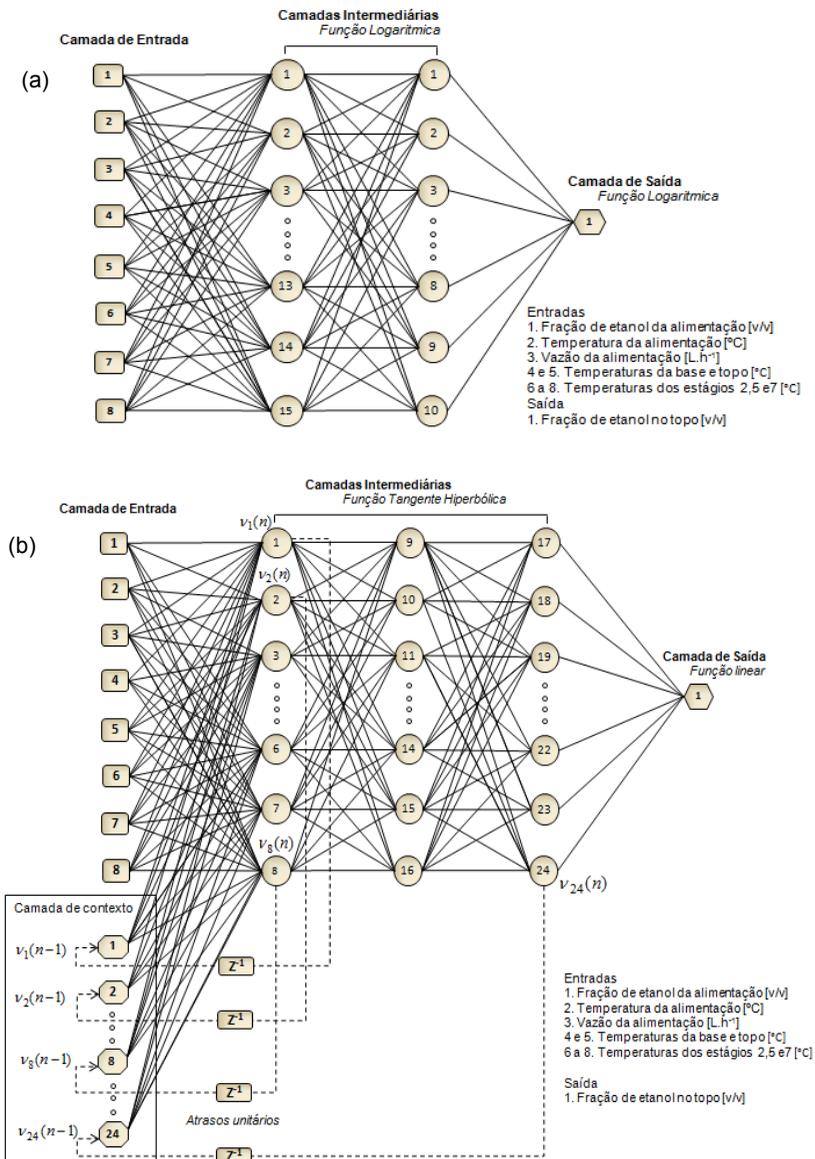


Figura 4.51. Ilustração das arquiteturas finais das 2 redes com melhor desempenho nesta etapa, (a) rede *feedforward* – MPL com 2 camadas intermediárias; (b) rede recorrente de Elman com 3 camadas intermediárias.

Na Figura 4.52 está representado o diagrama de blocos proposto para futura implementação do sensor por software na unidade piloto. Propõem-se um sistema de controle com algoritmo de controle inferencial para o controle de composição do destilado da coluna. As composições do topo estimadas pelo sensor baseado no modelo neural seriam usadas como entradas atuais para o sistema de controle (seria considerada como variável medida) e as temperaturas de alguns pratos da coluna, por exemplo, poderiam ser utilizadas como variáveis manipuladas pelo controlador. A pureza do produto desejado seria o ponto de ajuste do controlador (variável controlada).

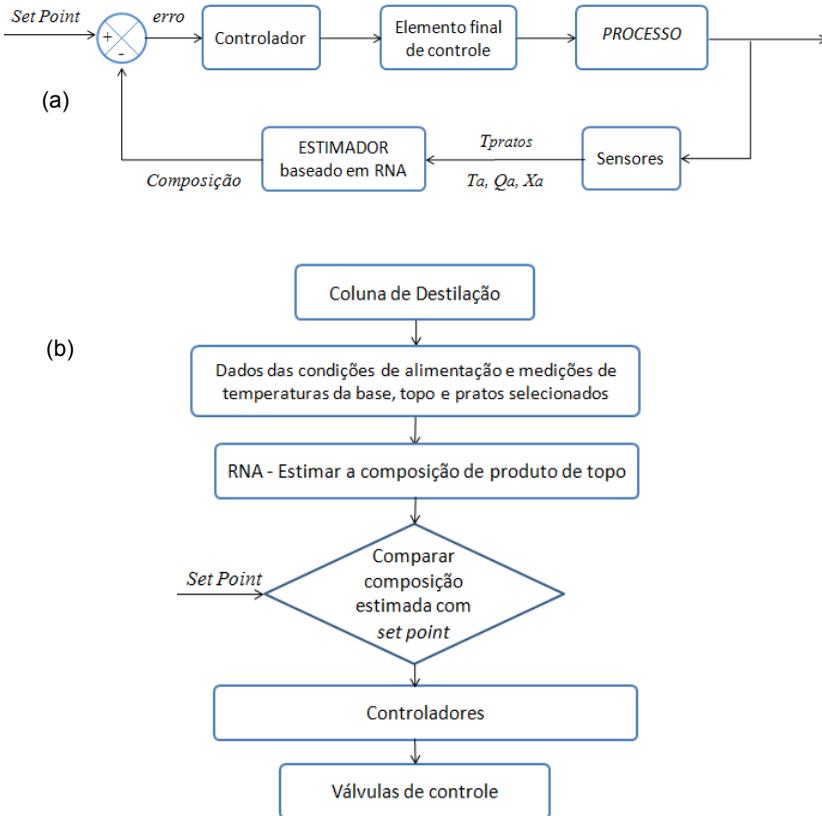


Figura 4.52. (a) Diagrama de blocos e (b) fluxograma simplificado das etapas, com a proposta do sensor por software para predição da composição de etanol no topo para ser aplicada no controle da unidade.

Com o intuito de aperfeiçoar o sistema aplicando técnicas de controle mais avançadas, torna-se necessário realizar uma otimização futura das variáveis que foram usadas como entradas da rede neural, ou seja, o caminho inverso do treinamento da rede, determinando qual o valor ótimo de cada variável de entrada para determinado valor de composição de topo (produto de interesse). Desta forma, pode-se estabelecer vários *set points* para as variáveis controladas, decorrentes de cada especificação de produto de topo requerido. Uma opção interessante seria usar as temperaturas dos pratos escolhidos como entradas da rede ( $T_b$ ,  $T_t$ ,  $T_2$ ,  $T_5$ ,  $T_7$ ) como variáveis controladas, fazendo-se uso para tal, do aquecimento individualizado em cada prato com o emprego da resistência elétrica.

Assim, quando ocorrer mudança do ponto de operação (mudança de campanha em unidades industriais, por exemplo, onde se deseja mudar a composição do produto requerido) o sensor virtual fornecerá a medida *on-line* de composição de topo estimada e o sistema de controle atuará de modo a ajustar os valores das temperaturas dos estágios pré-determinados pela otimização.

Outra abordagem de controle que também se apresenta como uma alternativa para ser empregada na futura implementação do sensor na unidade piloto, se refere ao sistema com lógica *fuzzy* o qual poderia ser empregado acoplado ao sistema de inferência, responsável por estimar a composição de topo da coluna. Desta forma, o sistema *fuzzy* poderia ser ajustado de maneira a gerar dinamicamente os *set points* para os controladores PID's de temperaturas, relacionados com as variáveis de entrada da rede, possibilitando assim, a realização de controle da qualidade do topo da coluna. Araújo Junior (2007) aplicou esta técnica em uma coluna debutanizadora simulada para controle da fração de i-pentano presente no GLP.

Em um controle *fuzzy* utilizam-se regras lógicas (SE <premissa>; ENTÃO <conclusão>) no algoritmo de controle em função das diversas faixas de valores que as variáveis de estado do problema podem assumir, com a intenção de descrever em uma rotina a experiência humana, intuição e heurística para controlar um processo, ao contrário dos controladores convencionais, em que o algoritmo de controle é descrito analiticamente por equações algébricas ou diferenciais.

Nas plantas industriais de processos petroquímicos, técnicas de controle tem sido introduzidas em um nível da escala de automação entre o controle regulatório e o sistema supervisório. Com isso, são mantidas as malhas do controle regulatório, pois são vistas como seguras

e confiáveis. As técnicas mais sofisticadas (chamadas de controle avançado) são aplicadas em nível de otimização e na geração de *set points* para o controle regulatório.

Atualmente, devido à disponibilidade de equipamentos com elevada capacidade de processamento, estas técnicas tem sido aplicadas com excelentes resultados. Diversos trabalhos são encontrados na literatura (SOMSUNG e PRATISHTHAMANDA, 2005; SALAHSHOOR e HAMZEHNEJAD, 2010), alguns empregando controladores *fuzzy*, na grande maioria atuando como supervisores, ou seja, em vez de gerar *set points*, realizam o ajuste dos parâmetros dos controladores geralmente da camada do controle regulatório.

Com o emprego destas técnicas conjuntas: sensor por software fornecendo o valor da variável controlada *on-line* ao sistema de controle e, aquecimento distribuído, com ações rápidas e localizadas sobre cada estágio diminuindo o tempo decorrente da propagação do efeito corretivo com ações somente na base, espera-se criar um ambiente inovador de controle aplicado a colunas de destilação, com redução de transientes e melhora do consumo energético.

#### 4.3.5 Avaliação final em relação ao modelo empírico obtido para ser empregado como analisador virtual

Após análise dos estudos, chega-se a uma rede neural recorrente de Elman, que reproduz de forma muito precisa o comportamento dinâmico do processo, podendo ser considerada um excelente modelo empírico da unidade piloto de destilação. Desta forma, poderá ser utilizada como preditor da variável de saída composição volumétrica de etanol no topo da coluna para fins de otimização e controle no processo de destilação. Esta rede, após ter sido treinada com 78.833 pontos oriundos de um simulador dinâmico, apresentou um erro bastante pequeno (em torno de  $10^{-7}$ ) para o conjunto de dados simulados no qual ela foi validada. Pode-se afirmar, portanto, que o objetivo de desenvolver um sensor por software com capacidade de estimar a composição futura do produto de topo de uma coluna de destilação simulada foi atingido.

Os resultados obtidos demonstram que a rede neural é um modelo empírico dinâmico adequado da coluna de destilação, ainda que esta seja um sistema acoplado e não-linear. Além disso, a rede possui característica robusta e com excelente capacidade de generalização e prevê o processo com boa acurácia. Esta rede pode ser utilizada como modelo para projeto de controladores, otimização de operação e

implementada como modelo para o sensor por software na unidade piloto de destilação com ações distribuídas de controle, empregando aquecimento individual em cada prato.

#### **4.4 MODELO EMPÍRICO DINÂMICO BASEADO EM REDES NEURAIIS – MALHA FECHADA**

Atingido o objetivo principal desta tese, que foi desenvolver um sensor por software em malha aberta baseado em redes neurais artificiais para modelar uma coluna com aquecimentos distribuídos, para que este possa ser empregado para projeto de controladores, e consequente otimização de operação de destilação, objetivou-se também encontrar uma rede neural capaz de prever a composição de topo do processo em malha fechada, levando em consideração todo acoplamento das variáveis e do sistema de controle da unidade. Para tal, foi utilizado o modelo do processo, desenvolvido no software comercial HYSYS<sup>®</sup>, em seu modo dinâmico, simulando a operação do processo em malha fechada, devidamente validado experimentalmente, conforme detalhado neste capítulo na *seção 4.2.1*. No simulador desenvolvido em malha fechada, foram obtidos os dados necessários para o treinamento e desenvolvimento da nova rede dinâmica, que consistiu num grande desafio, devido a complexidade e grande número de fatores/variáveis envolvidos.

O desenvolvimento desta rede possibilita outra opção, que seria implementar este sensor em um *loop* externo, com o intuito de inferir e fazer o monitoramento da composição de topo 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 e manutenção.

Foi dada atenção especial em relação à seleção das entradas das novas redes, pois com o sistema operando em malha fechada, a composição do topo sofre influência de outras variáveis do processo, principalmente das variáveis manipuladas do sistema de controle. Desta forma, foram testadas redes com 10 entradas ( $X_a$ ,  $T_a$ ,  $Q_a$ ,  $T_b$ ,  $T_t$ ,  $T_2$ ,  $T_5$ ,  $T_7$ ,  $\emptyset_{ref}$ ,  $R$ ) onde foram incluídos o calor do refeedor ( $\emptyset_{ref}$ ) e a vazão de refluxo ( $R$ ) em relação ao conjunto de variáveis que foi usado na rede em malha aberta. Acredita-se que as informações fornecidas por estas variáveis são relevantes para a construção dos padrões de treinamento pelo algoritmo empregado, pois a composição de topo é

função destas. A variável de saída foi mantida a mesma, composição volumétrica de etanol no topo na coluna ( $X_t$ ).

As temperaturas da coluna, as quais afetam diretamente a composição, sofrem uma influência dupla. Num primeiro momento são afetadas pelas perturbações aplicadas nas variáveis de entradas da corrente de alimentação ( $X_a$ ,  $T_a$ ,  $Q_a$ ) e, após, sofrem o impacto das variáveis manipuladas do sistema de controle ( $\emptyset_{ref}$ ,  $R$ ) tentando ajustar/manter o ponto de operação nominal do processo. A rede precisará entender estes ajustes para as variáveis: calor do refeedor ( $\emptyset_{ref}$ ) e refluxo ( $R$ ) e criar padrões durante o treinamento para poder correlacioná-los.

As redes com 8 entradas também foram avaliadas, pois acredita-se que ambos efeitos podem estar bem representados apenas pelas alterações dos perfis de temperatura. Nas redes com 8 entradas, as variáveis de entradas são as mesmas empregadas na rede em malha aberta, ou seja  $X_a$ ,  $T_a$ ,  $Q_a$ ,  $T_b$ ,  $T_t$ ,  $T_2$ ,  $T_5$ ,  $T_7$ . Os resultados de quais conjuntos de entradas (com 8 ou 10 entradas) são mais adequados, dependerão da forma com que a rede conseguirá organizar os dados e determinar padrões de treinamento, de modo que tenha capacidade de realizar previsões futuras com novos dados, dentro dos limites do treinamento.

