

# *Monografia de Graduação*

## **Sistema de Inferência Neural e Processamento Estatístico Multivariável Aplicado a Indústria do Petróleo**

**Diogo Leite Rebouças**

**Natal, junho de 2009**

# **Sistema de Inferência Neural e Processamento Estatístico Multivariável Aplicado a Indústria do Petróleo**

**Diogo Leite Rebouças**

Orientador: Prof. Dr. Fábio Meneghetti Ugulino de Araújo

**Trabalho de Conclusão de Curso** apresentado ao Corpo docente do Curso de Engenharia de Computação da Universidade Federal do Rio Grande do Norte como parte dos requisitos para obtenção do título de Bacharel em Engenharia de Computação.

**Natal – RN  
Junho/ 2009**

# **Sistema de Inferência Neural e Processamento Estatístico Multivariável Aplicado a Indústria do Petróleo**

**Diogo Leite Rebouças**

Trabalho de Conclusão de Curso aprovado em 26 de junho de 2009 pela banca examinadora composta pelos seguintes membros:

---

Prof. Dr. Fábio Meneghetti Ugolino de Araújo (Orientador) ..... DCA/UFRN

---

Prof. Dr. André Laurindo Maitelli ..... DCA/UFRN

---

Prof. Me. José Medeiros de Araújo Júnior ..... DCA/UFRN

*“Algo só é impossível até que  
alguém duvide e acabe provando o  
contrário.”*

*Sir Albert Einstein*

*Dedico este trabalho primeiramente a Deus, que me confortou nos momentos mais difíceis e que indicou o caminho certo ao longo de toda minha vida. Ao meu anjo da guarda, pela proteção e iluminação. À toda minha família e, em especial, aos meus pais, Marcondes e Aretusa, pela compreensão e dedicação ao longo desta e de outras jornadas. À minha namorada pelo apoio e incentivo durante a realização deste trabalho. Que não lhes falte saúde, paz e tranquilidade. Fica aqui minha gratidão eterna, pois sem vocês nada seria possível.*

# AGRADECIMENTOS

---

Aos meus pais, pelo incentivo e apoio incondicional ao longo de toda minha vida.

Ao meu orientador professor Fábio Meneghetti Ugulino de Araújo, sou grato pela orientação.

À todos os professores e funcionários do Departamento de Engenharia de Computação e Automação, pelo aprendizado e pela ajuda.

Ao professor Adelardo Adelino Dantas de Medeiros, por ter desenvolvido o modelo de tese do Programa de Pós Graduação em Engenharia Elétrica e de Computação.

Aos colegas de graduação, pelas críticas e sugestões.

À CAPES e ao CNPq, pelo apoio financeiro.

# RESUMO

---

O processamento de gás natural é uma das mais importantes atividades da indústria petroquímica. Em uma Unidade de Processamento de Gás Natural (UPGN) o produto que possui o maior potencial financeiro é o gás liquefeito de petróleo (GLP), apesar de outros produtos, como a gasolina natural e o gás residual de qualidade industrial, também serem obtidos como resultado do processo de fracionamento.

Devido a grande concorrência de mercado, as empresas buscam reduzir suas despesas quanto a produção e a quantidade de matéria prima utilizada, sem que o produto perca sua qualidade. Para que isso seja possível faz-se necessário o desenvolvimento de estratégias de controle eficientes.

Nas UPGNs, esse controle de qualidade é baseado nas composições dos produtos finais. No entanto, mesmo quando esta análise é feita a partir de instrumentos específicos, como os cromatógrafos, o tempo de medição e purgação das amostras inviabiliza a elaboração de estratégias de controle que melhorem o desempenho da planta.

O objetivo deste trabalho é desenvolver um sistema de inferência utilizando redes neurais artificiais (RNAs), que seja capaz de estimar as frações molares de etano e pentano no GLP e a fração molar do propano no gás residual. Para isso, será utilizada uma UPGN simulada no software HYSYS<sup>®</sup>, formada pelas colunas de destilação deetanizadora e debutanizadora.

Esse sistema deve realizar a estimativa das frações molares dos produtos a cada minuto, permitindo a aplicação de técnicas de controle multivariável com o intuito de controlar a qualidade do GLP e minimizar as perdas do processo. Para reduzir a complexidade da rede neural será utilizada a técnica estatística de Análise de Componentes Principais (ACP).

**Palavras-chave:** Inferência, Coluna de Destilação, Análise de Componentes Principais, Redes Neurais Artificiais.

# ABSTRACT

---

The natural gas processing is one of the most important activities of petrochemical industry. In a Natural Gas Processing Unit (NGPU) the liquefied petroleum gas (LPG) is the product that has the largest economical potencial, although other products, such as natural gasoline and industrial quality residual gas, are also obtained as a fractional process result.

Due to the high market competition, the companies are reducing their expenditures with production and quantity of raw material used, without the lost of product quality. To make this possible, it is necessary to develop efficient control strategies.

In NGPUs, that quality control is based on the final products compositions. However, even when this analysis is performed by specific instruments, such as chromatographs, the measurement and purgation time of the samples hinder the development of advanced control strategies.

The goal of this work is to develop an inferential system using artificial neural networks (ANNs) that can estimate the ethane and pentane molar fraction on LPG and the propane molar fraction on residual gas. The system will uses a NGPU simulated by HYSYS<sup>®</sup> software, composed by a deethanizer column in series with a debuthanizer column.

The system should estimate the products molar fractions every minute, allowing the application of multivariable control techniques in order to maintain the LPG quality and minimize the process loss. The statistical technique of principal Components Analysis is used to reduce the neural network complexity.

**Keywords:** Inference, Distillation Columns, Principal Components Analysis, Artificial Neural Networks.

# SUMÁRIO

---

<b>Sumário</b>	<b>i</b>
<b>Lista de Figuras</b>	<b>iii</b>
<b>Lista de Tabelas</b>	<b>v</b>
<b>Lista de Símbolos e Abreviaturas</b>	<b>vi</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Introdução . . . . .	1
1.2 Sistemas inferenciais para colunas de destilação . . . . .	2
<b>2 O Processamento de Gás Natural</b>	<b>5</b>
2.1 A indústria do gás . . . . .	5
2.2 O gás natural . . . . .	5
2.2.1 Produção e separação . . . . .	6
2.2.2 Condicionamento e o processamento . . . . .	7
2.3 Destilação fracionada . . . . .	8
2.4 UPGN simulada . . . . .	9
2.4.1 Coluna deetanizadora . . . . .	9
2.4.2 Coluna debutanizadora . . . . .	11
<b>3 Sistema de Inferência</b>	<b>13</b>
3.1 Redes neurais artificiais . . . . .	13
3.1.1 Identificação através de modelo neural . . . . .	15
3.1.2 Determinação da ordem do modelo . . . . .	16
3.1.3 Seleção do modelo . . . . .	16
3.2 Análise de componentes principais . . . . .	17
3.2.1 As componentes principais . . . . .	17
3.3 Composição do sistema . . . . .	19
3.3.1 Razão de redução . . . . .	20

3.4	Agregação dos módulos . . . . .	20
3.4.1	Estrutura 1 . . . . .	20
3.4.2	Estrutura 2 . . . . .	21
3.4.3	Estrutura 3 . . . . .	22
3.4.4	Estrutura 4 . . . . .	22
<b>4</b>	<b>Resultados</b>	<b>24</b>
4.1	Coleta dos dados . . . . .	24
4.2	A análise de componentes principais . . . . .	25
4.3	O treinamento da rede neural de inferência . . . . .	25
4.4	A Ordem do modelo e o processamento dos dados . . . . .	25
4.5	Critério de escolha . . . . .	26
4.6	Comparação das estruturas . . . . .	26
4.6.1	Estrutura 1 . . . . .	27
4.6.2	Estrutura 2 . . . . .	28
4.6.3	Estrutura 3 . . . . .	29
4.6.4	Estrutura 4 . . . . .	31
4.7	Análise das RNAs de quarta ordem . . . . .	34
4.8	Melhor estrutura e validação do sistema . . . . .	34
<b>5</b>	<b>Conclusões</b>	<b>38</b>
5.1	Conclusões . . . . .	38
5.2	Perspectivas . . . . .	39
	<b>Referências Bibliográficas</b>	<b>40</b>

## LISTA DE FIGURAS

---

2.1	Esquema de um separador bifásico. . . . .	7
2.2	Coluna ou torre de destilação. . . . .	8
2.3	Coluna deetanizadora simulada. . . . .	10
2.4	Coluna debutanizadora simulada. . . . .	11
3.1	Rede PMC totalmente conectada de duas camadas. . . . .	14
3.2	Esquema de uma rede neural com estrutura NNARX. . . . .	15
3.3	Esquema da estrutura 1. . . . .	21
3.4	Esquema da estrutura 2. . . . .	21
3.5	Esquema da estrutura 3. . . . .	22
3.6	Esquema da estrutura 4. . . . .	22
4.1	Estrutura 1 – Validação de propano ( $r = 4$ ). . . . .	27
4.2	Estrutura 1 – Validação de etano ( $r = 4$ ). . . . .	28
4.3	Estrutura 1 – Validação de pentano ( $r = 4$ ). . . . .	28
4.4	Estrutura 2 – Validação de propano ( $r = 4$ ). . . . .	29
4.5	Estrutura 2 – Validação de etano ( $r = 4$ ). . . . .	29
4.6	Estrutura 2 – Validação de pentano ( $r = 4$ ). . . . .	29
4.7	Estrutura 3 – Validação de propano ( $r = 4$ ). . . . .	30
4.8	Estrutura 3 – Validação de etano ( $r = 4$ ). . . . .	30
4.9	Estrutura 3 – Validação de pentano ( $r = 4$ ). . . . .	31
4.10	Estrutura 4 – Validação de propano ( $r = 4, \gamma = 0,84$ ). . . . .	32
4.11	Estrutura 4 – Validação de etano ( $r = 4, \gamma = 0,84$ ). . . . .	32
4.12	Estrutura 4 – Validação de pentano ( $r = 4, \gamma = 0,84$ ). . . . .	32
4.13	Estrutura 4 – Validação de propano ( $r = 4, \gamma = 0,75$ ). . . . .	33
4.14	Estrutura 4 – Validação de etano ( $r = 4, \gamma = 0,75$ ). . . . .	33
4.15	Estrutura 4 – Validação de pentano ( $r = 4, \gamma = 0,75$ ). . . . .	33
4.16	Validação de propano (Degrau negativo, Estrutura 4, $r = 4$ e $\gamma = 0,75$ ). . . . .	35
4.17	Validação de etano (Degrau negativo, Estrutura 4, $r = 4$ e $\gamma = 0,75$ ). . . . .	35
4.18	Validação de pentano (Degrau negativo, Estrutura 4, $r = 4$ e $\gamma = 0,75$ ). . . . .	36
4.19	Validação de propano (PRBS, Estrutura 4, $r = 4$ e $\gamma = 0,75$ ). . . . .	36