Os dados obtidos deste processo, agora em malha fechada, também são representativos de toda amplitude na qual se deseja que o sensor venha a operar. A faixa de operação da unidade no simulador com os controladores operando no modo automático pôde ser aumentada devido à ação de controle que evita que as variáveis se afastem do ponto de operação nominal do processo, mesmo quando aplicadas perturbações de maior magnitude. Este fato foi constatado durante a validação experimental do simulador em malha fechada, etapa que forneceu inúmeras informações para conhecimento do comportamento da dinâmica do processo em malha fechada.

A faixa de trabalho empregada pode ser considerada bastante representativa e abrangendo todos os pontos de operação da unidade piloto, empregados experimentalmente. Para obtenção do conjunto de dados do treinamento e validação foram aplicadas perturbações nas variáveis de entrada da coluna, usando combinações entre as variáveis: vazão, temperatura e composição volumétrica de etanol, conforme descrito com detalhes na metodologia *Capítulo III, seção 3.3.1.5, Tabela 3.4*.

O tempo de amostragem continuou sendo de 10 segundos, pois tendo conhecimento da dinâmica do processo, entende-se que este

tempo seja suficiente para identificação dos efeitos gerados pelas perturbações nas variáveis avaliadas. O intervalo de aplicação entre as perturbações foi determinado em função do tempo necessário até o sistema de controle rejeitar a perturbação e retornar ao estado estacionário anterior à perturbação. Este tempo é dependente da magnitude da perturbação aplicada e é variável para cada perturbação.

#### 4.4.1 Redes Neurais Multi-Layer Perceptron (MPL) e recorrentes de Elman

Nesta etapa foram testadas e avaliadas 78 redes neurais. Foram desenvolvidas redes recorrentes e MPL, ambas com 1, 2 e 3 camadas intermediárias, onde foi empregado o algoritmo de treinamento *Regularização de Bayesian* (BR). Para as redes recorrentes empregaram-se nas camadas intermediárias funções de transferência *tangente hiperbólica* com saída *linear*, e para redes MPL, a função de transferência foi a *logarítmica*. Estas escolhas foram baseadas nos melhores resultados obtidos nos treinamentos anteriores, realizados com as redes em malha aberta. Sabe-se que talvez estes parâmetros não obtenham o mesmo desempenho com as redes que serão testadas nesta etapa, em virtude da alteração do conjunto de dados. Porém, esta decisão foi tomada para reduzir o universo dos testes. Acredita-se que com as configurações testadas serão obtidas redes com boa capacidade de predição. As arquiteturas e números de neurônios testados também foram baseados nos melhores resultados obtidos para rede em malha aberta, no qual foi desenvolvido o *soft sensor*.

Nesta etapa, algumas redes MPL com 8 e 10 entradas com o algoritmo BR também foram avaliadas, porém com funções de transferência *tangente hiperbólica* com saída *linear*. No entanto, as redes MPL com estas funções de transferência apresentaram resultados com desempenho inferiores, quando comparadas com as redes testadas com função *logarítmica* nas camadas intermediárias e na saída. Em função disto, seus resultados não serão apresentados, sendo estas redes (26) descartadas.

Na etapa do treinamento a rede foi alimentada com 151.902 pontos, formando uma matriz de entrada  $X_{tre}$   $\langle 151.902 \times 10 \rangle$  para conjunto de 10 entradas e uma matriz  $X_{tre}$   $\langle 151.902 \times 8 \rangle$  para conjunto de 8 entradas. Em ambos os casos foi gerada uma matriz de saída  $Y_{tre}$   $\langle 151.902 \times 1 \rangle$ , conforme descrito na metodologia *Capítulo III, seção 3.3.2.3.5, Tabela 3.7*.

A Figura 4.53 representa os histogramas mostrando a distribuição dos conjuntos de dados usados para treinamento e para validação cruzada.

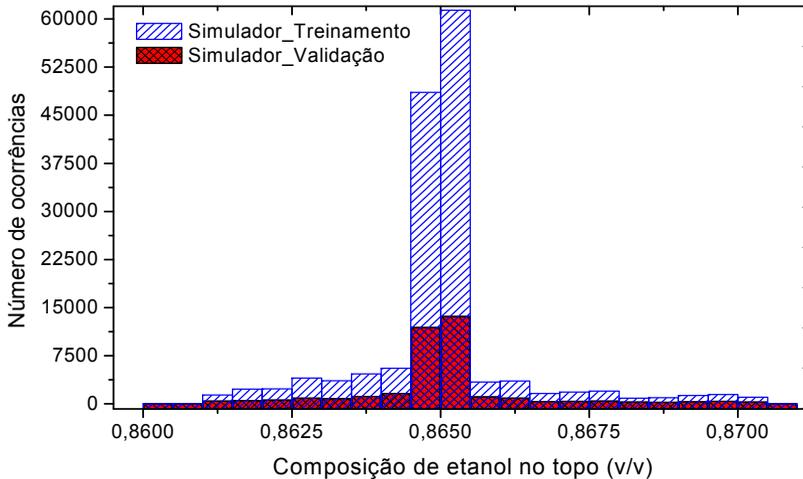


Figura 4.53. Histogramas da composição de etanol no topo da coluna para rede neural em malha fechada, dados gerados pelo simulador nas etapas do treinamento e validação.

Novamente observa-se grande quantidade de dados de composição com valor 0,865, equivalente ao valor desejado do produto de topo, em estado estacionário, na condição nominal de operação da coluna. Após cada perturbação, o sistema de controle atua sobre o processo, fazendo com que o sistema retorne ao ponto de operação não se afastando demasiadamente do valor desejado (*set point*). Este fato justifica a estreita faixa de composição do produto de topo. Observa-se ainda que os dados de validação contemplaram toda a faixa na qual as redes foram treinadas, o que confere maior confiabilidade aos resultados obtidos na etapa da validação.

#### 4.4.1.1 Medidas de desempenho das redes neurais avaliadas – Etapas de treinamento e validação cruzada

Os dados contidos na Figura 4.54 e na Figura 4.55, ilustram o comportamento do erro médio quadrático (MSE) entre redes MPL e recorrentes do tipo Elman, com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias. Em cada caso, foi

avaliado o desempenho de redes com 8 e 10 entradas, na etapa do treinamento e validação, respectivamente. Percebe-se, novamente, a tendência de redução do erro com o aumento do número de neurônios, comportamento este representado pelas setas contidas nas figuras. As redes MPL com uma camada intermediária apresentaram os maiores erros, tanto no treinamento quanto na validação. Seus limites superiores não aparecem na escala gráfica, para não comprometer a visualização das redes com melhor desempenho. No entanto, os valores ficaram na grandeza de  $10^{-6}$ . Estas redes (6) foram descartadas por não apresentarem capacidade de mapear todo conjunto de dados de forma satisfatória.

A nomenclatura das legendas contidas nas Figuras 4.54 e 4.55 representam: TIPO DE REDE; número de entradas; ALGORITMO DE TREINAMENTO; funções de transferência nas camadas intermediárias e saída.

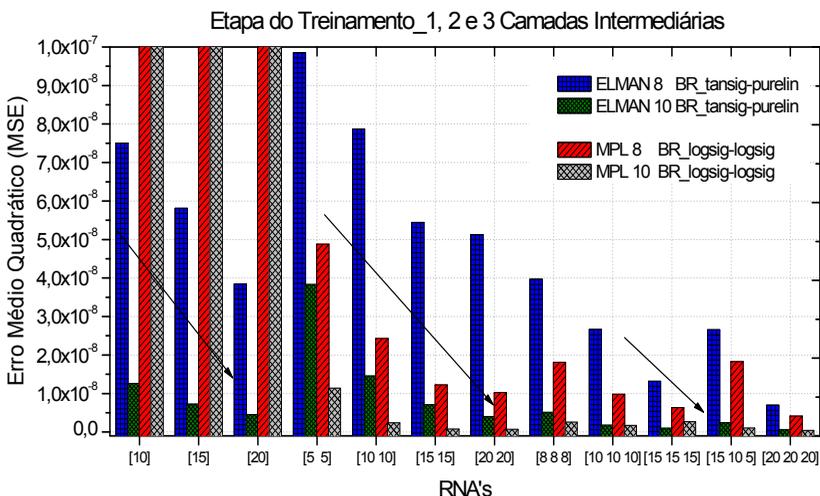


Figura 4.54. Comparativo do comportamento do erro médio quadrático (MSE) entre redes MPL e recorrentes tipo Elman, com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias, na etapa do treinamento.

Na etapa do treinamento, várias redes obtiveram bom desempenho. Porém, vale lembrar que redes com muitos neurônios nas camadas intermediárias, mesmo obtendo um bom treinamento, estão susceptíveis a sobre-treinamento e perdem a capacidade de generalização. O ideal é encontrar redes de tamanho moderado, com

bom treinamento e com alta capacidade de generalização. Na Figura 4.55, os quadrados tracejados representam algumas das redes avaliadas que contêm esta característica. Na Tabela 4.5 são apresentados os erros de 15 redes pré-selecionadas.

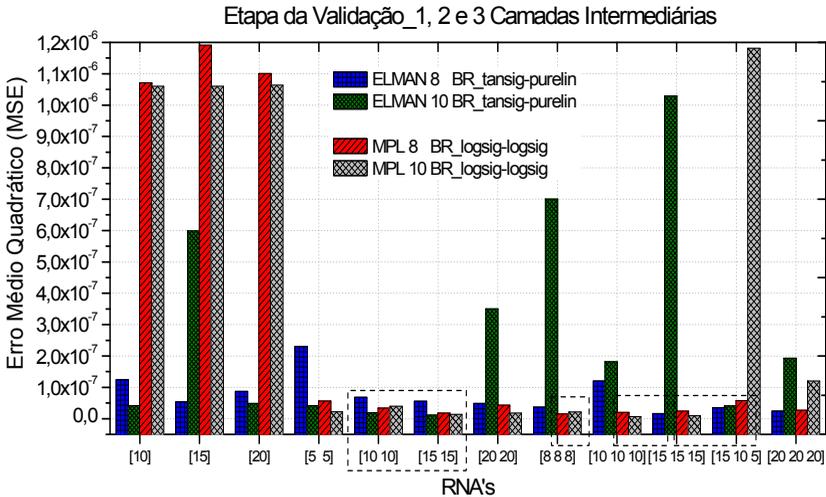


Figura 4.55. Comparativo do comportamento do erro médio quadrático (MSE) entre redes de MPL e Recorrentes com o incremento do número de neurônios para redes com 1, 2 e 3 camadas intermediárias, na etapa da validação.

Todas as redes contidas dentro dos tracejados da Figura 4.55 podem ser consideradas adequadas para predição da composição de etanol do topo. Dentre elas, a rede que obteve a melhor resposta para o conjunto de dados de validação, com o menor erro médio quadrático, foi a rede recorrente de Elman de 2 camadas intermediárias, com 15 neurônios em cada, que foi treinada com 10 entradas e com algoritmo BR. Também duas redes MPL treinadas com 10 entradas, contendo 3 camadas e 10 neurônios em cada e de 3 camadas com 15 neurônios em cada, obtiveram praticamente o mesmo desempenho e também possuem potencial para serem aplicadas como analisador virtual para monitoramento da composição de topo em tempo real.

Porém, a vantagem das redes recorrentes frente às redes *feedforward*, conforme já mencionado, é que elas incorporam a dinâmica do processo dentro da rede, através da realimentação interna durante seu treinamento. Este fato confere a elas capacidade superior de

predição de valores futuros e justamente por esta característica, a rede recorrente de Elman foi a escolhida.

Pela análise dos dados da Tabela 4.5 que contém além de informações sobre a arquitetura neural e o número de interações realizadas durante o treinamento, as medidas de desempenho das 15 redes com alta capacidade de generalização, e levando em consideração as colocações citadas acima, foi escolhida a rede de Elman de 2 camadas intermediárias, com 15 neurônios em cada, treinada com 10 entradas e com algoritmo BR para realizar o monitoramento da composição de etanol no topo da coluna e atuar como analisador virtual no processo em malha fechada.

Na Figura 4.56 é ilustrada a predição da rede neural de Elman escolhida (Elman-BR-*tansig-purelin* [15 15]) versus os resultados do simulador para fração de etanol no topo da coluna, na etapa do treinamento. Observa-se que os valores preditos pela rede praticamente se sobrepõe aos valores fornecidos pelo simulador. Para uma melhor observação da predição da rede foi realizada uma ampliação de uma seção do gráfico.

Na Figura 4.57 é apresentada a comparação entre a saída simulada e a predição da rede de Elman escolhida, que obteve um dos melhores desempenhos para os dados de validação cruzada, um dos motivos pela sua escolha, além dos já relatados anteriormente. Pela análise dos dados contidos na Figura 4.57, observa-se que nos 35.000 pontos, onde estão incorporadas todas as perturbações aplicadas, a rede foi capaz de descrever o comportamento dos dados simulados de forma eficaz em todas as regiões. Na ampliação de uma seção é possível verificar alguns pontos em que a predição da rede possui erros maiores, porém são aceitáveis, frente à magnitude da escala do qual a performance da rede está sendo avaliada e ao bom desempenho observado de forma geral. Isto mostra que a rede apresenta bom treinamento e capacidade de estimação adequada.

Tabela 4.5. Seleção das 15 redes com alta capacidade de generalização, com seus respectivos erros na etapa do treinamento e validação.

<i>Rede/ Algoritmo</i>	<i>Nº de Entradas</i>	<i>Nº de neurônios</i>	<i>Nº de épocas</i>	<i>*MSE_tre</i>	<i>**MSE_val</i>	<i>**MAPE_val</i>
MPL/BR	8	[10 10]	261	2,4280E-8	3,4279E-8	0,01035
MPL/BR	10	[10 10]	446	2,4122E-9	3,9733E-8	0,00841
MPL/BR	8	[15 15]	454	1,2218E-8	1,7624E-8	0,00980
MPL/BR	10	[15 15]	891	7,6617E-10	1,3587E-8	0,00520
MPL/BR	8	[8 8]	598	1,8118E-8	1,5220E-8	0,00798
MPL/BR	10	[8 8]	564	2,5591E-9	2,1752E-8	0,00879
MPL/BR	8	[10 10 10]	640	9,8351E-9	2,0129E-8	0,00850
MPL/BR	10	[10 10 10]	595	1,7236E-9	1,0623E-8	0,00470
MPL/BR	10	[15 15 15]	537	2,6795E-9	1,1009E-8	0,00513
ELMAN/BR	10	[20]	1294	4,5540E-9	4,8855E-8	0,01020
ELMAN/BR	10	[10 10]	1638	1,4549E-8	1,9359E-8	0,00990
ELMAN/BR	8	[15 15]	1513	5,4445E-8	5,5956E-8	0,01640
ELMAN/BR	10	[15 15]	1487	5,1180E-9	1,0573E-8	0,00466
ELMAN/BR	10	[15 10 5]	1245	2,4664E-9	4,0750E-8	0,00938
ELMAN/BR	8	[15 15 15]	421	1,3175E-8	1,6749E-8	0,00843

\*MSE\_tre = Erro médio quadrático na etapa do treinamento.