4.20 Validação de etano (PRBS, Estrutura 4, $r = 4$ e $\gamma = 0,75$ ). . . . .	37
4.21 Validação de pentano (PRBS, Estrutura 4, $r = 4$ e $\gamma = 0,75$ ). . . . .	37

# LISTA DE TABELAS

---

2.1	Composição de uma amostra do gás natural em Guamaré-RN. . . . .	8
3.1	Variáveis de processo escolhidas. . . . .	19
4.1	Detalhes sobre as validações. . . . .	26
4.2	Comparação da ACP para as estruturas propostas. . . . .	26
4.3	EMQs Estrutura 1 – Ordem 2, 3 e 4 ( $r = 2$ , $r = 3$ e $r = 4$ ). . . . .	27
4.4	EMQs Estrutura 2 ( $r = 4$ ). . . . .	28
4.5	EMQs Estrutura 3 – Ordem 2, 3 e 4 ( $r = 2$ , $r = 3$ e $r = 4$ ). . . . .	30
4.6	EMQs Estrutura 4 ( $r = 4$ ). . . . .	31
4.7	Melhores redes ( $r = 4$ ). . . . .	34
4.8	Características das RNAs ( $r = 4$ ). . . . .	34

# LISTA DE SÍMBOLOS E ABREVIATURAS

---

$\delta_i$	$i$ -ésima saída da rede neural
$\gamma$	Porcentagem da variância total
$\Lambda$	Matriz de covariância de $\mathbf{Y}$
$\lambda_i$	$i$ -ésimo autovalor de $\Lambda$
$\hat{\lambda}_i$	$i$ -ésimo autovalor de $\mathbf{S}$
$\hat{y}_i$	$i$ -ésimo valor estimado pelo sistema
$\hat{y}(t)$	Saída estimada pela rede neural
$\omega_{j,l}$	Peso sináptico que parte do neurônio $l$ e chega no neurônio $j$
$\phi$	Vetor de entradas da rede neural
$\phi_l$	$l$ -ésima entrada da rede neural
$\rho$	<i>biases</i>
$\rho_{i,j}$	<i>biases</i> que partem da camada $j$ e chegam na camada $i$
$\Sigma_{p \times p}$	Matriz de covariância de $\mathbf{X}$
$\theta$	Vetor de parâmetros ajustáveis da rede neural
$C_1$	Metano ( $C_1H_4$ )
$C_2$	Etano ( $C_2H_6$ )
$C_3$	Propano ( $C_3H_8$ )
$C_4$	Butano ( $C_4H_{10}$ )
$C_5$	Pentano ( $C_5H_{12}$ )

$C_{N+}$	Conjunto de subprodutos contendo cadeias carbônicas com $N$ ou mais carbonos
$CO_2$	Dióxido de carbono
$d$	Atraso de transporte
$E_j$	Variância total da $j$ -ésima componente principal amostral
$F, f$	Funções de ativação dos neurônios da rede neural
$f_l$	Função de ativação linear
$f_s$	Função de ativação sigmoideal
$f_t$	Função de ativação tangente hiperbólica
$H_2S$	Sulfeto de hidrogênio
$k$	Número de componentes principais
$m$	Ordem de entrada
$N$	Número de amostras de validação
$n$	Ordem de saída
$N_2$	Nitrogênio gasoso
$n_\phi$	Número de entradas da rede
$n_h$	Número de neurônios da camada oculta
$p$	Número de variáveis do sistema original
$r$	Ordem do modelo
$R_r$	Razão de redução
$u$	Entrada da planta
$V_i$	$i$ -ésima variável de processo escolhida
$y$	Saída da planta
$y_i$	$i$ -ésimo valor real de saída da planta
$\hat{\mathbf{e}}_i$	$i$ -ésimo autovetor normalizado de $\mathbf{S}$
$\hat{\mathbf{Y}}_j$	$j$ -ésima componente principal amostral estimada
$\mu_y$	Vetor de médias de $\mathbf{Y}$

$\mu$	Vetor de médias de $\mathbf{X}$
$\mathbf{e}_i$	$i$ -ésimo autovetor normalizado de $\mathbf{X}$
$\mathbf{O}$	Matriz ortogonal de dimensão $p \times p$ constituída dos autovetores normalizados de $\Sigma_{p \times p}$
$\mathbf{S}$	Matriz de covariância amostral
$\mathbf{X}$	Vetor aleatório composto por $p$ variáveis aleatórias
$\mathbf{Y}$	Vetor aleatório composto por $p$ combinações lineares das variáveis aleatórias de $\mathbf{X}$
<b>ACP</b>	Análise de Componentes Principais
<b>ANEEL</b>	Agência Nacional de Energia Elétrica
<b>ARMAX</b>	<i>AutoRegressive Moving Average eXternal input</i>
<b>ARX</b>	<i>AutoRegressive eXternal Input</i>
<b>EMQ</b>	Erro Médio Quadrático
<b>FIR</b>	<i>Finite Impulse Response</i>
<b>GLP</b>	Gás Liquefeito de Petróleo
<b>GMR</b>	Guamaré
<b>LGN</b>	Líquido de Gás Natural
<b>MQP</b>	Mínimos Quadrados Parciais
<b>NGE</b>	<i>Nonlinear Geometric Estimator</i>
<b>NNARMAX</b>	<i>Neural Network AutoRegressive Moving Average eXternal input</i>
<b>NNARX</b>	<i>Neural Network AutoRegressive eXternal Input</i>
<b>NNOE</b>	<i>Neural Network Output Error</i>
<b>NNSSIF</b>	<i>Neural Network State Space Innovations Form</i>
<b>OE</b>	<i>Output Error</i>
<b>PCR</b>	<i>Principal Components Regression</i>
<b>PID</b>	Proporcional, Integrativo e Derivativo
<b>PLS</b>	<i>Partial Least Squares</i>
<b>PMC</b>	Perceptron de Múltiplas Camadas,

<b>PRBS</b>	<i>Pseudo Random Binary Signal</i>
<b>PV</b>	<i>Process Variable</i>
<b>RNA</b>	Rede Neural Artificial
<b>SSIF</b>	<i>State Space Innovations Form</i>
<b>UPGN</b>	Unidade de Processamento de Gás Natural
<b>UTPF</b>	Unidade de Tratamento e Processamento de Fluidos
<b>VP</b>	Variável principal
<b>VS</b>	Variável secundária

# INTRODUÇÃO

---

## 1.1 Introdução

Segundo Thomas et al. (2001), a indústria do petróleo é, certamente, a mais fantástica atividade produtiva organizada em toda a existência do ser humano. Dentre as atividades realizadas por essa indústria, pode-se dizer que a destilação é uma das mais importantes uma vez que separa a alimentação de um determinado processo industrial em dois ou mais produtos de diferentes composições.

Essa operação, que exige grandes quantidades de energia, chega a demandar uma boa parte do total de investimentos e de custo energético. Entretanto, as necessidades energéticas podem ser significativamente reduzidas se a planta, mais especificamente as colunas de destilação, forem adequadamente projetadas ou quando existem estratégias de controle eficientes, capazes de manter as condições ideais de operação.

Para Pulis (2008), uma má estratégia de controle pode resultar em longos períodos de funcionamento com produtos mais puros do que o necessário e, conseqüentemente, um consumo maior de energia, ou produtos que não atendem as especificações definidas. Abdullah et al. (2007) mostra que um controle efetivo das colunas de destilação pode melhorar a qualidade do produto final, diminuir o consumo de energia e aumentar a capacidade de produção e a segurança do processo. No entanto, para realizar controle ou supervisão em plantas desse tipo faz-se necessário obter medições *online* de variáveis que muitas vezes são compostos químicos que estão diretamente relacionados com a qualidade dos produtos finais e, por esse motivo, são de grande interesse econômico.

Conforme discutido em Warne et al. (2004), mensurar ou estimar variáveis desse tipo passa a ser uma tarefa difícil a partir do momento em que existem limitações tecnológicas ou devido as técnicas de medição utilizadas. Joseph (1999) mostra ainda que os instrumentos que realizam a medição direta dessas composições são normalmente caros, de difícil manutenção e introduzem atrasos de medição significativos, o que impossibilita a criação de estratégias de monitoramento e controle de qualidade.

A falta ou os alto custos desses analisadores acabam incentivando o desenvolvimento de sensores virtuais, implementados através de *softwares* (*soft-sensors*), para medição indireta das variáveis primárias do processo. Uma possível maneira de se fazer essa medição é inferindo os valores das composições a partir das variáveis secundárias do processo. De maneira geral, os sistemas de inferência são construídos para realizar estimativas das variáveis primárias a partir das variáveis secundárias de fácil mensuração, tais como: temperaturas, pressões, níveis e vazões.

Para Warne et al. (2004) existem três tipos de abordagem para a construção desses *softwares*: a modelagem fenomenológica, os métodos de regressão estatística ou ainda a modelagem a partir de alguma técnica de inteligência artificial. A escolha de qual técnica utilizar varia de acordo com a aplicação. No entanto, Zamprognia et al. (2004) deixa claro que, como a maioria dos processos químicos são geralmente complexos de se modelar, devido as suas inerentes características de não-linearidade, os desenvolvedores optam, normalmente, por utilizar uma das duas últimas abordagens.

## 1.2 Sistemas inferenciais para colunas de destilação

Para Pulis (2008), um problema chave para controlar processos químicos é saber como representar, de maneira simples e confiável, a evolução dos estados do sistema (i.e., composições, reagentes, temperaturas, entre outros).

Em diversos processos desse tipo, somente algumas composições podem ser mensuradas em tempo adequado para que seja possível se construir estratégias de controle de qualidade. Conforme discutido anteriormente, isso se deve à ausência ou ao alto investimento necessário para se adquirir analisadores de composição, tais como os cromatógrafos a gás, que, mesmo quando disponíveis, introduzem atrasos de medição indesejáveis.

Diversos trabalhos demonstram a aplicabilidade dos sistemas inferenciais em processos com colunas de destilação. Baratti et al. (1998) utiliza um filtro de Kalman estendido para inferir as frações molares de etanol e butanol resultantes de uma coluna de destilação ternária.

Zhang (2001) apresenta um controlador inferencial realimentado em uma coluna de destilação simulada (metanol-água), em que as composições dos produtos de topo e de fundo são estimadas a partir de sistemas de inferência baseados em PCR (*Principal Components Regression*) e PLS (*Partial Least Squares*).

Mejdell & Skogestad (1991) discutem sobre o uso de estimadores baseados em PCR e PLS implementados em uma planta piloto. Em seguida, comparam sua performance com um *Kalman Bucy Filter*, desconsiderando a importância das informações contidas nas medições. Oisiovici & Cruz (2000) aplicam um filtro de Kalman discreto estendido para estimar a composição de produtos em colunas de destilação fracionada. Venkateswarlu & Avantika (2001) comparam diversas variações dos filtros de Kalman para realizar estimativas em colunas de destilação multicomponentes.

Em Tronci et al. (2005) foi apresentado um NGE (*Nonlinear Geometric Estimator*) de baixa ordem, baseado em geometria diferencial, para inferir a composição do produto em uma coluna de destilação binária (etanol-água).

Na área de inteligência artificial, Bahrtiya & Whiteley (2000) apresenta um estudo sobre a aplicação de redes neurais artificiais para o desenvolvimento de sistemas de inferência em uma refinaria de petróleo. Fortuna et al. (2007) descreve um *soft-sensor* com redes neurais Perceptron de Múltiplas Camadas (PMC) para estimar a octanagem da gasolina produzida em refinarias. Em Lima (2001) é apresentado um medidor inferencial para estimar a fração molar de pentano no Gás Liquefeito de Petróleo (GLP), produto de topo de uma coluna debutanizadora, em que é feito um estudo estatístico que utiliza a correlação para selecionar as entradas mais adequadas para uma rede neural PMC de inferência.

No trabalho a ser aqui apresentado, é proposto um sistema de inferência baseado em redes neurais artificiais, que tem por objetivo estimar valores de variáveis importantes de uma Unidade de Processamento de Gás Natural (UPGN), tornando possível a implementação de um sistema de controle inferencial baseado em técnicas de controle multivariável.

O sistema de inferência será composto por uma única rede neural que deverá estimar satisfatoriamente as frações molares de componentes dos subprodutos da UPGN. A UPGN em questão é formada por diversos elementos, variando de unidade para unidade e será simulada através do *software* HYSYS<sup>®</sup>. A simulação será baseada em uma UPGN real, denominada UPGN-II, situada em Guamaré-RN e será formada por uma coluna deetanizadora em série com uma coluna debutanizadora.

De maneira simplificada, a coluna deetanizadora recebe como carga o gás natural pré-processado e, através de um processo de destilação, produz em seu topo o gás residual, formado por metano ( $CH_4$ , ou simplesmente  $C_1$ ), etano ( $C_2H_6$ , ou simplesmente  $C_2$ ) e traços de propano ( $C_3H_8$ , ou simplesmente  $C_3$ ). Como seu produto de fundo, é obtido o líquido de gás natural, formado por  $C_{3+}$  e traços de  $C_2$ , que alimenta o estágio seguinte: a coluna debutanizadora. Nessa última, são obtidos o GLP, composto idealmente por propano e butano ( $C_4H_{10}$ , ou simplesmente  $C_4$ ) e a gasolina natural, composta por frações de pentano ( $C_5H_{12}$ , ou simplesmente  $C_5$ ) e mais pesadas ( $C_{5+}$ ), como produtos de topo e de fundo, respectivamente.

Na prática, o GLP, que corresponde ao produto de maior importância econômica para a UPGN, apresenta em sua composição alguns contaminantes, tais como o etano não vaporizado na coluna deetanizadora e o pentano vaporizado na coluna debutanizadora. O sistema neural de inferência proposto tem por objetivo realizar a estimativa das frações molares desses contaminantes de forma a controlar a qualidade do GLP bem como da fração molar do  $C_3$  no gás residual, considerado perda do sistema.

A planta escolhida como objeto de estudo é dotada de diversos instrumentos, conforme mostrado no capítulo 2. Apesar disso, existe uma carência quanto ao número de sensores de temperatura nos pratos de suas colunas. Isso ocorre devido a ausência desses sensores na planta real a qual a simulação foi baseada. Sendo assim, optou-se por manter essa peculiaridade com o intuito de estabelecer uma maior proximidade com o sistema real. Dessa forma, as variáveis secundárias serão determinadas a partir das variáveis de processo (PVs – *Process Variables*) dos controladores PIDs presentes nas malhas regulatórias das colunas, utilizando a técnica estatística de análise de componentes principais.

No decorrer deste trabalho será visto que a estrutura da rede neural selecionada possui entre suas entradas valores passados das variáveis primárias que se deseja estimar. Essa realimentação provoca um acúmulo de erros de estimativa que poderá levar o sistema à instabilidade. Conforme proposto em Linhares (2009), uma forma para solucionar esse problema é estabelecer uma metodologia de correção do sistema a partir de medições periódicas realizadas pelos cromatógrafos a gás presentes no processo em estudo.

No capítulo 2 será discutido a respeito do processamento de gás natural e da simulação da UPGN. O capítulo 3 apresenta algumas estruturas para o sistema de inferência proposto enquanto que o capítulo 4 mostra os resultados obtidos. Por fim, o capítulo 5 apresenta as considerações finais e novas perspectivas de trabalho.

# O PROCESSAMENTO DE GÁS NATURAL

---

Neste capítulo serão apresentadas informações sobre a produção e o processamento de gás natural bem como algumas características do gás produzido na UPGN II de Guamaré-RN. Em seguida será comentado sobre o funcionamento de uma UPGN, mostrando o esquema de funcionamento básico de um sistema de destilação fracionada e logo após será apresentada uma breve descrição sobre o sistema simulado.

## 2.1 A indústria do gás

A indústria do gás natural vem apresentando um crescimento significativo ao longo dos anos, o que reforça cada vez mais a posição desse produto como uma fonte de energia alternativa. Segundo a Agência Internacional de Energia, em 2003, a participação do gás natural no consumo mundial de energia era da ordem de 16,3%.

A Agência Nacional de Energia Elétrica (ANEEL) informa que em 2005 as reservas provadas do Brasil eram da ordem de 230 bilhões de  $m^3$ , dos quais 48% estão localizados no Estado do Rio de Janeiro, 20% no Amazonas, 9,6% na Bahia e 8% no Rio Grande do Norte. Para a ANEEL, a participação do gás natural na matriz energética brasileira ainda era pouco expressiva, da ordem de 5,6% do consumo final.

Para a Agência Nacional de Petróleo, Gás Natural de Biocombustíveis (ANP), as reservas provadas saltaram de 227,7 bilhões de  $m^3$  em 2007 para 364,2 bilhões de  $m^3$  em 2008. No mesmo período a produção de gás natural mais que duplicou, passando de 26,5 para 59 milhões de  $m^3$ /dia.

## 2.2 O gás natural

Semelhante aos demais combustíveis fósseis, o gás natural é uma mistura de hidrocarbonetos gasosos originados da decomposição de matéria orgânica fossilizada ao longo

de milhões de anos. Em seu estado bruto, é composto essencialmente por metano, com proporções variadas de etano, propano, butano e hidrocarbonetos mais pesados, além de dióxido de carbono ( $CO_2$ ), nitrogênio gasoso ( $N_2$ ), sulfeto de hidrogênio ( $H_2S$ ), água, ácido clorídrico, metanol e outras impurezas.

As principais características desse gás são a sua densidade em relação ao ar, o poder calorífico, o ponto de orvalho<sup>1</sup> da água e dos hidrocarbonetos, além de seus teores de carbono,  $CO_2$ , hidrogênio, oxigênio e compostos sulfurosos. Outras características intrínsecas importantes são os baixos índices de emissão de gases poluentes, rápida dispersão em caso de vazamentos e baixos índices de odor e contaminantes.

Devido as suas propriedades e características químicas o gás natural é um insumo altamente qualificado para ser usado como matéria prima. Na indústria do petróleo, o gás natural é utilizado para a produção de uma série de compostos.

### 2.2.1 Produção e separação

Para Thomas et al. (2001), ao longo da vida produtiva de um campo de petróleo ocorre, geralmente, a produção simultânea de gás, óleo e água. Devido ao interesse econômico ser direcionado à produção de hidrocarbonetos (gás e óleo), os campos marítimos e terrestres são dotados de certas “facilidades de produção”, que são as instalações destinadas a efetuar, sob condições controladas, o processamento primário de fluidos. Segundo ele a produção do gás pode ser considerada como resultado da composição de três partes, e tem seu início quando, em campo, as produções dos poços são submetidas a processos de separação.