\*\*MSE\_val e MAPE\_val = Erro médio quadrático e erro percentual médio absoluto na etapa da validação.

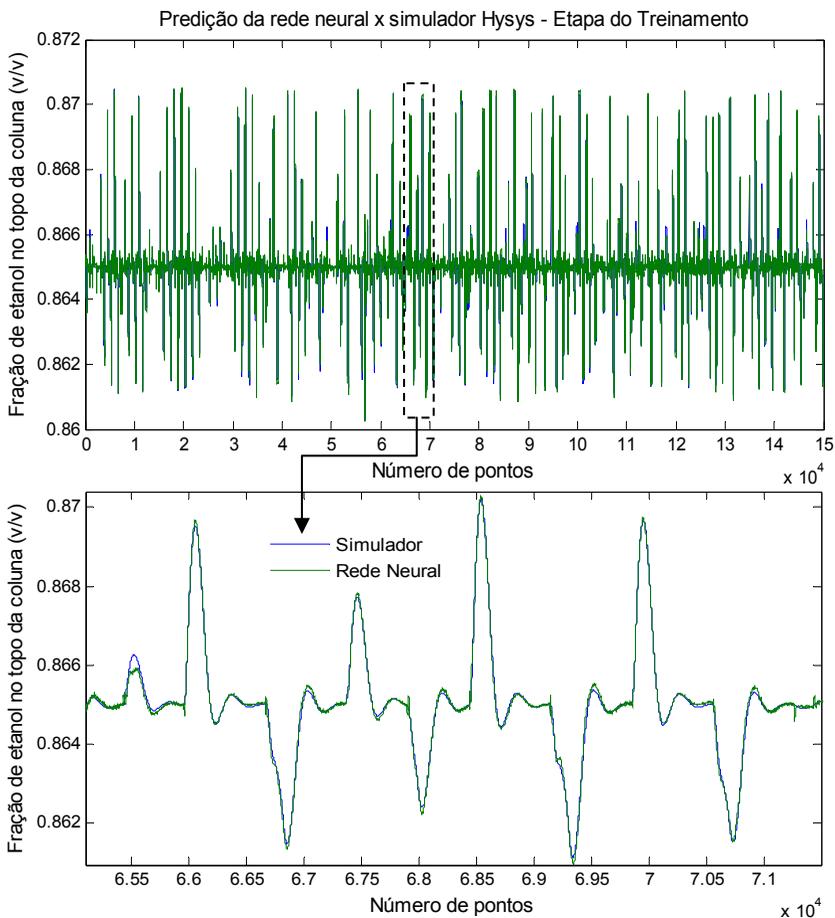


Figura 4.56. (a) Comparação entre a saída do simulador e a predição da rede recorrente selecionada (*Elman-BR-tansig-purelin* [15 15]) na etapa do treinamento; (b) ampliação de uma seção.

O bom treinamento que a rede obteve e a boa capacidade de predição em malha fechada tornam possível a sua utilização para realizar o monitoramento *on-line* da composição de produto de topo de etanol na unidade de destilação piloto, mantendo a atual configuração das malhas de controle.

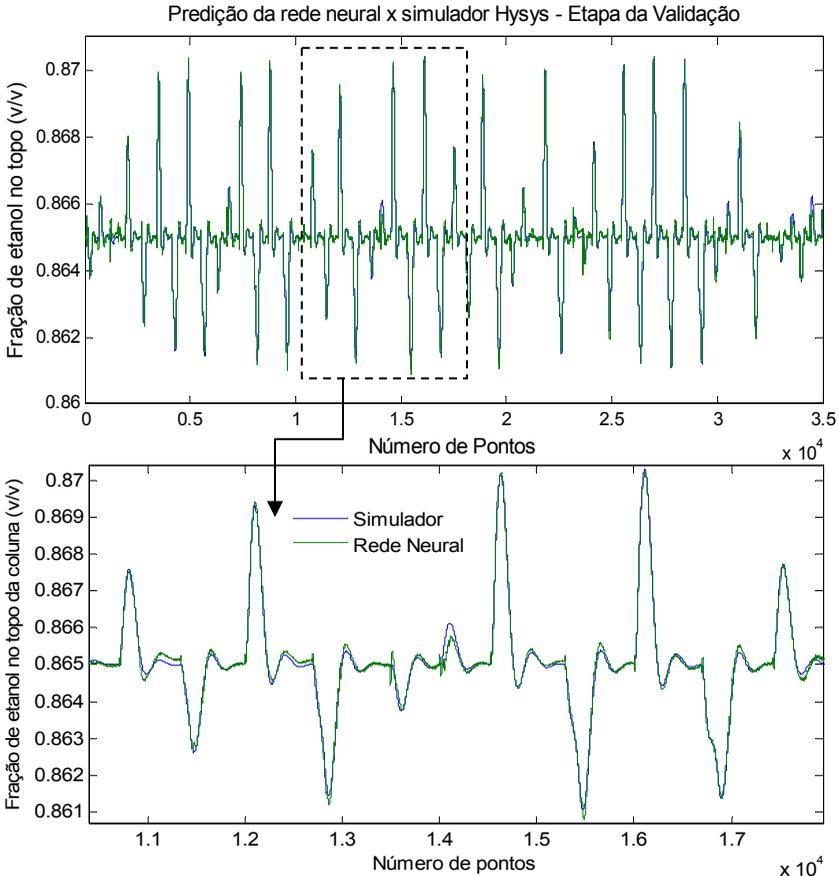


Figura 4.57. (a) Comparação entre a saída do simulador e a predição da rede recorrente selecionada (*Elman-BR-tansig-purelin* [15 15]), na etapa da validação; (b) ampliação de uma seção.

Uma ilustração da arquitetura final da rede que obteve o melhor desempenho nesta etapa, para o conjunto de dados da validação, pode ser observada na Figura 4.58.

Com relação à quantidade de unidades de contexto que a rede de Elman estabelece, conforme visto anteriormente, esta depende diretamente da quantidade de camadas intermediárias e do número de neurônios contidos em cada camada. Outro fator determinante é a decisão de fazer a realimentação de apenas algumas camadas ou de todas elas. Neste trabalho a decisão foi de fazer a realimentação de todas as camadas intermediárias, pois se entende que a rede ganha em termos

de armazenamento de informações que podem ser usadas para melhor prever o comportamento do sistema. Em relação à rede de Elman selecionada, ela possui 2 camadas intermediárias, com 15 neurônios em cada camada, conforme Figura 4.58. Desta forma, temos uma camada com 30 unidades de contexto contendo informações atrasadas no tempo em um passo, que constituem a memória da rede e complementam as 10 entradas dinâmicas da rede.

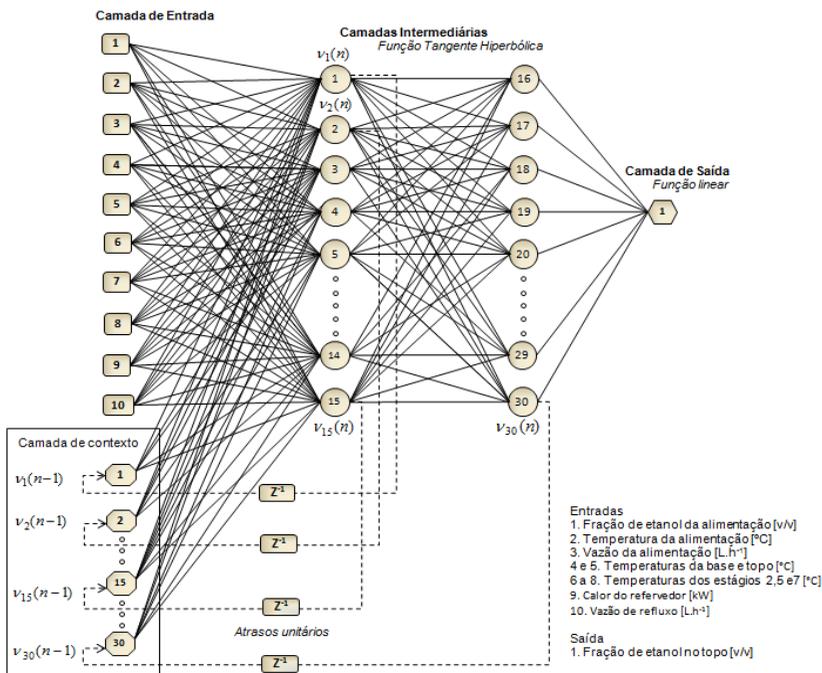


Figura 4.58. Ilustração da arquitetura final da rede recorrente de Elman selecionada para ser empregada como sensor virtual no monitoramento da composição de topo na unidade piloto de destilação.

A coluna simulada no software HYSYS<sup>®</sup> em malha fechada, com a implementação da configuração de controle distribuído caracterizando um processo diferenciado em relação às configurações convencionais, foi validada com sucesso com dados experimentais da planta piloto, representando de forma coerente o comportamento do processo. Desta forma foi possível a obtenção do conjunto de dados dinâmicos representativos da coluna operando em malha fechada.

Na Figura 4.59 é apresentado um diagrama esquemático da coluna de destilação com a proposta do analisador virtual para monitoramento da composição de etanol no topo. A nomenclatura das *tag's* referentes às malhas de controle implementadas na unidade piloto de destilação, ilustradas na Figura 4.59, são apresentadas no *Capítulo III, seção 3.2.2, Tabela 3.1*. As linhas em azul representam os sinais das variáveis de entrada do estimador com modelo do processo baseado em rede neural artificial.

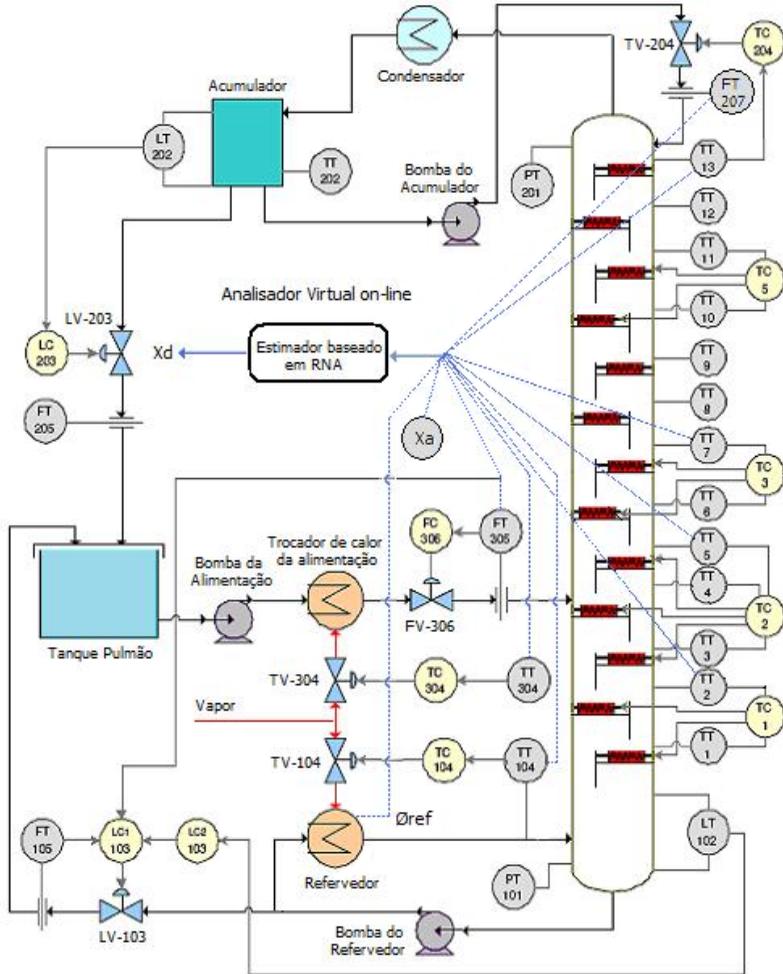


Figura 4.59. Diagrama esquemático da coluna de destilação com a proposta do analisador virtual para monitoramento da composição de etanol no topo.

Vale destacar que, para a implementação deste sistema, torna-se necessário a aquisição do medidor de vazão *fieldbus* para a corrente de refluxo, ilustrado na Figura pela *tag* FT-207. Outras adaptações na unidade e em relação aos softwares empregados também se fazem necessário e estão detalhadas no final deste documento como sugestões para trabalhos futuros.

#### 4.4.2 Avaliação final em relação ao modelo empírico obtido para ser empregado como analisador virtual em malha fechada

A metodologia proposta foi utilizada com sucesso na obtenção de um analisador virtual para a estimação da medida de composição do produto de topo de uma coluna de destilação binária. Os resultados apresentados indicam que o estimador de composição é consistente, com boa capacidade de generalização. Com o modelo empírico desenvolvido em malha fechada, pode-se optar por manter as mesmas malhas de controle e implementar o modelo apenas para fazer o monitoramento do produto de topo da unidade, ou seja, o sensor virtual funcionaria como um analisador de processo em linha, com a vantagem do baixo custo operacional e de manutenção.



# Capítulo V

## 5.1 CONCLUSÕES

Este trabalho apresentou o desenvolvimento de uma modelagem empírica para uma unidade piloto de destilação com características especiais, empregando redes neurais artificiais. As redes neurais foram escolhidas para tal, em virtude das vantagens que apresentam em termos de simplicidade e velocidade de predição, quando comparadas com modelos fenomenológicos e ainda pelo fato de necessitarem apenas da aquisição de um conjunto de dados representativo para modelar o processo. Com isso, se tornam ferramentas poderosas em vários setores da indústria de processos, inclusive a de petróleo e gás.

O desenvolvimento do modelo do processo com configuração de controle com ação distribuída foi realizado em ambiente HYSYS<sup>®</sup> Dinâmico. Apresentou-se como um desafio, visto a complexidade e acoplamentos, tanto do processo, quanto do software utilizado. Porém, todo esforço demandado nesta etapa foi recompensado com os bons resultados obtidos com o emprego deste simulador, que tornaram consistente a base de dados obtida, dando confiabilidade às demais etapas do projeto. Desta forma, pode-se considerar que a escolha do software foi adequada.

A validação das simulações foi realizada com sucesso. Para as simulações estáticas, onde se realizou um estudo específico, as respostas obtidas pelo software, quando submetidos a variações na carga, foram coerentes e seguiram o comportamento previsto do sistema etanol-água nas condições avaliadas. Serviram para comprovar a eficácia do simulador em reproduzir o processo e seus efeitos quando sujeito a perturbações.