A primeira parte é proveniente dos hidrocarbonetos que nas condições de temperatura e pressão do reservatório já se encontram no estado gasoso. A segunda parte são os hidrocarbonetos que se encontram dissolvidos em óleo e que, quando levados à superfície, são vaporizados. Por fim, a terceira parte são os hidrocarbonetos dissolvidos em água, os quais normalmente não são contabilizados nos cálculos de produção.

A separação do gás natural ocorre através de separadores bifásicos horizontais que por apresentarem uma maior área superficial de interface, permitem uma melhor separação líquido/gás e gás/líquido. Nesse tipo de separador, o fluido entra e choca-se com os defletores de entrada, provocando uma mudança brusca de velocidade e direção do fluido. A força da gravidade causa a separação das gotículas líquidas mais pesadas que deixam a corrente de gás e se acumulam no fundo do vaso, onde o líquido é coletado. Para que isso seja possível essa seção de coleta deve ocorrer dentro de um intervalo de tempo apropriado de modo que o gás consiga ser desprendido do líquido, indo para a parte superior do separador. Um esquema simples de funcionamento pode ser visualizado na figura 2.1.

A pressão do separador é mantida por um controlador que atua regulando o fluxo de saída do gás pela parte superior. O líquido separado deixa o vaso através de uma válvula cuja abertura e fechamento é regulado por um controle de nível [Thomas et al. 2001].

<sup>1</sup>O ponto de orvalho é a temperatura na qual ocorre a formação da primeira gota de líquido quando o gás sofre resfriamento ou compressão.

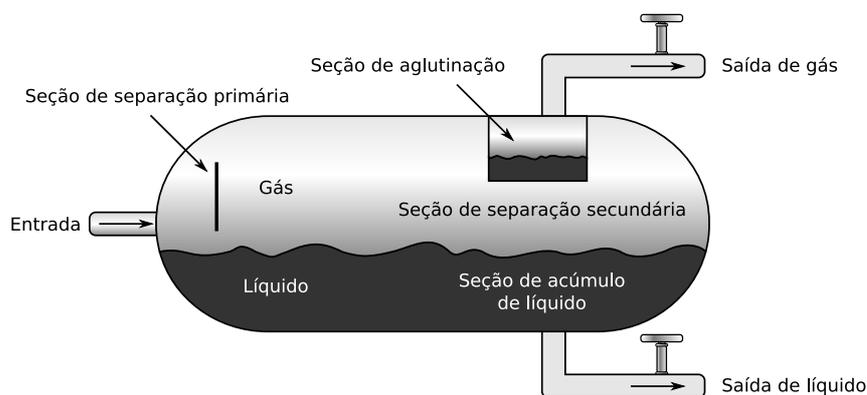


Figura 2.1: Esquema de um separador bifásico.

### 2.2.2 Condicionamento e o processamento

O condicionamento ou o tratamento do gás é o conjunto de processos físicos e/ou químicos aos quais o gás é submetido de modo a remover/reduzir os teores de contaminantes para atender as especificações de mercado. O condicionamento envolve processos de desidratação (absorção ou adsorção) e dessulfurização (absorção química ou física).

Livre da fase líquida, o gás natural é encaminhado para as unidades de processamento de gás natural (UPGN) para que seja promovida a separação das frações leves (metano e etano que constituem o gás residual) das pesadas, que apresentam um maior valor comercial. Segundo Thomas et al. (2001), parte do gás residual produzido é consumido nas áreas produtoras para elevação artificial de petróleo (*gas-lift*), para a recuperação secundária de reservatórios através da injeção em poços ou ainda como combustível. Caso a produção de gás residual seja maior do que a demanda das áreas produtoras, o excesso é transferido (se economicamente viável) ou conduzido para os queimadores.

A unidade de processamento de gás natural em estudo faz parte de um grupo de três UPGNs localizadas na unidade de tratamento e processamento de fluidos (UTPF) de Guamaré-RN. Essas unidades recebem petróleo e gás natural provenientes dos campos de extração marítimos (*offshore*) e terrestres (*onshore*), transformando-os em produtos de consumo para o mercado.

A tabela 2.1 é o resultado da análise de uma amostra do gás natural que chega para ser processado no Polo Petroquímico de Guamaré [Lima 2001]. Essa composição é influenciada pela entrada de novos campos de produção e o fechamento de poços, o que pode causar alterações na proporção de contribuição dos compostos na mistura total.

Nas UPGNs, geralmente, o gás natural passa por processos para a remoção de água e elementos oxidantes e em seguida é encaminhado para uma série de colunas ou torres de destilação fracionada para ser decomposto. Além do gás residual citado acima, também são normalmente produzidos a gasolina natural e o gás liquefeito de petróleo (GLP), produto de maior interesse econômico.

Tabela 2.1: Composição de uma amostra do gás natural em Guararé-RN.

Componente	Fórmula química	Fórmula Simplificada	% em mol
Metano	$CH_4$	$C_1$	76,15
Etano	$C_2H_6$	$C_2$	10,64
Propano	$C_3H_8$	$C_3$	5,88
Butano	$C_4H_{10}$	$C_4$	2,44
Dióxido de carbono	$CO_2$	–	2,32
Nitrogênio	$N_2$	–	1,38
Pentano	$C_5H_{12}$	$C_5$	0,87
Hexano	$C_6H_{14}$	$C_6$	0,30
Heptano	$C_7H_{16}$	$C_7$	0,10
Octano	$C_8H_{18}$	$C_8$	0,01

### 2.3 Destilação fracionada

Para Luyben (1990) as colunas destilação fazem parte do processo mais importante e popular dos processos químicos. Apesar de estarem sendo estudados ao longo de várias décadas, os processos de destilação ainda encantam pesquisadores da área de controle de processos. Isso ocorre devido ao fato da destilação ser a técnica de separação mais comumente utilizada pelas indústrias químicas e petroquímicas.

Esse processo é realizado através da adição ou subtração de calor baseado nas diferentes temperaturas de ebulição dos compostos a serem separados. Na indústria do petróleo, a destilação é normalmente realizada em grandes colunas cilíndricas verticais, denominadas colunas de destilação fracionada [Linhares 2009]. A figura 2.2 representa um diagrama esquemático de uma dessas colunas que, em conjunto com o refeedor (*reboiler*) e o condensador, forma um sistema de destilação.

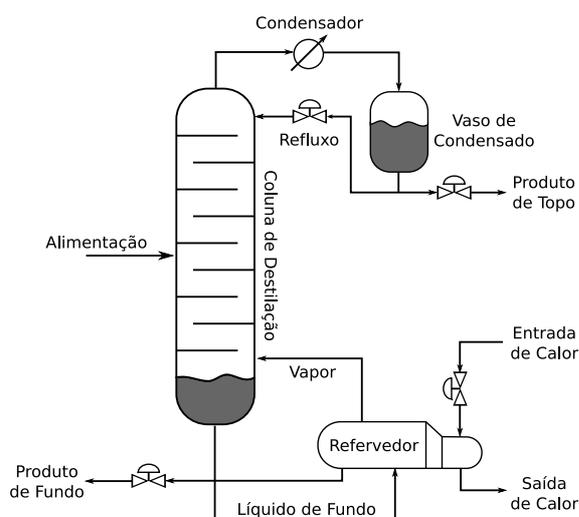


Figura 2.2: Coluna ou torre de destilação.

A coluna de destilação possui uma estrutura tubular cilíndrica de diâmetro variável com uma altura de 6 à 60 metros. Seu interior é normalmente preenchido com diversos pratos perfurados que determinam os “estágios” onde ocorre o contato entre líquidos e vapores. A alimentação da coluna é normalmente realizada em um dos estágios intermediários.

O refeedor é responsável pela vaporização do líquido de fundo a partir de um processo de troca de calor. Essa troca é realizada através de um fluxo de vapor que retorna à coluna e um fluxo de líquido que é retirado dela. O líquido retirado é denominado produto de fundo da coluna de destilação e é geralmente composto pelas frações mais pesadas da mistura.

Os componentes mais leves que atingem o topo da coluna na forma de vapor são resfriados pelo condensador através de um processo de troca de calor, geralmente utilizando água fria ou ar, e em seguida são armazenados no vaso de condensado. O líquido presente nesse reservatório dá origem a dois fluxos dos quais os componentes mais leves constituem o produto de topo da coluna e os mais pesados retornam para a coluna. Segundo Linhares (2009), esse último é conhecido como refluxo e é de fundamental importância para a eficiência da destilação, aumentando o rendimento do processo.

Os vapores originados do refeedor seguem ascendentemente e ao ultrapassar cada um dos estágios tornam-se mais ricos em componentes leves. Analogamente, o fluxo que segue descendentemente torna-se cada vez mais rico em frações pesadas. Para Lima (2001) esses fenômenos ocorrem devido ao equilíbrio líquido-vapor, como reflexo das diferentes temperaturas em cada estágio.

## 2.4 UPGN simulada

As colunas demetanizadoras, deetanizadoras, depropanizadoras e debutanizadoras são exemplos de colunas de destilação normalmente encontradas em UPGNs. Cada UPGN possui uma estrutura complexa e diferenciada de acordo com as características do gás natural que a alimenta e com seus produtos finais.

A UPGN-II GMR, simulada neste trabalho, é constituída por uma coluna deetanizadora em série com uma coluna debutanizadora. A unidade contém ainda uma coluna depropanizadora, mas essa não faz parte da operação rotineira da UPGN, entrando em operação apenas quando é necessário corrigir o nível de propano no ciclo de refrigeração da UTPF e, por essa razão, será desconsiderada nesse estudo. As duas colunas de interesse foram simuladas computacionalmente no *software* de processos químicos HYSYS®.

As colunas de destilação que serão apresentadas são do tipo contínuo, de forma que o processo é alimentado continuamente. Diferentemente da destilação em batelada, o processo contínuo possui um melhor rendimento [Júnior 2007].

### 2.4.1 Coluna deetanizadora

A primeira fase do processo de destilação se dá na coluna deetanizadora. Nessa, são esperados como produtos de topo e de fundo, respectivamente, o gás residual e o líquido de gás natural (LGN). A figura 2.3 representa um esquema da coluna simulada bem como

o conjunto de controladores PID e demais instrumentos. A nomenclatura adotada para cada um dos instrumentos segue a nomenclatura da planta real.

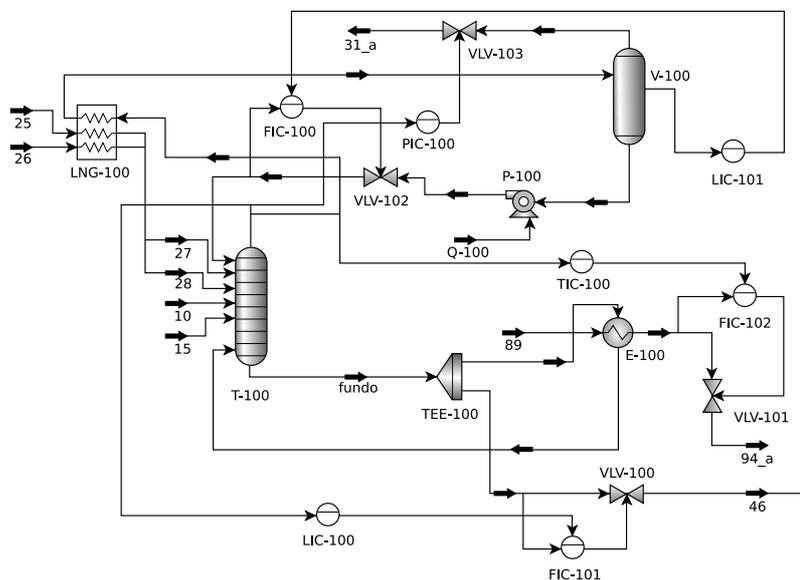


Figura 2.3: Coluna de destilação simulada.

A alimentação da coluna é feita em diversos pontos (linhas 10, 15, 27 e 28), fornecendo ao processo um tipo de gás composto por hidrocarbonetos variados ( $C_1+$ ).

Como explicado anteriormente, os componentes mais leves do gás atingirão o topo formando o produto de topo da coluna. Ao passar pelo trocador de calor LNG-100, parte desse produto é depositado no vaso de condensado V-100 e parte do líquido do vaso retorna à coluna. O controle da vazão de refluxo é realizado pelo controlador em cascata formado pelos controladores de nível LIC-101 (mestre) e de vazão FIC-100 (escravo). Baseado na pressão de topo da coluna, o controlador de pressão PIC-100 atua ajustando o nível de abertura da válvula VLV-103, determinando a quantidade de gás residual a ser liberado do vaso de condensado para comercialização.

Na parte inferior da coluna, os componentes mais pesados do gás de carga são encaminhados em sua forma líquida para o retervedor E-100 de modo que após a troca de calor o vapor resultante retorne à coluna. A troca de calor do E-100 é feita utilizando-se um fluxo de óleo térmico cuja vazão é controlada pela válvula VLV-101, que por sua vez é ajustada de acordo com o controlador de vazão FIC-102. Assim como o FIC-100 citado acima, o FIC-102 que opera em cascata, sendo escravo do controlador TIC-100.

O restante do líquido que não foi vaporizado compõe o LNG que é encaminhado para a próxima fase da destilação, a coluna debutanizadora. A vazão do LNG para a próxima coluna é regulada a partir do nível de líquido presente no fundo da coluna, sendo ajustada pelo controlador em cascata formado pelos controladores de nível LIC-100 (mestre) e de vazão FIC-101 (escravo). O LNG obtido é composto essencialmente por  $C_{3+}$  e  $C_2$  não vaporizado.



A destilação realizada pela coluna debutanizadora visa a obter como produto de topo os componentes mais leves de sua carga. No entanto, devido ao aquecimento da coluna, parte dos componentes pesados, principalmente os pentanos, acabam contaminando o GLP. Por outro lado, se a coluna não for aquecida suficientemente, os produtos leves não serão vaporizados, sendo depositados junto ao produto de fundo Júnior (2007). Além dos pentanos, o GLP também apresenta em sua composição os etanos presentes no LGN que, por serem agora expostos à temperaturas mais elevadas, são completamente retirados do processo.

O pentano presente no GLP é formado, principalmente, pela soma das frações molares do i-pentano e do n-pentano, que apresentam a mesma tendência de variação em função das condições de operação Linhares (2009).

Para que o GLP atenda as especificações de mercado, as frações molares de seus contaminantes devem ser mantidas entre pequenas faixas de aceitação preestabelecidas pela ANP. Na UPGN-II GMR, um dos poucos instrumentos capazes de realizar a medição das frações molares dos componentes do GLP, os cromatógrafos a gás, são compartilhados entre diversas linhas de produção e, por esse motivo, os intervalos de medição são normalmente grandes, tornando inviável a criação de estratégias de controle e supervisão mais sofisticadas. Ter acesso a essas informações permitiria controlar a qualidade do GLP, aumentando a lucratividade da UPGN.

# SISTEMA DE INFERÊNCIA

---

Neste capítulo serão mostrados detalhes de algumas ferramentas que foram utilizadas para a composição do sistema de inferência, tais como as redes neurais artificiais e análise de componentes principais.

Em seguida a atenção será voltada para o modelo utilizado para a identificação da dinâmica do processo. Através desse procedimento de identificação será possível realizar as inferências das frações molares do propano no gás residual e do etano e pentano no GLP.

A parte final do capítulo mostrará em detalhes quais foram as estruturas escolhidas para a coleta de resultados descrevendo o porque da utilização de cada uma delas.

## 3.1 Redes neurais artificiais

Segundo Haykin (2000), as redes neurais artificiais (RNAs) são estruturas paralelas, maciçamente distribuídas, constituídas por unidades simples de processamento conhecida como neurônios. Essas estruturas se assemelham ao cérebro humano devido a sua capacidade de “adquirir conhecimento” a partir do ambiente em que se encontra. Esse aprendizado ocorre através de um ajuste das forças de conexões, ou pesos sinápticos, que existe entre os neurônios. São essas conexões que armazenam os conhecimentos adquiridos pela rede.

Dentre as diversas aplicações das RNAs, podem ser citados exemplos para a classificação de padrões, filtragem de sinais, análise de imagens e controle/identificação de sistemas dinâmicos. Dentre as justificativas de utilização dessas estruturas citadas por Haykin (2000), destacam-se: sua característica intrínseca de não-linearidade, sua capacidade de generalização e adaptabilidade, a tolerância à falhas e a facilidade para realizar o mapeamento de relações entrada-saída.

Devido a grande complexidade existente em muitos problemas físicos reais, desenvolver um modelo matemático que represente adequadamente a dinâmica do processo é

uma tarefa praticamente impossível. Para Linhares et al. (2008), as RNAs, através de seu processo de aprendizagem e de sua capacidade de aproximação universal, conseguem representar a função correspondente à dinâmica do sistema com relativa simplicidade.

Essa capacidade de identificação se assemelha a funcionalidade de um sistema de inferência, que nada mais é do que um modelo capaz de reproduzir as relações dinâmicas entre variáveis secundárias e primárias.

Diversas são as arquiteturas de redes neurais existentes, tais como: as redes de funções de base radial, as redes de Kohonen, máquinas de vetor de suporte, entre outras [Linhares 2009]. Neste trabalho será utilizada a arquitetura de perceptron de múltiplas camadas (PMC), que já vem sendo aplicada com sucesso em diversos trabalhos realizados na área de inferência, como mostrado em Bawazeer & Zilouchian (1997), Bo et al. (2003) e Fortuna et al. (2005). A figura 3.1 representa um modelo esquemático dessa arquitetura.

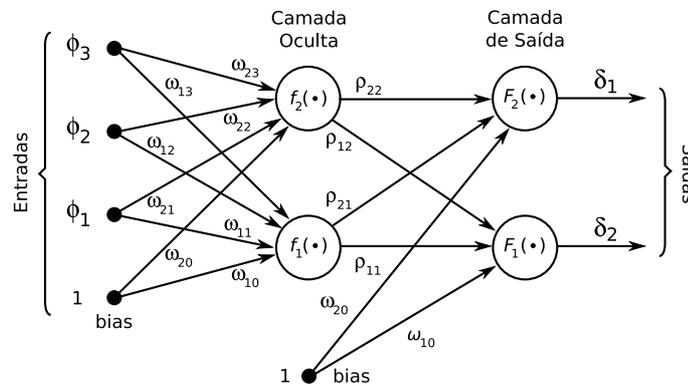


Figura 3.1: Rede PMC totalmente conectada de duas camadas.

Segundo Nørgaard et al. (2000), uma rede PMC básica possui seus neurônios dispostos em camadas recebendo como entrada as saídas dos neurônios da camada imediatamente anterior ou, no caso da primeira camada, as entradas da rede. Por possuir essa configuração essas redes são conhecidas como redes *feedforward*.

Como mostrado na figura 3.1, a segunda camada é conhecida como camada de saída, pois produz as saídas da rede neural. A primeira camada é conhecida como camada oculta ou intermediária por estar “escondida” entre as entradas da rede ( $\phi_1$ ,  $\phi_2$  e  $\phi_3$ ) e a camada de saída. Em Cybenko (1989) foi demonstrado que qualquer função contínua pode ser aproximada por uma rede neural PMC que possua uma camada oculta com funções de ativação sigmoideal ou tangente hiperbólica.