Nas simulações dinâmicas utilizaram-se os dados experimentais da planta piloto, no regime transiente, para realizar a validação. Para tal, foram utilizados os perfis de temperaturas experimentais da base, topo e prato 2 da coluna, além de avaliar o comportamento das variáveis manipuladas para cada malha de controle. Os resultados novamente foram coerentes, seguiram as tendências dos perfis experimentais para as duas perturbações na carga, indicando a acertada escolha do simulador.

Definiu-se por utilizar o ambiente MATLAB<sup>®</sup> para desenvolver as redes neurais. Através do software escolhido foi possível avaliar os diversos métodos e algoritmos de treinamento e aprendizado de forma adequada. Foi feito uso deste por ser considerado um ambiente de computação numérica de alto desempenho e que interage diretamente com o ambiente MS-Windows.

Dentre as opções dos modelos de redes disponíveis, foram escolhidas para este estudo redes *feedforward* do tipo *Multi-Layer Perceptron* (MPL) e redes recorrentes do tipo Elman. Esta seleção teve como intuito permitir a comparação de uma rede considerada padrão para problemas de identificação de sistemas, a MPL, com uma rede muito utilizada em processos dinâmicos, a rede recorrente de Elman.

Na etapa inicial deste trabalho foram avaliadas apenas as redes MPL. Pela análise de seus resultados conclui-se que quando utilizado o algoritmo de treinamento *Levenberg-Marquardt* – LM, para o conjunto de dados em que a rede foi treinada, a função *logarítmica* aplicada nas camadas intermediárias e na camada de saída foi a que mais se adaptou aos dados e que obteve o melhor desempenho comparativamente com as combinações das funções *tangente hiperbólica* e *linear*. Desta forma, esta função foi escolhida para ser empregada na etapa seguinte, na avaliação de algoritmos de treinamento e métodos de otimização.

Com o estudo de diversos algoritmos de treinamento, além de encontrar o algoritmo mais adaptado ao conjunto de dados disponíveis, houve a contribuição científica para as pesquisas na área. A escolha do algoritmo mais adequado não é trivial e depende da complexidade e da natureza do problema, do tamanho do conjunto de dados disponível para treinamento, do número de parâmetros ajustáveis da rede e do erro mínimo desejado. Em função dos resultados obtidos chega-se a conclusão que os algoritmos de treinamento *Levenberg-Marquardt* (LM) e *Regularização de Bayesian* (BR) que empregam o *Método quasi-Newton* se destacam dentre os algoritmos avaliados, resultando em treinamentos com erros médios quadráticos muito inferiores aos demais.

O algoritmo *Regularização de Bayesian* não necessita obrigatoriamente de um conjunto de teste, pois usa uma função de treinamento modificada para melhorar seu desempenho, onde a superfície de erro é suavizada, aumentando a generalização. Com o emprego destes foi gerado redes com melhor capacidade de generalização em relação aos demais algoritmos avaliados. Sendo assim, recomenda-se seu uso quando o universo do conjunto de dados não for demasiadamente grande, pois neste caso, a carga computacional e o tempo de treinamento necessários, podem inviabilizar seu emprego. No

estudo específico deste trabalho, o conjunto de dados pode ser considerado de dimensão moderada, e a questão do uso da memória e processamento não foi uma questão limitante observada para emprego destes algoritmos.

Dando ênfase a questão de recurso computacional, quando não se possui os recursos adequados para tal, e for admissível pelo sistema que a precisão da estimação não seja tão rigorosa, o emprego do algoritmo *Resilient Backpropagation* (RP) se torna uma opção relevante e aconselhável, com base nos resultados deste trabalho. Esta indicação é feita em função deste ser considerado um algoritmo simples com rápida convergência e pouca memória requisitada. Neste trabalho, seu desempenho foi considerado muito bom frente ao conjunto de dados específicos, quando comparado aos demais algoritmos avaliados.

Resultados pouco satisfatórios, comparativamente, ficaram com os algoritmos *Taxa de Aprendizado Adaptativa* (GDX), *One Step Secant Method* (OSS) e *Gradiente Conjugado em Escala* (SCG). Deve-se avaliar o emprego destes algoritmos para cada caso específico em que são recomendados. De forma geral, o algoritmo BR obteve respostas que apresentaram o menor erro médio quadrático para o conjunto de dados.

Na etapa onde foram avaliadas redes MPL e redes recorrentes, ambas com o algoritmo de treinamento BR, concluiu-se que a rede recorrente de Elman tem capacidade superior de predição para o conjunto de dados os quais lhe foram apresentadas, em comparação com as redes MPL, as quais também obtiveram bons resultados. Esta característica das redes de Elman é atribuída à realimentação interna.

A rede de Elman selecionada, treinada com dados do simulador desenvolvido em malha aberta, contendo 3 camadas intermediárias, com oito neurônios em cada camada que usa a função de transferência *tangente sigmóide* nas camadas intermediárias e com saída *linear*, com o algoritmo de treinamento BR, é a rede mais adequada para o conjunto de dados usados na validação. Reproduziu de forma fiel e com boa acurácia o comportamento dinâmico do processo, mesmo este sendo um sistema acoplado e não-linear. Por estes motivos foi considerada um excelente modelo empírico dinâmico, sendo selecionada para ser aplicada como sensor virtual no processo de destilação estudado. Desta forma, poderá ser utilizada como preditor da variável de saída composição volumétrica de etanol no topo da coluna para fins de otimização e controle no processo de destilação em questão.

Em relação à arquitetura da rede neural selecionada, ao avaliar-se os procedimentos que foram aplicados até sua determinação, com as variações entre algoritmos de treinamento, número de camadas, número

de neurônios e combinações das funções de transferências, conclui-se pelos resultados obtidos que o nível de complexidade foi adequado, e que várias redes foram encontradas com real possibilidade de executar a função de modelar o processo, algumas com maior precisão da estimação da propriedade foco e outras menor precisão, decorrentes de suas distintas capacidades de generalização.

Verificou-se também que redes com arquiteturas muito robustas (elevado número de camadas e neurônios) acabam perdendo a capacidade de generalização, mesmo obtendo um bom resultado no treinamento. Este fato ocorre porque a rede acaba se especializando no conjunto de dados, memorizam os exemplos de treinamento, mas não aprendem a generalizar em situações novas e ocorre o sobre-treinamento ou *overfitting*, devido ao grande número de parâmetros ajustáveis. Quanto maior for a RNA usada, mais complexas são as funções que ela pode criar. Logo, recomenda-se buscar redes com tamanhos moderados que forneçam o erro mínimo admissível para o processo, adotando-se como critério de seleção o desempenho na etapa de validação.

A justificativa pelo grande número de redes neurais e arquiteturas avaliadas se deve ao fato de que, na implementação do sensor na planta o desempenho destas pode variar, em função disto, redes distintas podem ser consideradas.

Do ponto de vista industrial, dependendo da aplicação e, em casos que a precisão da estimação não seja tão rigorosa, as redes com arquiteturas mais básicas, como é o caso da MPL, com apenas uma camada intermediária e poucos neurônios podem ser empregadas, desde que possuam capacidade de mapear os dados de entrada/saída de forma satisfatória. Esta análise também é válida em relação aos algoritmos de treinamentos. Mesmos os que obtiveram resultados com erros médios quadráticos superiores, também podem ser considerados aptos a serem empregados em redes para aplicação industrial, visto que a grandeza do erro de estimação é aceitável para grande parte dos processos. Desta forma, o algoritmo *Resilient Backpropagation* (RP) se destacou, tornando-se uma opção relevante e aconselhada para uso industrial.

Pode-se concluir, portanto que o objetivo de desenvolver um sensor por software com capacidade de estimar a composição futura do produto de topo de uma coluna de destilação simulada para fins de controle, foi atingido. Este modelo empírico neural foi capaz de representar o comportamento da unidade com o controle distribuído. O bom treinamento que a rede obteve e a capacidade de predição em malha aberta tornam possível que esta seja empregada como um simulador para treinamento, ou que sua predição seja disponibilizada *on-line* para

promover projetos de controladores não lineares, auxiliando na otimização da operação na unidade piloto de destilação.

Finalmente, o recurso computacional exigido pelos estimadores é muito baixo. Para validar o conjunto de dados da validação cruzada, que possui um conjunto de pontos relativamente grande (16.610 pontos - equivalente a aproximadamente 12 perturbações) de modo geral, as redes neurais necessitaram de um tempo inferior a 30 segundos, o que os torna atraentes para uso da estimação *on-line*.

Acredita-se que este sensor já tenha capacidade e possa ser utilizado para estimação na planta piloto, tendo sido treinado apenas com dados oriundos de um simulador, pelos motivos relatados anteriormente. Porém, com o transcorrer do tempo, poderá haver disponibilidades de dados referentes a análises cromatográficas nos quais se pode criar um histórico destes, durante experimentos em situações conhecidas nas quais se coletam dados confiáveis do processo que poderão ser utilizados para melhorar o desempenho do sensor. Além disto, atualmente existem técnicas de treinamentos parciais e retreinamento de redes (ZANATA, 2005) que poderão ser empregados com o sensor já implementado e em operação, caso o desempenho do sensor esteja abaixo do esperado. Desta forma, o sensor aprenderá padrões ligeiramente diferentes daqueles com que foi treinado, no sentido de ajustá-lo com maior precisão frente às condições reais da coluna de destilação piloto.

Comumente, as unidades de destilação empregam técnicas de análise *off-line* com emprego de cromatógrafos, e fazem o ajuste da planta e controladores desta forma, embora com significativo atraso e podendo gerar produtos fora da especificação durante este período. De forma análoga, seriam interessantes e extremamente relevantes para futuras etapas tais medidas para aferir o resultado do sensor de tempos em tempos, como uma espécie de realimentação para correção do valor, após implementá-lo na unidade piloto.

Com a implementação futura do sensor na planta, possibilitando a predição da composição com o auxílio do controle distribuído, tem-se a possibilidade real de incrementar os ganhos já comprovados desta nova concepção de distribuição de calor. Busca-se minimizar ainda mais os transientes de operação e, conseqüentemente melhorar o consumo energético da unidade, decorrente de ações mais rápidas, diretas e localizadas sobre cada estágio, diminuindo o tempo da ação corretiva.

A real implementação de estratégias avançadas de controle em ambientes industriais, mesmo que vantajosa, ainda é lenta e requer uma série de avaliações prévias e um planejamento adequado para sua

implantação no processo. Assim sendo, existe grande relevância do estudo destes modelos empíricos com processos simulados, de modo a buscarem espaço dentro dos processos industriais.

Em relação ao analisador virtual para ser empregado na coluna operando em malha fechada com o atual sistema de controle, o mesmo foi desenvolvido a partir do modelo construído também com o emprego do software HYSYS<sup>®</sup> Dinâmico. Novamente o software foi usado com sucesso para obtenção de dados usados no treinamento de redes neurais.

O objetivo desta etapa foi encontrar uma rede neural capaz de prever a composição de topo do processo em malha fechada, levando em consideração todo acoplamento e grande número de fatores/variáveis tanto do processo quanto do sistema de controle da unidade. Com o processo em malha fechada, outros fatores influenciam a composição de topo, e desta forma, foram testadas redes com um número maior de variáveis de entrada.

Novamente testaram-se redes MPL e recorrentes, sendo que o desempenho dos dois tipos de redes foi muito similar, a ponto de ambas possuírem capacidade de serem empregadas como modelo do sensor. Porém, em virtude da vantagem das redes recorrentes frente às redes *feedforward*, de incorporarem a dinâmica do processo dentro da rede através da realimentação interna, a rede recorrente de Elman foi novamente a escolhida. Dentre as redes de Elman avaliadas, a que obteve a melhor resposta para o conjunto de dados de validação, com o menor erro médio quadrático, foi a rede recorrente de 2 camadas intermediárias, com 15 neurônios em cada, que foi treinada com 10 entradas e com algoritmo *Regularização de Bayesian* (BR). Optou-se por configurar a rede recorrente com a realimentação de todas as camadas intermediárias, pois se entende que, desta forma, a rede ganha em termos de armazenamento de informações, as quais podem ser usadas para melhor prever o comportamento do sistema.

O bom treinamento obtido pela rede selecionada e a boa capacidade de generalização e conseqüente predição em malha fechada, tornam possível que esta seja empregada com sucesso como modelo do analisador virtual. Desta forma, pode-se realizar o monitoramento *on-line* da composição de produto de todo na unidade de destilação piloto, mantendo a atual configuração das malhas de controle. A grande vantagem de se usar um analisador virtual é o baixo custo operacional e de manutenção, quando comparados aos analisadores de processo *on-line*.

# Capítulo VI

## 6.1. SUGESTÕES PARA TRABALHOS FUTUROS

De posse destes modelos empíricos desenvolvidos para a unidade piloto de destilação, o primeiro, em malha aberta, verifica-se a possibilidade de incremento da proposta com ação distribuída associada implementação de técnicas mais complexas de controle avançado multivariável e inferencial.

A segunda opção, com o modelo desenvolvido em malha fechada, pode-se optar por manter as mesmas malhas de controle e implementar o modelo apenas para fazer o monitoramento do produto de topo da unidade, ou seja, o sensor virtual funcionaria como um analisador em linha, com a grande vantagem de baixo custo operacional e de manutenção.

Estes estudos constituem uma grande contribuição à modelagem de processos e sua implementação em unidades pilotos geram pesquisas que impactam no meio industrial, ganhando respaldo e abrindo possibilidades para serem implementados em ambientes industriais.

Outra questão relevante a ser mencionada é o fato de em processos reais, a coleta de dados apresentar ruídos inerentes aos sensores, os quais podem ocasionar um aumento do erro da predição da rede. Desta forma, uma sugestão para tornar o modelo simulado em ambiente HYSYS<sup>®</sup> dinâmico mais realístico, seria a inclusão de um termo para gerar ruído nos dados simulados, aproximando-os do comportamento real do processo.

No caso da implementação futura do sensor na unidade para fins de controle, talvez seja interessante a aplicação de filtros antes da entrada do sensor virtual para amenizar as flutuações comentadas anteriormente.

No transcorrer deste estudo foram identificados alguns pontos em relação à unidade experimental que necessitam de atenção e podem ser melhorados, visando a implementação futura do sensor virtual de composição. Estes serão descritos a seguir:

1. Aquisição de um sensor de vazão *fieldbus* para ser instalado na linha de produto de fundo para medição da vazão de líquido que entra no refeedor, possibilitando a melhor caracterização da

corrente e, conseqüentemente, um balanço energético mais preciso, além de garantir o balanço de massa;

2. Aquisição de um sensor de vazão *fieldbus* para ser instalado na linha de refluxo, possibilitando a medição precisa desta vazão (a qual é fundamental também no cálculo do calor retirado no topo), permitindo que esta informação seja disponibilizada também ao sensor virtual em malha fechada.