A equação 3.1 expressa matematicamente o funcionamento de uma rede neural PMC em que  $\delta_i$  representa a saída  $i$  da rede.

$$\delta_i(t) = g_i[\phi, \theta] = F_i \left[ \sum_{j=1}^{n_h} \rho_{i,j} f_j \left( \sum_{l=1}^{n_\phi} \omega_{j,l} \phi_l + \omega_{j,0} \right) + \rho_{i,0} \right] \quad (3.1)$$

O vetor de parâmetros  $\theta$  contém os pesos sinápticos e *biases* ( $\omega_{j,l}$ ,  $\rho_{i,j}$ ). Os neurônios da camada escondida e o número de entradas da rede são, respectivamente,  $n_h$  e  $n_\phi$ , enquanto que  $F_i$  e  $f_j$  são as funções de ativação dos neurônios das camadas de

saída e escondida, respectivamente. As funções de ativação utilizadas são mostradas nas equações 3.2, 3.3 e 3.4.

$$f_i(x) = x \quad (3.2)$$

$$f_s(x) = \frac{1}{1 + e^{-x}} \quad (3.3)$$

$$f_t(x) = \tanh(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (3.4)$$

### 3.1.1 Identificação através de modelo neural

Segundo Linhares (2009), a realização de inferência utilizando RNAs pode ser vista como um problema de identificação, uma vez que a rede neural treinada deve ser capaz de representar, satisfatoriamente, a dinâmica existente entre as variáveis secundárias e primárias do processo.

As estruturas de modelo baseadas em rede neural apropriadas para a identificação de sistemas não-lineares, são generalizações de modelos lineares. Para Lucena (2005), essas estruturas são caracterizadas por seu vetor de regressão, que nada mais é do que um vetor que contém as variáveis utilizadas para estimar a saída do sistema. Dependendo da escolha do vetor de regressão, diferentes estruturas de modelo neural podem surgir. Estruturas FIR (*Finite Impulse Response*), ARX (*AutoRegressive eXternal input*), ARMAX (*AutoRegressive Moving Average eXternal input*), OE (*Output Error*) e SSIF (*State Space Innovations Form*) são algumas das estruturas lineares mais conhecidas. Se o vetor de regressão for selecionado para modelos ARX, a estrutura do modelo será chamada NNARX (*Neural Network ARX*). Do mesmo modo, existirão também modelos NNFIR, NNARMAX, NNOE e NNSSIF.

Neste trabalho foi utilizado um modelo de rede baseado no NNARX descrito em Nørgaard et al. (2000). A figura 3.2 representa um esquema simplificado do modelo adotado. A presença de regressores no modelo, relaciona a saída da rede com seus valores passados de entrada e saída. A utilização desses regressores é de fundamental importância para a identificação de sistemas.

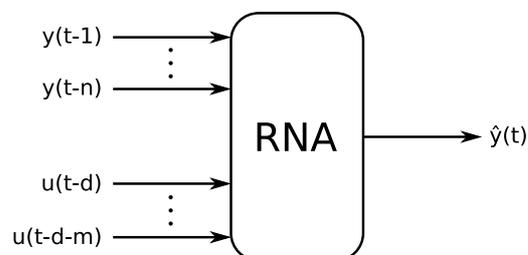


Figura 3.2: Esquema de uma rede neural com estrutura NNARX.

A expressão matemática que descreve o modelo não-linear pode ser descrita como representado na equação 3.5.

$$\hat{y}(t) = f(y(t-1), \dots, y(t-n), u(t-d), \dots, u(t-d-m)) \quad (3.5)$$

Nessa equação  $\hat{y}$  representa a saída estimada,  $d$  o atraso de transporte,  $n$  a ordem da saída,  $m$  a ordem de entrada da planta,  $f(\cdot)$  uma função não linear mapeada pela rede neural,  $y$  a saída da planta e  $u$  a entrada da planta.

A estimativa gerada pela estrutura NNARX é sempre estável, uma vez que representa relações puramente algébricas entre a estimativa e as medições passadas de entradas e saídas do processos, não existindo a realimentação da saída estimada.

Entretanto, como falado no final capítulo 2, as medições das frações molares das variáveis de interesse nem sempre estão disponíveis, pois os cromatógrafos apresentam longos intervalos de medição. Assim, se a estrutura original do NNARX fosse mantida, não seria possível reduzir o período de medição do cromatógrafo. Com o intuito de contornar esse problema, os valores passados estimados pelo próprio modelo são também utilizados como entrada da RNA. Esse fato pode fazer com que o sistema se torne instável ao longo do tempo, fornecendo estimativas das frações molares que divergem cada vez mais de seus valores reais. Em Linhares (2009) é proposto um sistema de correção do erro em tempo real baseado nas medições periódicas dos cromatógrafos de linha presentes no processo.

### 3.1.2 Determinação da ordem do modelo

Em sistemas de identificação *blackbox*, como o utilizado neste trabalho, é de fundamental importância que a ordem do modelo seja determinada adequadamente. Uma escolha inadequada poderá fazer com que a dinâmica do processo não seja assimilada pela rede neural de inferência, fazendo com que as estimativas geradas possuam erro médio consideravelmente alto. Segundo Arruda et al. (2003), existe uma ordem de modelo ótima que permite obter o menor erro entre o modelo estimado e o sistema real.

A ordem de entrada ( $m$ ) e de saída ( $n$ ) de um sistema desse tipo pode ser obtida a partir da realização de testes ou simulações com o processo. Por facilidade de representação, considerar-se-á a ordem de entrada igual a ordem de saída ( $m = n$ ). No capítulo 4 poderão ser vistos alguns resultados dos testes que foram realizados para a determinação da ordem do sistema proposto.

### 3.1.3 Seleção do modelo

Existem diversas estruturas de modelagem, cada uma com suas vantagens e desvantagens. Segundo Nørgaard et al. (2000), o problema da seleção das estruturas pode ser dividido em duas partes. A primeira delas é a parte da seleção da “família” de estruturas de modelagem. Dependendo do sistema, as estruturas podem ser: modelos lineares, redes PMC, redes de função de base radial, *wavelets* etc. Já a segunda, é a parte da seleção do subconjunto da família escolhida.

Resumidamente, após a seleção da estrutura, segue-se uma sequência etapas até que o modelo possa representar adequadamente o sistema, de acordo com algum critério específico, como por exemplo, o erro médio quadrático (EMQ) de estimativa, que é calculado a partir da equação 3.6 e utilizado no algoritmo de treinamento da RNA para determinar o ponto de parada (fim do treinamento).

$$\text{EMQ} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{y_i - \hat{y}_i}{y_i} \right)^2} \quad (3.6)$$

Nessa equação,  $N$  representa o número de amostras de validação,  $y_i$  o valor real e  $\hat{y}_i$  o valor estimado.

## 3.2 Análise de componentes principais

Segundo Jolliffe (2002), a análise de componentes principais (ACP) é, provavelmente, a mais antiga e conhecida técnica de análise de dados multivariáveis. Introduzida por Pearson (1901) e desenvolvida independentemente por Hotelling (1933), essa técnica só passou a ser largamente utilizada com o advento dos computadores, possuindo uma grande variedade de aplicações, tais como: compressão de dados, processamento de imagens, reconhecimento de padrões, entre outras.

Para Kano et al. (2001), as técnicas de ACP e mínimos quadrados parciais (MQP) vem sendo largamente utilizadas na modelagem, monitoramento e controle de processos quimiométricos, uma vez que podem lidar, sem problemas, com variáveis de entrada correlacionadas.

De acordo com Warne et al. (2004), o objetivo principal da ACP é o de reduzir a dimensão de um conjunto de dados constituído por um grande número de variáveis inter-relacionadas, tentando manter, tanto quanto possível, a variação do conjunto de dados originais.

Como mostrado em Mingoti (2005), essa redução é feita ao se mapear um sistema composto por  $p$  variáveis em  $k$  combinações lineares ( $k < p$ ) não correlacionadas, chamadas de componentes principais. Através desse mapeamento, o sistema de variabilidade do vetor aleatório composto das  $p$  variáveis originais pode ser aproximado pelo sistema de variabilidade do vetor aleatório que contém as  $k$  componentes principais. A qualidade dessa aproximação pode ser medida a partir da proporção da variância total das variáveis mantidas no sistema.

As  $k$  componentes principais que representarão o sistema podem ser obtidas através da matriz de covariância ou de correlação. A opção por uma dessas duas matrizes normalmente é feita de acordo com a discrepância causada pelas diferentes unidades de medida das variáveis do sistema. Nos casos em que as componentes principais obtidas são influenciadas pelas variáveis de maior variância, costuma-se utilizar a matriz de correlação à matriz de covariância. No entanto, o procedimento para obter as componentes principais é semelhante para ambos os casos.

### 3.2.1 As componentes principais

Seja  $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)'$  um vetor aleatório, com vetor de médias  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_p)'$  e matriz de covariância  $\boldsymbol{\Sigma}_{p \times p}$ , com respectivos autovetores normalizados  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p$ , isto é, os autovetores  $\mathbf{e}_i$  satisfazem as seguintes condições:

1.  $\mathbf{e}_i' \mathbf{e}_j = 0$  para todo  $i \neq j$
2.  $\mathbf{e}_i' \mathbf{e}_i = 1$  para todo  $i = 1, 2, \dots, p$
3.  $\boldsymbol{\Sigma}_{p \times p} \mathbf{e}_i = \lambda_i \mathbf{e}_i$  para todo  $i = 1, 2, \dots, p$

sendo o autovetor  $\mathbf{e}_i$  denotado por  $\mathbf{e}_i = (e_{i1}, e_{i2}, \dots, e_{ip})'$ . Considere o vetor aleatório  $\mathbf{Y} = \mathbf{O}'\mathbf{X}$ , em que  $\mathbf{O}_{p \times p}$  é a matriz ortogonal de dimensão  $p \times p$ , constituída dos autovetores normalizados da matriz  $\boldsymbol{\Sigma}_{p \times p}$ , isto é,

$$\mathbf{O}_{p \times p} = \begin{bmatrix} e_{11} & e_{21} & \dots & e_{p1} \\ e_{12} & e_{22} & \dots & e_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1p} & e_{2p} & \dots & e_{pp} \end{bmatrix} \quad (3.7)$$

Sendo assim, pode-se dizer que  $\mathbf{Y}$  é um vetor composto por  $p$  combinações lineares das variáveis aleatórias do vetor  $\mathbf{X}$ , que possui vetor de médias  $\boldsymbol{\mu}_y = \mathbf{O}'\boldsymbol{\mu}$  e matriz de covariância  $\boldsymbol{\Lambda}_{p \times p}$ , tal que,

$$\boldsymbol{\Lambda}_{p \times p} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_p \end{bmatrix} \quad (3.8)$$

Dessa forma, as variáveis aleatórias que compõem  $\mathbf{Y}$  são não correlacionadas entre si. Assim sendo, surge a idéia de se utilizar combinações lineares em  $\mathbf{Y}$ , como uma forma alternativa de se representar o vetor  $\mathbf{X}$ , reduzindo-se o espaço de variáveis.

As  $k$  combinações lineares (componentes principais) escolhidas para representar o vetor  $\mathbf{X}$  podem ser relacionadas com a “energia” do sistema e são escolhidas de acordo com os autovalores  $\lambda_1, \lambda_2, \dots, \lambda_p$ .

Na prática, a matriz  $\boldsymbol{\Sigma}_{p \times p}$  é desconhecida e precisa ser estimada através dos dados amostrais coletados. Em geral, a matriz  $\boldsymbol{\Sigma}_{p \times p}$  é estimada pela matriz de covariância amostral  $\mathbf{S}_{p \times p}$ . Considerando então  $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p$  como sendo os autovalores da matriz  $\mathbf{S}_{p \times p}$ , e  $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_p$ , os respectivos autovetores normalizados, a  $j$ -ésima componente principal amostral estimada é definida por:

$$\hat{\mathbf{Y}}_j = \hat{\mathbf{e}}_j' \mathbf{X} = \hat{e}_{j1} \mathbf{X}_1 + \hat{e}_{j2} \mathbf{X}_2 + \dots + \hat{e}_{jp} \mathbf{X}_p \quad (3.9)$$

com  $j = 1, 2, \dots, p$ , e a variância total (energia) da  $j$ -ésima componente principal amostral é dada por:

$$E_j = \frac{\hat{\lambda}_j}{\sum_{i=1}^p \hat{\lambda}_i} \quad (3.10)$$

Segundo Mingoti (2005), para se reduzir a dimensão de  $p$  para  $k$  componentes principais, deve-se adotar um critério de escolha. Como falado anteriormente, uma forma comum de se escolher as  $k$  componentes é através da porcentagem  $\gamma$  da variância total, em que  $0 < \gamma < 1$ . Assim sendo, busca-se o menor valor de  $k$ , tal que:

$$\frac{\sum_{i=1}^k \hat{\lambda}_i}{\sum_{j=1}^p \hat{\lambda}_j} \geq \gamma \quad (3.11)$$

### 3.3 Composição do sistema

Um dos principais pontos de um projeto de sistema de inferência é a seleção das variáveis secundárias que serão utilizadas para realizar a estimativa das variáveis primárias do processo. Em sistemas aplicados à colunas de destilação é comum escolher temperaturas de diferentes estágios da coluna, bem como da vazão de refluxo do condensador e a vazão de óleo térmico do refeedor.

Assim como a planta real, o sistema simulado neste trabalho apresenta poucos sensores de temperatura. Por esse motivo, apesar do *software* de simulação fornecer perfis completos de temperatura de ambas as colunas, optou-se por utilizar os valores de temperatura que possam ser obtidos através dos sensores existentes na UPGN-II GMR.

A partir das figuras 2.3 e 2.4, pode-se perceber que as variáveis secundárias restringem-se às variáveis de processos (PVs) dos controladores. A tabela 3.1 mostra quais foram as variáveis escolhidas.

Tabela 3.1: Variáveis de processo escolhidas.

Variável	Controlador	Variável secundária
$V_1$	PIC-100	Pressão de topo
$V_2$	FIC-100	Vazão de refluxo
$V_3$	TIC-100	Temperatura do estágio 40
$V_4$	FIC-101	Vazão de LGN
$V_5$	TIC-102-2	Temperatura do estágio 16
$V_6$	LIC-102-2	Volume líquido do estágio 28
$V_7$	FIC-101-2	Vazão de refluxo
$V_8$	LIC-100-2	Nível do condensado

De posse dessas informações, serão montadas diferentes estruturas que caracterizarão o sistema de inferência como um todo. De maneira geral, o sistema será composto por

um módulo de ACP e um módulo RNA, receberá como entrada as variáveis de processo e terá como saída as estimativas das variáveis primárias de interesse.

### 3.3.1 Razão de redução

Tendo definido como será composto o sistema, a partir da ordem do modelo e do número de variáveis de entrada e saída, pode-se definir a razão de redução como sendo:

$$R_r = 1 - \frac{k}{r(V_e + V_s)} \quad (3.12)$$

Nessa equação  $k$  representa o número de componentes principais,  $r$  a ordem do modelo,  $V_e$  e  $V_s$  o número de variáveis de entrada e saída do módulo ACP.

O valor da razão de redução será utilizado no capítulo 4 como uma forma de comparar a eficiência do módulo ACP quando a ordem do modelo aumenta. Quando multiplicado por 100% ( $R_r \times 100\%$ ), esse valor expressa a porcentagem de redução do número de variáveis originais para o número de componentes principais selecionadas.

## 3.4 Agregação dos módulos

Uma vez explicado o funcionamento dos módulos ACP e RNA, é fácil perceber que a agregação desses módulos fará com que as  $p$  variáveis do processo sejam reduzidas para  $k$  componentes principais, diminuindo assim a complexidade da rede neural de inferência. Essa rede, configurada como uma estrutura de identificação NNARX, deverá modelar as dinâmicas entre as variáveis secundárias e primárias do processo, realizando estimativas confiáveis a cada minuto.

Com o intuito de avaliar diferentes desempenhos do sistema, serão propostas 4 estruturas de agregação dos módulos ACP e RNA. É importante destacar que existem diversas combinações de agregação desses módulos que não estarão sendo avaliadas neste trabalho. A avaliação dos resultados para cada uma dessas propostas será feita ao longo do capítulo 4.

Assim como exposto na seção 3.1.2, por facilidade de representação, a ordem de entrada será igual a ordem de saída.

As figuras 3.3 a 3.6 representam os diagramas esquemáticos das 4 estruturas que estarão sendo avaliadas. Nas figuras, VS representará as variáveis secundárias (ou variáveis de processo), VP as variáveis primárias,  $n$  o instante de amostragem,  $r$  a ordem do modelo,  $m$  o número de variáveis secundárias,  $j$  o número de variáveis primárias e  $k$  o número de componentes principais. Dessa forma, o número de variáveis do processo pode ser definido como  $p = m + j$ .

### 3.4.1 Estrutura 1

Para que a complexidade do módulo RNA seja reduzida ao máximo, a primeira estrutura será composta por um módulo ACP que tem como entrada as variáveis primárias, secundárias e seus respectivos regressores.

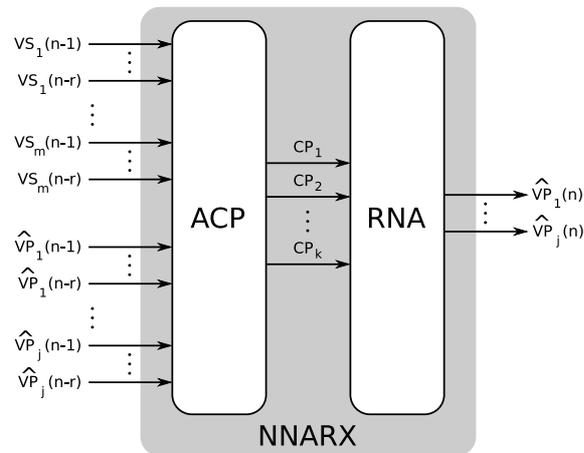


Figura 3.3: Esquema da estrutura 1.

Deseja-se saber aqui se o módulo ACP tem a capacidade de filtrar as informações necessárias dos regressores, visto que, esses são de fundamental importância para a identificação da dinâmica do sistema (seção 3.1.1).

Dependendo da capacidade de redução do módulo ACP, essa estrutura permitirá que o sistema seja testado com ordens consideravelmente elevadas, uma vez que a entrada do módulo RNA será composta apenas pelas  $k$  componentes principais selecionadas.

### 3.4.2 Estrutura 2

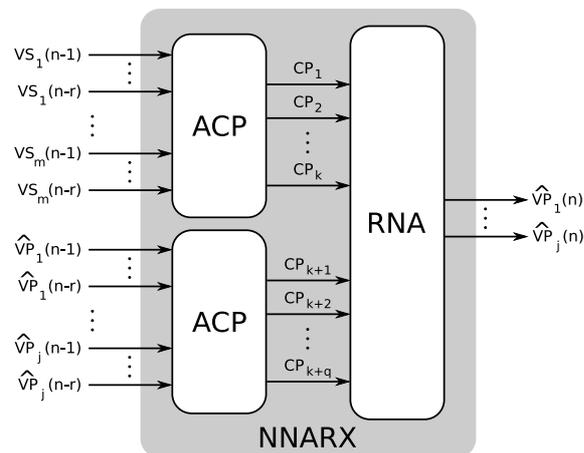


Figura 3.4: Esquema da estrutura 2.

Nessa estrutura, assim como na estrutura anterior, buscou-se reduzir a complexidade do módulo RNA ao máximo. Entretanto, com essa estrutura poderá ser analisada a capacidade dissociativa do módulo ACP.

A ideia é fazer com que as informações de entrada sejam separadas das informações de saída a partir de dois módulos ACP distintos.