Além das modificações na unidade experimental, pode-se aperfeiçoar as técnicas computacionais aplicadas atualmente, tais como:

1. Realização de uma otimização nas variáveis de entrada do sensor com o propósito de determinar os pontos ótimos de cada variável (*set point*) para distintas condições de produto de topo requerido. Esta etapa é fundamental na implementação do sensor desenvolvido e para aplicação de técnicas de controle mais avançadas na unidade em questão;
2. Uma abordagem de controle que também se apresenta como uma alternativa para ser empregada na futura implementação do sensor na unidade piloto se refere ao sistema com lógica *fuzzy*, o qual poderia ser empregado acoplado ao sistema de inferência, responsável por estimar a composição de topo da coluna. Desta forma, o sistema *fuzzy* poderia ser ajustado de maneira a gerar dinamicamente os *set points* (conforme otimização proposta no item anterior) para os controladores PID's de temperaturas, relacionados com as variáveis de entrada da rede.
3. Após a inclusão dos sensores adequados na corrente de produto de fundo, desenvolver e aprimorar o balanço energético realizado por Werle (2007) com modificação do programa de cálculo de energia, levando em consideração o fato da mistura não se comportar idealmente. Realizar a implementação do modelo energético com interface com software supervisor. Desta forma, disponibilizar os valores calculados de carga térmica na base, de acordo com as condições operacionais, fornecendo o valor atualizado desta variável para o sensor virtual. Este tópico é necessário, caso o sensor seja implementado em malha fechada, no qual o calor da base constitui uma das entradas do sensor virtual.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABDULLAH, Z.; NORASHID, A.; ZAINAL, A. Nonlinear modelling application in distillation column, **Chemical Product and Process Modeling**. v.2, ed.12, 2007. DOI: 10.2202/1934-2659.1082.

ABDUL-WAHAB, S.; ELKAMEL, A.; MADHURANTHAKAM, C. R.; AL-OTAIBI, M.B. Building inferential estimators for modeling product quality in a crude oil desalting and dehydration process. **Chemical Engineering and Processing**, v.45, p.568-577, 2006.

ADETOLA, V.; GUAY, M. Integration of real time optimization and model predictive control. **Journal of Process Control**, v.20, p.125–133, 2010.

ANSARI, R. M.; TADÉ, M. O. **Nonlinear Model-Based Process Control: Application in Petroleum Refining**, Springer, London, 2000.

ANASTASIADIS, A.D.G.D.; MAGOULAS; M. N; RAHATIS, V. New globally convergent training scheme based on the resilient propagation algorithm. **Neurocomputing**. v.64, p. 253-270, 2004.

ARDALANI-FARSA, M.; ZOLFAGHARI, S. Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks. **Neurocomputing**. v.73, p.2540-2553, 2010.

ARAUJO JUNIOR, J. M. **Técnicas de Inteligência Artificial para a geração Dinâmica de Set Points para uma Coluna de Destilação**. (Dissertação de mestrado). 65 p. Universidade Federal do Rio Grande do Norte – Programa de Pós-Graduação em Engenharia Elétrica, 2007.

BAHAR, A.; ÖZGEN, C. State estimation for a reactive batch distillation column, In: **Proceedings of the 17<sup>th</sup> World Congress The International Federation of Automatic Control**, Seoul, Korea, July p.6-11, 2008.

BARATTI, R.; BERTUCCO, A.; ROLD, A.D.; MORBIDELLI, M. Development of a composition estimator for binary distillation columns. Application to a Pilot Plant. **Chemical Engineering Science** v.50 (10), p.1541–1550, 1995.

BATTITI, R. First- and second-order methods for learning: between steepest descent and Newton's method. **Neural Computing**, v.4, p.141–166, 1992.

BAI, S. Dynamic data reconciliation: Alternative to the Kalman filter. **Journal of Process Control**, v.16, 2006.

BARBOSA, C.H. Inference of distillation column products quality using Bayesian networks. **Neural networks, IJCNN'02**. Proceedings of the International Joint Conference. p. 86-91, 2002.

BAROLO, M.; PISTILLO, A.; TROTTA, A., Issues in the development of a composition estimator for a middle vessel batch column. In: Biegler, L. T. **Preprints IFAC Sympos, ADCHEM**, Pisa (Italy), June 14–16, p. 959-964, 2000.

BAROLO, M.; BERTO, F. Composition control in batch distillation: Binary and multicomponent mixtures. **Industrial Engineering Chemical Research**, v.37, p.4689-4698, 1998.

BEALE, E. M. L. **A derivation of conjugate gradients**. In: Lootsma, F.A. (Ed.), Numerical Methods for Nonlinear Optimization. Academic Press, London, 1972.

BISHOP, C. M. **Neural Networks for Pattern Recognition**. Oxford. Clarendon Press., 1995.

BORGES, M. Analisadores de processo. WNAF 2004 – **Workshop nacional sobre analisadores de processos em linha**, Salvador, Bahia: UFBA, p.37, 2004.

BOLF, N.; IVANDIC, M.; GALINEC, G. Softsensors for diesel fuel Property estimation. In: Petton, R.J., Maquin, D. **16<sup>th</sup> Mediterranean Conference on Control and Automation**. IEEE, p.1804, 2008.

BJÖRN, I.N.; GRÉN, U.; SVENSON, F. Simulation and experimental study of intermediate heat exchange in a sieve tray distillation column. **Computers & Chemical Engineering**, v. 26, p. 499-505, 2002.

BOSTON, J. F.; SULLIVAN, JR, S. L. An improved algorithm for solving the mass balance equations in multistage separation processes. **Canadian Journal of Chemical Engineering**, v.50, p.663-671, 1972.

BROWN, R. G.; HWANG, P. Y.C. **Introduction to random signals and applied Kalman filtering**. 2<sup>ed</sup>. Jonh Wiley and Sons, 1992.

CANETE, F. J.; SAZ-OROZCO, P.D.; GONZALEZ-PEREZ, S. Application of adaptive neurofuzzy control using soft sensors to continuous distillation, In: **18<sup>th</sup> European Symposium on Computer Applied Process Engineering-ESCAPE18**, Elsevier, Lyon, France, 2008.

CARCANO, E. C.; BARTOLINI, P.; MUSELLI, M.; PIRODDI, L. Jordan recurrent neural network versus IHACRES in modelling daily streamflows. **Journal of Hydrology**, v. 362, p. 291- 307, 2008.

CHENG, W.; HUANG, J.; LIOU, C. Segmentation of DNA using simple recurrent neural network. **Knowledge-Based Systems. In Press, Corrected Proof**, 2011.

CHATTERJEE, T.; SARAF, D.N. On-line estimation of product properties for crude distillation units, **Journal Process Control**, v.14, p.61-77, 2004.

CHERIF, A.; CARDOT, H.; BONÉ, R. SOM time series clustering and prediction with recurrent neural networks. **Neurocomputing**, v.74, p.1936-1944, 2011.

CLAUMANN, C. A. **Desenvolvimento e aplicações de redes neurais wavelets e da teoria de regularização na modelagem de processos**. 167 f. Tese (Doutorado em Engenharia Química) – Departamento de Engenharia Química e Alimentos, Universidade Federal de Santa Catarina, Florianópolis, 2003.

CONSTANTINO, S. F. **Foundation Fieldbus**. Departamento de Engenharia Eletrônica. UFMG. Disponível em: <http://www.cpdee.ufmg.br/~seixas/PaginaSDA/Download/DownloadFiles/FF.pdf>. Acesso em: Novembro, 2010.

CONZ, V. **Desenvolvimento de analisadores virtuais aplicados a colunas de destilação industriais**. Dissertação (mestrado) Programa de Pós-Graduação em Engenharia Química, Departamento de Engenharia Química, Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

DAI, X.; WANG, W.; DING, Y.; SUN, Z. Assumed inherent sensor - inversion based ANN dynamic soft-sensing method and its application in erythromycin fermentation process. **Computers and Chemical Engineering**. v.30, p. 1203– 1225, 2006.

DAI, C.; WEIRONG, CHEN, W.; ZHU, Y.; JIANG, Z.; YOU, Z. Seeker optimization algorithm for tuning the structure and parameters of neural networks. **Neurocomputing**, v.74, p. 876-883, 2011.

DAMODARAN, P.; VÉLEZ-GALLEGO, M. C. A simulated annealing algorithm to minimize makespan of parallel batch processing machines with unequal job ready times. **Expert Systems with Applications**. v.39, p. 1451-1458, 2012.

DENNIS, J. E., SCHNABEL, R.B. **Numerical methods for unconstrained optimization and nonlinear equations**. Prentice-Hall, Edgewood Cliffs , 1983.

DE KOEIJER, G.; KJELSTRUP, S. Minimizing entropy production rate in binary tray Diabatic distillation. **Ind. Engineering Chemical Research**. v.41, p. 5826-5834, 2002.

DE RIJKE, A. **Development of a concentric internally heat integrated distillation column**. Tese (Doutorado em Engenharia Química). 191 p. Universidade Tecnológica de Delft, Holanda, 2007.

DESAI, K.; BADHE, Y; TAMBE, S. S.; KULKARNI, B. D. Soft-sensor development for fed-batch bioreactors using support vector regression. **Biochemical Eng. Journal**, v. 27, p. 225–239, 2006.

DUFOUR, P., BHARTIYA, S., DHURJATI, P. S., DOYLE, F. J. Neural network-based software sensor: training set design and application to a continuous pulp digester. **Control Engineering Practice**, v. 13, p. 135–143, 2005.

ENWEREMADU, C.; WAHEED, A; OJEDIRAN, J. Parametric study of an ethanol–water distillation column with direct vapour recompression heat pump. **Energy for Sustainable Development**. v 13, p. 96–105, 2009.

EVANGELISTA, M. A.; JUNIOR, F. N.; ARRUDA, L.V.R. Development of inferencial distillation models using multivariable statistical methods, **Proceedings of the 40<sup>th</sup> IEEE Conference on Decision and Control** 4. p.3722-3727, Orlando, Flórida, 2001.

ELMAN, J. L. Finding structure in time. **Cognitive Science**. v.14, p.179-211, 1990.

EWING, G. W. **Métodos Instrumentais de Análise Química**. Editora Edgar. Blücher - São Paulo, 1972.

FACCO, P.; DOPLICHER, F.; BEZZO, F.; BAROLO, M. Moving average PLS soft sensor for online product quality estimation in an industrial batch polymerization process. **Journal of Process Control**, v. 19, p. 520–529, 2008.

FIEG, G. Composition control of distillation columns with a sidestream by using gas chromatographs. **Chemical Engineering and Processing**. v. 41 p.123–133, 2002.

FILETI, A. M. F.; PEDROSA, L.S.; PEREIRA, J. A. F. R. A self tuning controller for multicomponent batch distillation with soft sensor inference based on a neural network. **Computers & Chemical Engineering Supplement**, v.23, p.261-264, 1999.

FIKAR, M. Application of iterative dynamic programming to optimal control of a distillation column. **Canadian Journal of Chemical Engineering**. v.76, p.1110-1117, 1998.

FLETCHER, R.; REEVES, C. M. Function mimization by conjugate gradients. **Comput. Journal**. v.7, p.149–154, 1964.

FLETCHER, R. **Practical methods of optimization**, 2<sup>nd</sup> edition, John Wiley & Sons, New York, NY, 1987.

FORESEE, F. D., HAGAN, M. T. Gauss-Newton approximation to Bayesian regularization. **Proceedings of the International Joint Conference on Neural Networks**, 1997.

FORTUNA, L.; GRAZIANI, S.; XIBILIA, M.G. Virtual instruments in refineries. **IEEE Instrumentation & Measurement Magazine**. v.8, p.26–34, October, 2005.

FORTUNA, L.; GRAZIANI, S.; RIZZO, A.; XIBILIA, M.G. Softsensor for a propylene splitter with seasonal variations, soft sensors for monitoring and control of industrial processes. In: **Advances in Industrial Control**. Springer, London, 2007.

GELADI, P.; KOWALSKI, B. R., Partial least-squares regression: a tutorial. **Analytica Chimica Acta**. v.185 p.1, 1986.

GIORDANO, F.; LA ROCCA, M.; PERNA, C. Forecasting Nonlinear Time Series with Neural Network Sieve Bootstrap. **Computational Statistics & Data Analysis**, v. 51, n<sup>o</sup> 8, p. 3871 - 3884, 2007.

GONÇALVES, M. L. **Métodos de classificação não-supervisionada de imagens de sensoriamento remoto usando mapas auto-organizáveis de Kohonen**. Tese de Doutorado. (Doutorado em Engenharia Elétrica e de Computação). Departamento de Engenharia de Computação e Automação Industrial. Universidade Estadual de Campinas, 2009.

GUPTA, S.; RAY, S.; SAMANTA, A.N. Nonlinear control of debutanizer column using profile position observer. **Computers & Chemical Engineering**. v.33, p.1202-1211, 2008.

HAYKIN, S. **Neural Networks**, 2<sup>nd</sup>. Prentice Hall, Upper Saddle River, New Jersey, 1999.

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

HAGAN, M.T., MENHAJ, M.B. Training feedforward networks with the Marquardt algorithm. **IEEE Transactions Neural Networks**. v.5, p.989–993, 1994.

HARTNETT, M. K.; LIGHTBODY, G.; IRWIN, G.W. Dynamic inferential estimation using principal components regression (PCR). **Chemometrics and Intelligent Laboratory Systems**. v.40, p.215, 1998.

HENLEY, E. J.; SEADER, J. D. **Equilibrium-Stage Separation Operation in Chemical Engineering**. John Wiley & Sons, Canada, 1981.

HEBB, D. O. **The organization of behavior: a neuropsychological theory**. Wiley. New York, 1949.

HO, T. J.; HUANG, C.T.; LIN, J.M.; LEE, L.S. Dynamic simulation for internally heat-integrated distillation columns (HIDiC) for propylene propane system. **Computers & Chemical Engineering**. v.33, p.1187–1201, 2009.

HOSEN, M. A.; HUSSAIN, M.A.; MJALLI, F. S. Control of polystyrene batch reactors using neural network based model predictive control (NNMPC): An experimental investigation. **Control Engineering Practice**, v.19, p.454-467, 2011.

HOPFIELD, J. J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities, **Proceedings of the national Academy of Sciences, USA**, v.79, p. 2554- 2558, 1982.

HOLCOMB, T. R.; MORARI, M. PLS/Neural Networks. **Computers & Chemical Engineering**, v.16, n.4, p.393-411, 1992.

HUMPHREY, J. L.; SEIBERT, A. F.; KOORT, R. A. Separation technologies advances and priorities, Final Report for US Department of Energy. **Office of Industrial Technologist**, Washington DC, DE 91-010, 1991.

JANA, K. A. Heat integrated distillation operation. **Applied Energy**. v.87, p.1477-1494, 2010.

JAMES, S. C.; LEGGE, R. L.; BUDMAN, H. On-line estimation in bioreactors: A review. **Reviews in Chemical Engineering**, v.16, n.4, p.311-340, 2000.

JAIN, L. Recurrent neural networks: Design and Applications. Departments of Physics and Computer Science and Information Systems American University Washington, D.C.: CRC Press, 2001.

JAZWINSKI, A. H. **Stochastic processes and filtering theory**. New York and London: Academic Press, 1970.

JESUS, W. M., NEVES JR, F., ARRUDA, L. V. R., FABRO, J. A. Controle de uma coluna de destilação no *software* HYSYS através de algoritmos em C++. **VIII Seminário de Iniciação Científica e Tecnológica**. Curitiba Anais do seminário, 4 f., 2003.

JIANXUL, L.; HIUHE, S. Soft Sensing Modeling using neuro fuzzy based on Rought set Theory. **Proceeding of the American Control Conference**, Anchorage: USA. v.1, p.543-548, 2002.

JOSEPH, B.; BROSILOW, C. B., Inferential control of process. Part I: Steady state analysis and design. Part 2: The structure and dynamics of inferential control systems. Part 3: Construction of suboptimal dynamic estimators, **AIChE Journal**. v.24, n.3, p.485-509, 1978.

JORDAN, M. I., Attractor dynamics and parallelism in a connectionist sequential machine. In **Proceedings of the Eight Annual Conference of the Cognitive Science Society**. p.531-546, 1986.

KALMAN, R. E. A new approach to linear Filtering and Prediction Problems. Transaction of the ASME – **Journal of Basic Engineering**. p. 35-45, March, 1960.

KANO, M., MIYAZAKI, K., HASEBE, S., HASHIMOTO, I., Inferential control system of distillation compositions using dynamic partial least squares regression. **Journal of Process Control**. v.10, p.157-166, 2000.

KANO, M., SHOWCHAIYA, N., HASEBE, S., HASIMOTO, I. Inferential control of distillation composition: selection of model and control configuration. **Control Engineering Practice**. v.11, p.927–933, 2003.

KALID, R. A. **Apostila do Curso de Controle de Processos**. Publicação do Departamento de Engenharia Química da UFBA, Salvador, Bahia, Brasil, 1999. Disponível em: <http://www.lacoi.ufba.br>.

KANSHA, Y.; TSURU, N.; FUSHIMI, C.; TSUTSUMI, A. Integrated process module for distillation processes based on self-heat recuperation technology. **Journal Chemical Engineering Japan**. v. 43(6), p.502-507, 2010.

KERMANI, B.G.; SCHIFFMAN, S.S.; TROY NAGLE, H., Performance of the Levenberg–Marquardt neural network training method in electronic nose applications. **Sensors and Actuators B: Chemical** v.110, p.13–22, 2005.

KHAZRAEE, S. M.; JAHANMIRI, A. H. D. Composition estimation of reactive batch distillation by using adaptive neuro-fuzzy inference system. **Product engineering and chemical technology, Chinese Journal of Chemical Engineering**, v.18 (4) p.703-710, 2010.

KIM, C.; FRIEDLY, C. F. Dynamic modeling of large staged systems. **Industrial and Engineering Chemistry Process Design and Development**. v.13, p. 177-188, 1974.

KIM, Y. H.; KIM, J. S. Constrained nonlinear control of a binary distillation column. **Journal of Chemical Engineering of Japan**, v. 31, p.417-424, 1998.

KOEIJER, G.; RØSJORDE, A., KJELSTRUP, S. Distribution of heat exchange in optimum diabatic distillation columns. **Energy**, v. 29, p. 2425–2440, 2004.

KOEIJER, G.; RIVERO, R. Entropy production and exergy loss in experimental distillation columns. **Chemical Engineering Science**. v. 58, p.1587 - 1597, 2003.

KOHONEN, T. Clustering, taxonomy and topological maps of patterns. In: Proceedings of the 6<sup>th</sup> International Conference on Patterns Recognition (6ICPR). Washington, DC: **IEEE Computer Soc. Press.**, p. 114-128, 1982.

KUMAR, U. A.; PALIWAL, M. Neural networks and statistical techniques: A review of applications. **Expert Systems with Applications**. v. 36, p. 2–17, 2009.

KRESTA, J. V.; MARLIN, T.E.; MACGREGOR, J. F. Development of inferential process models using PLS. **Computers & Chemical Engineering**, v.18 p.597-611, 1994.

LANGERHORST, M. **Modelagem matemática e simulação dinâmica de colunas de fracionamento de aromáticos com integração energética**. Dissertação (Mestrado em Engenharia Química). Universidade Federal da Bahia - Escola Politécnica – UFBA, 2000.

LEE, Y. H.; HONG, S. H. Dependency on prioritized data in the delay analysis of foundation fieldbus. **Control Engineering Practice**. v.18, p. 845–851, 2011.

LINHARES, L. L. S, ARAÚJO JUNIOR, J. M., ARAÚJO, F. M. U. Redes neurais artificiais para identificação da fração molar de pentano na composição do GLP. VII SBAI – **Congresso Brasileiro de Automação Inteligente**, Florianópolis, 2007.

LINHARES, L. L. S. **Sistema híbrido de inferência baseado em análise de componentes principais e redes neurais artificiais aplicado a plantas de processamento de gás natural**. Dissertação, 119 p. (Mestrado em Engenharia Elétrica). Universidade Federal do Rio Grande do Norte - UFRN, 2010.

LONGUI, L.G.S.; DIEHL, F.C.; LUSA, L.P.; SECCHI, A.R.; MUNIZ, L.A.R. Simulação operacional de uma torre de destilação atmosférica via Aspen Plus e avaliação de modelos de analisadores virtuais. **Revista Controle & Automação**. v.20 n<sup>o</sup>.3/Julho, Agosto e Setembro, 2009.

LOTUFO, F. A.; GARCIA, C. Sensores Virtuais ou Soft Sensors: Uma Introdução. In: **7<sup>th</sup> Brazilian Conference on Dynamics, Control and Applications**. Presidente Prudente - SP, Brazil, May, 2008.

LU, W.; FISHER, D. G. Least-squares output estimation with multirate sampling. **IEEE Transactions Automatic Control**. v.34, p.669, 1989.

LU, W.; FISHER, D. G. Multirate adaptive inferential estimation. **IEE Proceedings-D, Control Theory and Application**. v.139, p.181, 1992.

LU, N., YANG, Y., GAO, F., WANG, F. Multirate dynamic inferential modeling for multivariable processes. **Chemical Engineering Science**. v.59, p.855–864, 2004.

LUYBEN, W.L. **Process modeling, simulation and control for chemical engineers**. McGraw-Hill International Editions, Singapore, 1989.

MACMURRAY, J. C., HIMMELBLAU, D. M. Modeling and control of a packed distillation column using artificial neural networks, **Computers & Chemical Engineering**. v.19, n.10, p.1088, 1994.

MAHFOULD, M., KANDIAH, S., LINDKENS, D. A. Fuzzy model-based predictive control using an ARX structure with feedforward. **Fuzzy sets and systems**. v.125, p.39-59, 2002.

MARCON, S. M.; TRIERWEILER, J. O. EKF e CEKF: Comparação entre duas formulações do Filtro de Kalman Estendido, **Anais do Congresso Brasileiro de Automação**, 2002.

MAJI, P. Efficient Design of Neural Network Tree using a New Splitting Criterion. **Neurocomputing**, v.71, p. 787- 800, 2007.

MACKAY, D.J.C. Bayesian interpolation. **Neural Computing**. v.4, p.415–447, 1992.

MARTINS, M. A. F.; PESSOA, R. W. S.; KALID, R. A. Análise Estatística do uso de um filtro de Kalman a um sistema de medição com ruídos. In: **V Congresso Brasileiro de Metrologia**, Salvador. Novembro, 2009.

MARANGONI, C. **Implementação de uma estratégia de controle com ação distribuída em uma coluna de destilação**. Tese. (Doutorado em Engenharia Química). Programa de Pós-Graduação do Departamento de Engenharia Química e Alimentos, Universidade Federal de Santa Catarina – UFSC. Florianópolis, 2005.

MATSUDA, K.; KAWAZUISHI, K.; KANSHA, Y.; FUSHIMI, C.; NAGAO, M.; KUNIKIYO, H.; MASUDA, F.; TSUTSUMI, A. Advanced energy saving in distillation process with self-heat recuperation Technology. **Energy**. v.36, p. 4640-4645, 2011.

MENEGUELO, A. P. **Contribuições à análise e modelagem de operações transientes de colunas de destilação**. Tese de Doutorado – Departamento de Engenharia Química e Engenharia de Alimentos da UFSC, julho de 2007.

MELO, J. D. **Aspectos da implementação de redes neurais**, DCA-CT-UFRN, 2004.

MEDEIROS, M. C.; E. PEDREIRA, C. E. What are the effects of forecasting linear time series with neural networks? **Engineering Intelligent Systems**. v.9, p.237-242, 2001.

METZ, F. L. **Dinâmica e mecânica estatística de modelos de redes neurais**. Tese de Doutorado (Doutorado em Física). Departamento de Pós-Graduação em Física – Universidade Federal do Rio Grande do Sul – UFRGS, 2008.

MENDES, E.; CARVALHO, A.; MATIAS, A. Credit Assessment Using Evolutionary MLP Networks. Proceedings of Computational Finance 1997 - **The Fifth International Conference on Neural Networks in the Capital Markets**, December, 1997.

MENDES, E.; CARVALHO, A. Target Recognition using Evolutionary Neural Networks. **Proceedings of V Brazilian Symposium on Neural Networks**, December, 1998.

MENEZES JÚNIOR, J. M. P. **Redes neurais dinâmicas para predição e modelagem não-linear de séries temporais**. Dissertação 132p. (Mestrado em Engenharia Teleinformática). Departamento de Pós-Graduação em Engenharia de Teleinformática da Universidade Federal do Ceará, 2006.

MEJDELL, T.; SKOGESTAD, S. Estimation of distillation compositions from multiple temperature measurements using Partial-Least-Squares Regression. **Industrial Engineering Chemical Research**. v.30, p. 2543-2555, 1991.

MEJDELL, T.; ANDERSSON, B. Using temperature profile for product quality estimation on a distillation column. **ISA Paper**. p.93-254, 1993.

MØLLER, M.F. A scaled conjugate gradient algorithm for fast supervised learning. **Neural Networks**. v.6, p.525–53, 1993.

MOREJON, R. A.; PRINCIPE, J. C. Advanced search algorithms for information-theoretic learning with kernel-based estimators. **IEEE Transactions on Neural Networks**. v.15, p.874 - 884, 2004.

MOURA, L.G. **Modelagem empírica de colunas de destilação utilizando redes neurais de wavelets para otimização e controle de processos**. Dissertação. (Mestrado em Engenharia Química). Engenharia Química da Universidade Federal de Santa Catarina - UFSC, Florianópolis, 2003.

MOHR, T. **Desenvolvimento de um analisador virtual para uma planta de polietileno de alta densidade**. Dissertação. (Mestrado em Engenharia Química). Departamento de Pós-Graduação da Universidade Federal do Rio Grande do Sul –UFRGS, Porto Alegre, 2004.

MORAES, C. A. S. **Modelagem, controle e minimização do consumo de energia de uma unidade de fracionamento de nafta**. Tese. (Doutorado em Engenharia Elétrica). Faculdade de Engenharia Elétrica e Computação - Universidade Estadual de Campinas - UNICAMP, 2004.

MUKHERJEE, I; SRIKANTA R, S. Comparing the performance of neural networks developed by using *Levenberg–Marquardt* and Quasi-Newton with the gradient descent algorithm for modeling a multiple response grinding process. **Expert Systems with Applications**. v.39, p. 2397–2407, 2012.

MURTONIEMI, E.; MERKKU, P.; YLIRUUSI, J. Comparison of four different neural network training algorithms in modelling the fluidized bed granulation process. **Lab. Microcomput.** v.12, p.69–76, 1993.

NASCIMENTO, M. L. **Controle multivariável de colunas de destilação de petróleo.** Dissertação (Mestrado em Eng. Química), Escola Politécnica da Universidade de São Paulo. São Paulo, 130p., 1994.

NAGAI, E. **Identificação automática de modelos fuzzy inferenciais.** Tese (Doutorado em Engenharia Elétrica). Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná, UTFPR, Curitiba, Brasil, 2006.

NOORAI, A.; ROMAGNOLI, J.; FIGUEROA, J. Process, identification, uncertainty characterisation and robustness analysis of a pilot scale distillation column. **Journal of Process Control.** v.9, p.247-264, 1999.

NOCEDAL, J. **Theory of algorithms for unconstrained optimization,** Acta Numerica, v.1, p. 199-242, 1991.

NGUYEN, D.; WIDROW, B. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. **IJCNN: International Joint Conference on Neural Networks,** p.21-26, 1990.

OLIVEIRA, S. B. M.; MARQUES, R. P.; PARISE, J. A. R. Modeling of an ethanol–water distillation column with vapor recompression. **International Journal of Energy Research.** v.25 p.845–858, 2001.

OLIVEIRA, D. L.; PARISOTTO, I. G. B.; MARANGONI, C.; MACHADO, R. A. F. Estudo da Perda de Calor em uma Coluna de Destilação com Controle por Aquecimento Distribuído. **VII Congresso Brasileiro de Engenharia Química em Iniciação Científica,** 2007.

OISIOVICI, R. M., CRUZ, S. L. State estimation of batch distillation columns using a extended Kalman filter. **Chemical Engineering Science.** v.55, p.4667-4680, 2000.

OSIOVICI, R. M.; CRUZ, S. L. Inferential control of high-purity multicomponent batch distillation columns using an extended Kalman filter, **Ind. Engineering Chemical Research**, v.40, p.2628-2639, 2001.

ONDREY, G. Simulation and modeling spread their wings. **Chemical Engineering**, n. 5, p. 27-31, 2005.

OSORIO, D.; J. PÉREZ-CORREA, J. R.; EDUARDO AGOSIN, E.; CABRERA, M. Soft-sensor for on-line estimation of ethanol concentrations in wine stills. **Journal of Food Engineering**, v.87, p.571–577, 2008.

OJA, M.; KASKI, S.; KOHONEN, T. Bibliography os self-organizing map (SOM) papers: 1998-2001 addendum. **Neural Computing Suveys**, v.3, pp 1-156, 2003.

PASETTI, G. **Instrumentação, controle e supervisão de uma coluna de destilação piloto utilizando tecnologia Foundation Fieldbus**. Dissertação. (Mestrado em Engenharia Elétrica). Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina – UFSC. Florianópolis, 2005.

PATIL, S. R; NIGAM, M.J. Soft sensor for multicomponent distillation column using neural network and genetic algorithm based techniques. **International Journal of Computational Intelligence Research**, July. 2009.

PASSOS, G.; COHN, P. E., GORRINI, A. S. **Curso de Instrumentação Analítica**. Apostila do Instituto Brasileiro de Petróleo (IBP), RJ, 3ª Ed, 1988.

PASQUOTTO, J. L. D. **Previsão de séries temporais no varejo brasileiro: investigação comparativa da aplicação de redes neurais recorrentes de Elman**. Dissertação. (Mestrado em Administração). Programa de Pós-Graduação Departamento de Administração e Contabilidade da Universidade de São Paulo, USP, 2010.

PALADE, V.; AO, S.I. Ensemble of Elman neural networks and support vector machines for reverse engineering of gene regulatory networks. **Applied Soft Computing**, v.11, p. 1718-1726, 2011.

PENDHARKAR, P. C. Exhaustive and heuristic search approaches for learning a software defect prediction model . **Engineering Applications of Artificial Intelligence**, v.23, Issue 1, p.34-40, 2010.

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.

PINTÉR, J. D. Calibrating artificial neural networks by global optimization. **Expert Systems with Applications**, v.39, p. 25-32, 2012.

PINO, R.; PARRENO, J.; GOMEZ, A.; PRIORE, P. Forecasting Next-day Price of Electricity in the Spanish Energy Market using Artificial Neural Networks. **Engineering Applications of Artificial Intelligence**, v. 21, n<sup>o</sup>1, p. 53 - 62, 2008.

PINTO, R.; LIMA, E. L. Inferência e controle de variáveis dificilmente mensuráveis. **Seminário de Automação do IBP**, p.146-151, 1993.

PINTO, L. T. **Aplicação de colunas de múltiplo-efeito na destilação do sistema etanol-água. economia de energia.** Dissertação. (Mestrado em Engenharia de Produção). Universidade Federal de Santa Catarina – UFSC, Florianópolis, 1987.

PINCETI, P. Fieldbus: More than a communication link. **IEEE Instrumentation & Measurement Magazine**. v.7(1), p.17 - 23, 2004.

PONCE, G. H. S. F. **Estudo de nova Configuração de Colunas De Destilação aumentando Eficiência Energética.** Dissertação. (Mestrado em Engenharia Química), Universidade Estadual de Campinas, Faculdade de Engenharia Química, 2011.

PORFÍRIO, C.R.; ODLOAK, D. Optimizing model predictive control of an industrial distillation column. **Control Engineering Practice**. v.19, p. 1137-1146, 2011.

POWELL, M.J.D. Restart procedures for the conjugate gradient method. **Mathematical Programming**. v.12, p.241–254, 1997.

PULIDO, J. L. **Estudo de um novo conceito de coluna de destilação: coluna de destilação com integração interna de calor (CDIIC)**. Dissertação (Mestrado em Engenharia Química). Programa de Pós-Graduação em Engenharia Química da Universidade Estadual de Campinas –UNICAMP, 2011.

PLUMB, P. A.; ROWE, R.C.; YORK, P.; BROWN, M. Optimisation 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.

PREMKUMAR, R.; RANGAIAH, G. P. Retrofitting conventional column systems to dividing-Wall Columns. **Chemical Engineering Research and Design**. v.87, p. 47–60, 2009.

QIN, S. J.; McAVOY, T. J. Nonlinear PLS modeling using neural networks. **Computers & Chemical Engineering**. v.16, p. 379-391, 1992.

QUINTERO-MARMOL, E.; LUYBEN, W. L.; GEORGARKIS, C. Applications of an extended Luenberger observer to the control of multicomponent batch distillation. **Industrial Engineering Chemical Research**. v.30, p.1870-1880, 1991.

QUINTERO-MARMOL, E.; LUYBEN, W.L. Inferential Model-Based Control of Multicomponent Batch Distillation. **Chemical Engineering Science**, v.47 (4), p.887–898, 1992.

RAVI, V.; REDDY, P.J.; DUTTA, D. Application of fuzzy nonlinear goal programming to a refinery model. **Computers & Chemical Engineering**. v.22, p.709-712, 1997.

RAMCHANDRAN, S.; RHINEHART, R., Do Neural Networks offer Something for you? **InTech**, nov./95, 1995.

REFENES, A. N.; AZEMA-BARAC, M.; CHEN, L. KAROUSOS, S.A., Currency Exchange Rate Prediction and Neural Network Design Strategies. **Neural Computing and Applications**. v.1, p.46-58, Springer-Verlag, London, 1993.

RIZZO, A.; XILIBIA, M.G. An innovative intelligent system for sensor validation in Tokamak machines. **IEEE Transactions on Control System Technology**. v.10 (3), 2002.

RIGGS, J. B. Improve Distillation Column Control, **Chemical Engineering Process**. October, p. 31-47. 1998.

RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster backpropagation learning: the rprop algorithm. **IEEE International Conference on Neural Networks**. v1, p. 586-591, 1993.

ROCHA, G. B. **Aplicação do controle preditivo bilinear compensado a uma coluna de fracionamento de butadieno 1,3**. Dissertação. (Mestrado em Engenharia Elétrica). Departamento de Pós-Graduação em Engenharia Elétrica, UFBA, Salvador, BA, Brasil, 2003.

ROGINA, A.; SISCO, I.; MOHLER, I.; UJEVIC, Z.; BOLF, N. Soft sensor for continuous product quality estimation (in crude distillation unit). **Chemical Engineering Research and Design**, *In Press, Corrected Proof*, 2011.

RONG, B. G. Synthesis of dividing-wall columns (DWC) for multicomponent distillations - A systematic approach. **Chemical Engineering Research and Design**, v.89, p.1281-1294, 2011

ROSJORDE A.; KJELSTRUP, S. The second law optimal state of a diabatic binary tray distillation column. **Chemical Engineering Science**, v.60, p. 1199-1210, 2005.

RUMELHART, D.E.; McCLELLAND, J. L. **Paralled Distributed Processing**. v.1, The MIT Press., 1986.

RUÉLA, A. L. **Deconvolução de perfis de peço através de Rede Neural Artificial recorrente**. Dissertação. (Mestrado em Geofísica). Departamento de Pós-Graduação de Geofísica. Universidade Federal do Pará, 2005.

SANTOS, M. **Cromatógrafos para controle e otimização de processos**. WNAF 2004 – Workshop nacional sobre analisadores de processos em linha, Salvador, Bahia: UFBA, p. 44, 2004.

SANTOS, M. A.; SEIXAS, J. M.; PEREIRA, B. B.; MEDRONHO, R. A. Usando redes neurais artificiais e regressão logística na predição da hepatite A. **Revista Brasileira de Epidemiologia**. v.2, p.117-26, 2005.

SALAHSHOOR, K.; HAMZEHNEJAD, M. A novel online affine model identification of multivariable processes using adaptive neuro-fuzzy networks. **Chemical Engineering Research and Design**. v. 88, p. 155–169, 2010.

SALAMON, P.; STEVEN, E.; RIVERO, R.; RENDON, C.; HOFFMANN, K.H.; SCHALLER.M.; ANDRESEN, B. Optimization of a Diabatic Distillation Column with Sequential Heat Exchangers. **Engineering Chemical Research**. v.43, p.7566–7571. DOI: 10.1021/ie0495933, 2004.

SCHÜHLI, E. S. **Reconhecimento de gestos de maestro utilizando redes neurais artificiais parcialmente recorrentes**. Mestrado (Dissertação de Mestrado). Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Paraná, Curitiba, PR, 2005.

SCHALLER, M.; HOFFMANN, K. H.; SIRAGUSA, G.; SALAMON, P.; ANDRESEN, B. Numerically optimized performance of diabatic distillation columns. **Computers & Chemical Engineering**, v. 25, p. 1537-1548, 2001.

SEBORG, D. E.; EDGAR, T. F.; MELLICAMP, D. A. **Process Dynamics and Control**, John Wiley & Sons, Singapore, 1989.

SECCHI, A.R.; CARDOZO, N. S. M.; ALMEIDA, N. E.; FINKLER, T. F. An Algorithm for Automatic Selection and Estimation of Model Parameters. Proceedings of the **International Symposium on Advanced Control of Chemical Processes -ADCHEM 2006**. v. 2, p. 789-794. Gramado, Brasil, 2006.

SEIDER, W. D.; SEADER, J. D.; LEWIN, D.R. **Product & Process Design Principles: Synthesis, Analysis and Evaluation**, 2<sup>nd</sup> Ed., John Wiley, New York, 2003.

SHAHAVAND, A.; DERAKHSHAN, F.F.; SOTOUDEH, F. Application of artificial neural networks for simulation of experimental CO<sub>2</sub> absorption data in a packed column. **Journal of Natural Gas Science and Engineering**. v.3, p. 518-529, 2011.

SHU, L., CHEN, L.; SUN, F. Performance optimization of a diabatic distillation-column by allocating a sequential heat-exchanger inventory. **Applied Energy**. v.84, p. 893–903, 2007.

SILVA, F. A. T. F. **Rede morfológica não-supervisionada-RMNS**. Tese (Doutorado em Pesquisas Espaciais). Instituto Nacional de Pesquisas Espaciais – INPE, 1998.

SILVA, D. R. C. **Redes neurais artificiais no ambiente de redes industriais foundation fieldbus usando blocos funcionais padrões**. Dissertação. (Mestrado em Engenharia Elétrica). Universidade Federal do Rio Grande do Norte - Natal, 2005.

SINHA, R.; RIGGS, J. B. High purity distillation control using nonlinear process model based control. **Adv. Inst. Proceed.** Lubbock, EUA, v.44, p.765-772, 1989.

SINGH, V.; GUPTA, I.; GUPTA, H. O. ANN based estimator for distillation - inferential control. **Computers & Chemical Engineering and Processing**. v.44, p.785-795, 2005.

SINGH, V.; GUPTA, I.; GUPTA, H. O. ANN-based estimator for distillation using *Levenberg–Marquardt* approach. **Engineering Applications of Artificial Intelligence**. v. 20, p.249-259, 2007.

SOARES, A. S. **Predição de séries temporais econômicas por meio de redes neurais artificiais e transformada de wavelet: combinando modelo técnico e fundamentalista**. Dissertação (Mestrado em Engenharia Elétrica). Escola de Engenharia de São Carlos, Universidade São Paulo, 2008.

SOUZA, L. G. M. **Proposição e avaliação de algoritmos de filtragem adaptativa baseados na rede de Kohonen**. Dissertação (Mestrado em Engenharia de Teleinformática). Programa de Pós-Graduação em Engenharia de Teleinformática. Universidade Federal do Ceará, 2005.

SOPHOS, A. G.; STEPHANOPOULOS, G.; MORARI, M. Synthesis of optimum distillation sequences with heat integration schemes. In **71<sup>st</sup> Annual AIChE Meeting**. 1978.

SOMSUNG, K.; PRATISHTHAMANDA, S. Design and implementation of fuzzy supervisory PI controllers using fuzzy c-means clustering combined with of fuzzy gain scheduling for a binary distillation column. **IEEE Região 10**, Tencon, p.1-6. Novembro, 2005.

SONG, J. Scheduling of Actual Size Refinery Processes Considering Environmental Impacts with Multi-objective Optimization. **Industrial Engineering Chemical Research**. v. 41, p. 4794-4806, 2002.

SON, J. S.; LEE, D.M; CHOI, S.K. A study on genetic algorithm to select architecture of a optimal neural network in the hot rolling process. **Journal of Materials Processing Technology**. p. 643-648, 2004.

SKOGESTAD, S. The Dos and Don't of Distillation Column Control. **IChemE, Symposium Series**, n. 152, p. 28-43, 2007a.

SKOGESTAD, S.; HORI, S. Selection of control structure and temperature location for two-product distillation columns. **Trans IChemE, Part A, Chemical Engineering Research and Design**, 85(A3) p. 293-306, 2007b.

SUGIT, S. J.; DAOUTIDIS, P. Dynamics and control of vapor recompression distillation. **Journal of Process Control**. v. 19, p.1737-1750, 2009.

SU, C. L; YANG, S. M.; HUANG, W. L. A two-stage algorithm integrating genetic algorithm and modified Newton method for neural network training in engineering systems. **Expert Systems with Applications**, v.38, p. 12189-12194, 2011.

SUNDARESAN, K. R.; KRISHNASWANY P. R. Estimation of time delay time constant parameters in time, frequency and Laplace domains, **Canadian Journal Chemical Engineering**. p.56, 257, 1977.

SUPLANIT, B. Design of internally heat-integrated distillation column (HIDiC): Uniform heat transfer area versus uniform heat distribution. **Energy**. v.35, p. 1505-1514, 2010.

SUPLANIT, B. Optimal heat distribution in the internally heat-integrated distillation column (HIDiC). **Energy**. v.36, p. 4171-4181, 2011.

SKOGESTAD, S. Dynamics and Control Distillation Columns – A Critical Survey. **Third IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Process**, MC, p.1-25, April, 1992.

SZIGETI, F.; RIOS-BOLIVAR, A., RENNOLA, L. **Dual composition control of a binary distillation columns based on nonlinear state observer**. Proceedings of the 39<sup>th</sup> IEEE Conference on Decision and Control Sydney, Australia December, 2000.

TARSINAFFO, P. M. **Estruturas de integração neural *feedforward* testadas em problemas de controle preditivo**. Tese (Doutorado em Engenharia e Tecnologia Espaciais). Departamento Pós-Graduação em Engenharia e Tecnologia Espaciais/Mecânica Espacial e Controle. INPE – São José dos Campos, 2003.

TELEKEN, J. G. **Análise da fluidodinâmica do prato perfurado de uma unidade piloto de destilação**. Dissertação (Mestrado em Engenharia Química). Departamento de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina – UFSC, Florianópolis, 2009.

TERÁN, R. A. C. **Desenvolvimento de um sensor virtual para a predição das propriedades físico-químicas finais em reações de polimerização em suspensão**. Dissertação (Mestrado em Engenharia Química). Departamento de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina – UFSC, Florianópolis, 2009.

TSOI, A. C.; BACK, A. D. Discrete-time recurrent neural network architectures: a unifying review. **Neurocomputing**. v. 15, n.3, p.183-223, 1997.

TSUN-JEN HO. **Design and modeling studies of internally heat-integrated distillation column**. Thesis (Doctorate in Chemical and Material Engineering). Department of Chemical and Material Engineering, National Central University, Japão, 2010.

ÜBEYLİ, M.; ÜBEYLİ, E. D. Using recurrent neural networks for estimation of minor actinides' transmutation in a high power density fusion reactor, **Expert Systems with Applications**, v.37, p. 2742–2746, 2010.

VANLEEUEWEN, J. A.; JONKER, R. J.; GILL R., Octane Number Prediction Based on Gas Chromatographic Analysis with Non-linear Regression Techniques. **Chemometrics & Intelligent Laboratory Systems**. v.25, p. 325-340, 1994.

VASAN, A.; RAJU, K.S. Comparative analysis of Simulated Annealing, Simulated Quenching and Genetic Algorithms for optimal reservoir operation. **Applied Soft Computing** v.9, p. 274–28, 2009.

VENTIN, F. F. **Controle robusto de uma torre estabilizadora de nafta. Dissertação**, 126p (Mestrado em Engenharia Química). Departamento de Pós- Graduação em Engenharia Química (COPPE) da Universidade Federal do Rio de Janeiro, 2010.

VENKATESWARLU, C.; AVANTIKA, S. Optimal state estimation of multicomponent batch distillation, **Chemical Engineering Science**. v. 56, p.5771–5786, 2001.

VENKATESWARLU, C.; KUMAR, B.J. Composition estimation of multicomponent reactive batch distillation with optimal sensor configuration. **Chemical Engineering Science**. v.61, p.5560-5574, 2006.

VEAUX, R. D.; PSICHOGIOS, D. C.; UNGAR, L.H. A comparison of two non parametric estimation schemes: Mars and neural networks. **Computers & Chemical Engineering**. v.17, n.18, p.819-837, 1993

WANG, D., SRINIVASA, R.; LIU, J.; GURU, P.N.S.; LEONG, K. M. Data-driven soft sensor approach for quality prediction in a refinery process. **IEEE International Conference on Industrial informatics**. p.230-235, 2006.

WANG. X.; RONGFU L.; HUIHE S. **Designing a soft sensor for a distillation column with the fuzzy distributed radial basis function neural network**. 35th Conference on Decision and Control, 1714-1719, 1996.

WANG, X.; HUIHE, S. Soft Sensors in Fluid Catalytic Cracking. **3<sup>rd</sup> World Congress on Intelligent Control and Automation**. China, p. 1492-1496, 2000.

WAH, E. F.; HARRIOT, P. Understanding and prediction of the dynamic behavior of distillation columns. **Industrial and Engineering Chemistry Process Design and Development**. v.9, p. 396, 1970.

WELCH, G.; BISHOP, G. **An introduction to the Kalman filter**. University of North Carolina at Chapel Hill, April, 2004.

WEBER, R; BROSILOW, C. The use of secondary measurements to improve control, **AIChE** v.3, n.18, p. 614–627, 1972.

WERLE, L. O. **Minimização dos transientes através do aquecimento distribuído em uma coluna de destilação**. Dissertação. (Mestrado em Engenharia Química) - Universidade Federal de Santa Catarina – UFSC. Florianópolis, 2007.

WEI-ZHONG, A.; XI-GANG, Y. A simulated annealing-based approach to the optimal synthesis of heat-integrated distillation sequences, **Computers and Chemical Engineering**. v. 33 p. 199–212, 2009.

WOLD, S.; KETTNEH-WOLD, N.; SKAGERBERG, B. Non-linear PLS modeling. **Chemometrics and Intelligent Laboratory Systems**. v.7, p.53-65, 1989.

YAMAZAKI, T. Application of fuzzy modeling to distillation control. **IEEE International Conference of Fuzzy Systems**. v.1, p.654-659, 1996.

YILDIZ, U.; GURKAN, U.A.; OZGEN, C.; LEBLEBICIOGLU, K. State Estimator Design for Multicomponent Batch Distillation Columns. **Chemical Eng. Research and Design**. v.83 p. 1–12, 2005.

ZIEGLER, J. B. ; NICHOLS, N. B. Optimum settings for automatic controllers, **ASME Transactions**, v.64, p. 759-768. *Original paper*, 1942.

ZAMPROGNA, E.; BAROLO, M.; SEBORG, D. E. Composition estimations in a middle-vessel batch distillation column using artificial neural networks. **Chemical Engineering Research and Design -Trans IChemE**, v.79, part A, n.6, p.689-696, 2001a.

ZAMPROGNA, E. **Development of virtual sensors for batch distillation monitoring and control using multivariate regression techniques**, Ph.D. Dissertation (Master Degree in Chemical Engineering) Department of Chemical Engineering Principles and Practice, University of Padova, Italy, 2001b.

ZAMPROGNA, E.; BAROLO, M.; SEBORG, D. E. Estimating product composition profiles in batch distillation via partial least squares regression. **Control Engineering Practice** v.12, pp 917–929, 2004.

ZANATA, D. R. P. **Desenvolvimento de sensor virtual empregando redes neurais para medição da composição em uma coluna de destilação**. Dissertação, 229 p. (Mestrado em Telecomunicações e Controle). Escola Politécnica da Universidade de São Paulo - Departamento de Engenharia de Telecomunicações e Controle, 2005.

ZHAO, S. J.; XU, Y. M. Levenberg–Marquardt algorithm for nonlinear principal component analysis neural network through inputs training, in: **Fifth World Congress on Intelligent Control and Automation**. v. 4, p. 3278–3281, 2004.

ZHAO, F.; TIAN, Y.C.; TADÉ, M. O.; LI, H. A time-delay compensation strategy for process with uncertainties. **Computers & Chemical Engineering**. v.26, p.1835-1846, 2000.

ZHANG, J.; MARTIN, E. B.; MORRIS, A. J. KIPARISSIDES, C. Inferential estimation of polymer quality using stacked neural networks. **Computers & Chemical Engineering**. v.21, p.1025, 1997.

ZHANG, J. Inferential feedback control of distillation composition based on PCR and PLS models. **Proceedings of American Control Conference**. v. 2, Arlington, USA, p. 1196–1201, 2001.



## APÊNDICE

### A.1 Código base do treinamento da rede neural feedforward MPL (LM\_TRE)

```
%----- Rede Neural Artificial (CÓDIGO BASE)----- %
%
% - PROJETO: Analisadores virtuais baseados em
modelo neural ----- %
% ----- Rede Neural Dinâmica em Malha aberta:---- %
% -8 Entradas (Ta; Qa; Xea; Tb, T2, T5, T7, Tt) -- %
% 1 Saída (Xet topo) ----- %
%
% - - Algoritmo: Levenberg-Marquardt (LM) ----- %
% --- Função de ativação: Logarítmica (LOGSIG) -- %
% ----- Rede Neural Artificial: feedforward MPL de
duas camadas [15 10]----- %

% ----- Carregando os dados ----- %
% ----- Valores para X e Y ----- %
load X.mat;
load Y.mat;
X = X';
Y = Y';
% - Normalização dos conjuntos de dados X e Y
Função logarítmica [0,1] -- ----- %

[P,ps1] = mapminmax(X,0.1,0.9); %X
[T,ts1] = mapminmax(Y,0.1,0.9); %Y

% ----- Topologia da rede neural----- %
% - Distribuição do número de dados para o
treinamento, validação e teste----- %

[trainV,val,test] = dividevec(P,T,0.15,0.15);
net10 = newff(P,T,[15 10],{'logsig'
'logsig'},'trainlm','learngdm' 'mse');
```

```

% ----- Parâmetros da RNA ----- %

net10.trainParam.epochs = 10000;      % Maximum
number of epochs to train
net10.trainParam.goal = 1e-10;        % Erro mínimo
(Performance goal)
net10.trainParam.show = 10;           % Tempos de
visualização
net10.trainParam.max_fail = 6;         % Maximum
validation failures
net10.trainParam.mem_reduc = 1;        % Factor to
use for memory/speed
net10.trainParam.min_grad = 1e-10;    % Minimum
performance gradient
net10.trainParam.mu = 0.001;          % Initial mu
net10.trainParam.mu_dec = 0.1;        % mu decrease
factor
net10.trainParam.mu_inc = 10;          % mu increase
factor
net10.trainParam.mu_max = 1e10;       % Maximum mu
net10.trainParam.show = 2;            % Epochs
between displays
net10.trainParam.showCommandLine = 0; % Generate
command-line output
net10.trainParam.showWindow = 1;      % Show
training GUI
net10.trainParam.time = inf;           % Maximum
time to train in seconds

% ----- Treinamento da rede neural ----- %

net10 = initnw (net10,1);
[net10,tr10] =
train(net10,trainV.P,trainV.T,[],[],val,test);
save net10;
save tr10;
close all;
clc;

```

## A.2 Código base do treinamento da rede neural recorrente Elman (BR\_TRE)

```

% ----- Rede Neural Artificial (CÓDIGO BASE) ----- %
%
% PROJETO: Analisadores virtuais baseados em modelo
neural ----- %
% -- Rede Neural Dinâmica em MALHA ABERTA: ----- %
% - 8 Entradas (Ta; Qa; Xea; Tb, T2, T5, T7, Tt) -%
% --1 Saída (Xet topo) ----- %
%
% ---Algoritmo: Regularização de Bayesian (BR) -- %
% - Função de ativação: Tangente hiperbólica e
linear (TANSIG - PURELIN)----- %
% --- Rede Neural Artificial Recorrente de Elman de
três camadas [8 8 8]----- %

% ----- Carregando os dados ----- %
% -----Valores para X e Y ----- %
load X.mat;
load Y.mat;
X = X';
Y = Y';
% - Normalização dos conjuntos de dados X e Y
Função tansig [-1,1] ----- %

[P,ps1] = mapminmax(X,-0.9,0.9); %X
[T,ts1] = mapminmax(Y,-0.9,0.9); %Y

% ----- Topologia da rede neural----- %
% - Distribuição do número de dados para o
treinamento, validação e teste -----%

[trainV,val,test] = dividevec(P,T,0.15,0.15);
net18 = newelm(P,T,[8 8 8],{'tansig'
'purelin'},'trainbr','learnngdm' 'mse');

```

```

% ----- Parâmetros da RNA ----- %

net18.trainParam.epochs = 10000;      % Maximum
number of epochs to train
net18.trainParam.goal = 1e-10;        % Erro mínimo
(Performance goal)
net18.trainParam.show = 10;           % Tempos de
visualização
net18.trainParam.max_fail = 6;         % Maximum
validation failures
net18.trainParam.mem_reduc = 1;        % Factor to
use for memory/speed
net18.trainParam.min_grad = 1e-10;    % Minimum
performance gradient
net18.trainParam.mu = 0.001;          % Initial mu
net18.trainParam.mu_dec = 0.1;        % mu decrease
factor
net18.trainParam.mu_inc = 10;          % mu increase
factor
net18.trainParam.mu_max = 1e10;       % Maximum mu
net18.trainParam.show = 2;            % Epochs
between displays
net18.trainParam.showCommandLine = 0; % Generate
command-line output
net18.trainParam.showWindow = 1;      % Show
training GUI
net18.trainParam.time = inf;           % Maximum
time to train in seconds

% ----- Treinamento da rede neural ----- %

net18 = initnw (net18,1);
[net18,tr18] =
train(net18,trainV.P,trainV.T,[],[],val,test);
save net18;
save tr18;
close all;
clc;

```

### A.3 Código para cálculo da saída da rede neural para dados de treinamento (Y00) e da validação cruzada (Y01F)

```

% ----- Limpando a área de trabalho ----- %
clear all;
close all;
clc;
% ----- Chamada da RNA ----- %
load net10.mat;
load tr10.mat;

% -Carregamento do conjunto de dados de entrada X-%

load X.mat;          % dados do Treinamento
X = X';
load X_VAL.mat;     % dados da Validação
X_VAL = X_VAL';

% -Carregamento do conjunto de dados de saída Y %

load Y.mat;          % dados do Treinamento
load Y_VAL.mat;     % dados da validação

% ----- Normalização dos dados Tan [-1,1] ,Sig
[0,1], log [0,1] ----- %

P0 = mapminmax('apply',X,ps1);
P2 = mapminmax('apply',X_VAL,ps1);

%----- %

A0 = sim(net10,P0);
Y0 = mapminmax('reverse',A0,ts1);

A1 = sim(net10,P2);
Y1 = mapminmax('reverse',A1,ts1);

```

```
% --Geração das Matrizes de saída: Simulado versus
Neural----- %

Y00F= [Y Y0'];
Y01F= [Y_VAL Y1'];

save Y00F.mat; % treinamento      (saída simulada
versus neural)

save Y01F.mat; % validação cruzada (saída simulada
versus neural)

%----- Plota os Gráficos -----%

figure(1);plot(Y00F(1:78833,1:2));
figure(2);plot(Y01F(1:18610,1:2));
```

#### A.4 Código para cálculo do desempenho da rede neural (MSE\_tre; ME\_val; MAE\_val; MSE\_val; MPE\_val; MAPE\_val)

```
% ----- Limpando área de trabalho ----- %
clear all;
close all;
clc;
% ----- Carregamento de Y00 e Y01F ----- %

load Y00F.mat; % treinamento, matriz 2 colunas-
saída simulada versus neural
load Y01F.mat; % validação, matriz 2 colunas-saída
simulada versus neural

% ----- Cálculos de MSE_etapa TREINAMENTO----- %
Y00Fa = Y00F(1:78833,1);
Y00Fb = Y00F(1:78833,2);
n0=78833;

MSE0=0;
format short
MSE0 = ((Y00Fa-Y00Fb).^2);
MSE_tre=sum(MSE0)/n0;

% ----- Cálculos de ME_etapa VALIDAÇÃO ----- %

Y01Fa = Y01F(1:18610,1);
Y01Fb = Y01F(1:18610,2);
n1=18610;

ME1=0;
format short
ME1 = ((Y01Fa-Y01Fb));
ME_val=sum(ME1)/n1;

% ----- Cálculos de MAE_etapa VALIDAÇÃO ----- %

Y01Fa = Y01F(1:18610,1);
Y01Fb = Y01F(1:18610,2);
n2=18610;
MAE1=0;
format short
```

```

MAE1 = abs(Y01Fa-Y01Fb);
MAE_val=sum(MAE1)/n2;

% -----Cálculos de MSE_etapa VALIDAÇÃO ----- %

Y01Fa = Y01F(1:18610,1);
Y01Fb = Y01F(1:18610,2);
n3=18610;
MSE1=0;
format short
MSE1 = ((Y01Fa-Y01Fb).^2);
MSE_val=sum(MSE1)/n3;

% ----- Cálculos de MPE_etapa VALIDAÇÃO ----- %

Y01Fa = Y01F(1:18610,1);
Y01Fb = Y01F(1:18610,2);
n4=18610;
PE1=0;
format short
PE1 = ((Y01Fa-Y01Fb)./Y01Fa)*100;
MPE_val = sum(PE1)/n4;

% -----Cálculos de MAPE_etapa VALIDAÇÃO ----- %

Y01Fa = Y01F(1:18610,1);
Y01Fb = Y01F(1:18610,2);
n5=18610;
PE4=0;
format short
PE4 = abs((Y01Fa-Y01Fb)./Y01Fa).*100;
MAPE_val=sum(PE4)/n5;

% SALVAR matriz com dados do desempenho das redes %

MEDIDAS_DE_DESEMPENHO = [MSE_tre ME_val MAE_val
MSE_val MPE_val MAPE_val];

```