### 3.4.3 Estrutura 3

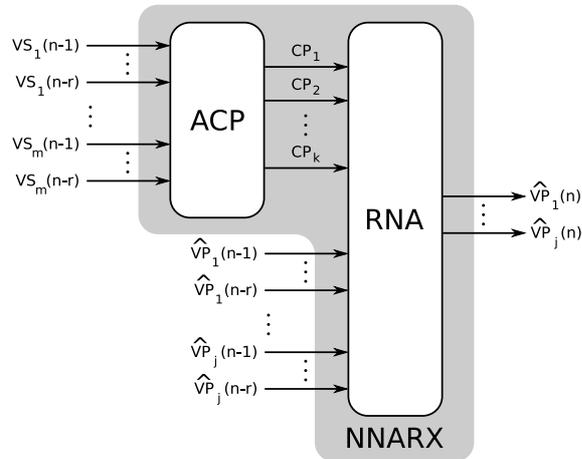


Figura 3.5: Esquema da estrutura 3.

Essa estrutura foi idealizada com o intuito de avaliar a importância das informações dos regressores de saída para a rede neural de inferência. Segundo Nørgaard et al. (2000), essas informações são de fundamental importância para que a rede assimile a dinâmica do processo e realize estimativas compatíveis com os valores reais.

Para que isso fosse possível, manteve-se os regressores das variáveis secundárias no módulo ACP e excluiu-se o módulo ACP dos regressores das variáveis primárias, o que fez com que a complexidade do módulo RNA aumentasse.

### 3.4.4 Estrutura 4

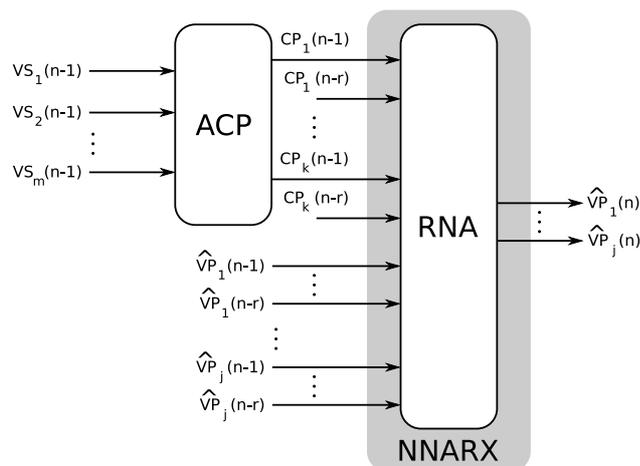


Figura 3.6: Esquema da estrutura 4.

Diferentemente das estruturas anteriores, essa proposta permite que o módulo RNA

permaneça o mais próximo possível da estrutura NNARX original, na qual os regressores de entrada e saída são conectados diretamente à rede.

Com essa estrutura o módulo ACP será capaz de realizar a redução somente das informações de entrada e não mais de seus regressores. Em teoria, isso fará com que as informações de entrada e saída sejam assimiladas mais facilmente pela rede, diminuindo o erro médio de estimativa. Por outro lado, a complexidade do módulo RNA aumentará significativamente.

# RESULTADOS

---

Neste capítulo será feita uma análise comparativa dos resultados das estruturas propostas no capítulo 3 com o intuito de realizar a inferência neural das frações molares do propano no gás residual bem como do etano e do pentano no GLP. No final do capítulo, a melhor das estruturas será escolhida para que seja feita uma análise detalhada de seus resultados.

## 4.1 Coleta dos dados

Na identificação de comportamentos dinâmicos, o primeiro passo a ser dado é a obtenção das amostras experimentais do processo que servirão para o treinamento supervisionado da rede neural de inferência.

A coleta de dados foi feita a partir da estimulação do sistema simulado (colunas de-etanizadora e debutanizadora) através da aplicação de sinais binários pseudo aleatórios (*pseudo random binary signal* – PRBS) aos *set points* dos controladores PIDs listados na tabela 3.1. Com esse procedimento foi possível proporcionar variações nas PVs desses controladores e, conseqüentemente, nas frações molares de  $C_3$  no gás residual e  $C_2$  e  $C_5$  no GLP.

Em todas as simulações foram mantidas as condições de segurança do processo, principalmente quanto aos limites de pressão e temperatura, permitindo que fosse obtido um grupo de dados com 3.000 amostras que incluam medições das variáveis secundárias e primárias do sistema simulado.

Os dados coletados foram aplicados ao módulo ACP, determinando as componentes principais que formaram as entradas do módulo RNA conforme ilustrado nas figuras 3.3, 3.4, 3.5 e 3.6.

## 4.2 A análise de componentes principais

Após a coleta dos dados foi possível obter a matriz  $S_{8 \times 8}$  da qual foram extraídos os autovalores  $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_8$  e seus respectivos autovetores normalizados  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_8$ . De posse desses valores, foram obtidas as componentes principais e, a partir do valor de  $\gamma$  da equação 3.11, foram selecionadas as  $k$  componentes de interesse.

Em Warne et al. (2004) é mostrado que a escolha do valor de  $\gamma$  é feita heurísticamente e que um dos critérios de escolha utilizados por ele é o critério de Jolliffe (2002). Nesse caso a porcentagem de variância total pode variar de 70 a 95%, estando diretamente relacionada com a quantidade de informação retida pelas componentes principais.

## 4.3 O treinamento da rede neural de inferência

O treinamento da rede neural de inferência foi realizado no *toolbox* de redes neurais do *software* matemático MATLAB<sup>®</sup> utilizando o algoritmo de treinamento de retropropagação do erro.

Ao final de cada etapa de treinamento a rede neural era submetida a 5 testes de validação para avaliar sua capacidade de generalização. A partir dos resultados encontrados concluiu-se que a estrutura NNARX de melhor desempenho é formada por uma única camada oculta com função de ativação sigmoideal. Os resultados obtidos para algumas das redes treinadas serão expostos nas seções seguintes.

## 4.4 A Ordem do modelo e o processamento dos dados

Assim como mostrado na seção 3.1.2 a ordem ( $r$ ) escolhida para o modelo influencia diretamente na inferência neural realizada pelo sistema. A escolha adequada dessa ordem permite que a rede neural de inferência assimile o comportamento da dinâmica do processo e realize estimativas confiáveis até mesmo na presença de ruídos.

Uma maneira simples, e muitas vezes eficiente, para determinar a ordem do modelo é através da realização de testes. Neste caso, existem duas formas para se fazer isso: iniciar os testes com uma ordem baixa ou com uma ordem alta, fazendo com que a ordem do sistema aumente ou diminua ao longo dos testes.

Neste trabalho a opção escolhida foi a segunda, com  $r_i = 4$  e  $r_f = 2$  para as estruturas 1 e 3, e  $r_i = r_f = 4$  para as estruturas 2 e 4. Para cada caso, exceto na estrutura 2, foram treinadas três redes, modificando apenas o número de neurônios da camada oculta. Cada rede foi treinada três vezes, para que a melhor delas fosse escolhida, e em cada treinamento eram feitas cinco validações.

A tabela 4.1 exibe um resumo das validações realizadas ao longo do treinamento. Nessa tabela, a porcentagem exibida entre parênteses para a estrutura 4 refere-se a porcentagem de variância total  $\gamma$  mostrada na equação 3.11. Esse valor é novamente utilizado na tabela 4.2, agora com o nome de “% Limite da ACP”, que mostra comparação da análise de componentes principais realizada sobre os dados coletados.

Tabela 4.1: Detalhes sobre as validações.

Estrutura	Número de ordens	Número de redes	Número de treinamentos	Número de validações	Total de validações
1	3	3	3	5	135
2	1	5	3	5	75
3	3	3	3	5	135
4 (84%)	1	3	3	5	45
4 (75%)	1	3	3	5	45
<b>Total</b>					<b>435</b>

Tabela 4.2: Comparação da ACP para as estruturas propostas.

Estrutura	% Limite da ACP	Ordem	$k$	$r(V_e + V_s)$	$R_r$
1	95%	4	10	44	0,77
		3	9	33	0,73
		2	9	22	0,59
2	95%	4	9 e 3	32 e 12	0,72 e 0,75
3	95%	4	8	32	0,75
		3	8	24	0,67
		2	7	16	0,56
4	84%	4	5	8	0,37
	75%	4	4	8	0,50

A partir da tabela 4.2 é fácil perceber que quanto maior a ordem do modelo maior será a razão de redução por parte do módulo ACP. Ou seja, o aumento da ordem do modelo fez com que a eficiência do módulo ACP também aumentasse, isolando as informações importantes e suas componentes principais e diminuindo a complexidade do módulo RNA.

## 4.5 Critério de escolha

Como mostrado na tabela 4.1, foram realizadas 5 validações para cada rede treinada. Em cada uma das 435 validações, as três variáveis de interesse eram estimadas, gerando, assim,  $3 \times 435 = 1305$  gráficos distintos que precisariam ser analisados.

Dessa forma, foi necessário estabelecer um critério de escolha para selecionar a melhor rede para cada estrutura. A forma escolhida foi a utilização do EMQ definido pela equação 3.6, avaliando cada uma das variáveis de interesse em separado.

## 4.6 Comparação das estruturas

Nas tabelas 4.3 a 4.6 as linhas em destaque se referem as melhores redes selecionadas para aquela ordem. Nessas tabelas,  $N_N$  representa o número de neurônios da camada

oculta,  $E_{et}$  o erro médio da validação de etano em porcentagem,  $E_{pr}$  o erro médio da validação de propano em porcentagem e  $E_{pe}$  o erro médio da validação do pentano em porcentagem.

#### 4.6.1 Estrutura 1

A tabela 4.3 mostra os EMQs de validação para as melhores redes da estrutura 1.

Tabela 4.3: EMQs Estrutura 1 – Ordem 2, 3 e 4 ( $r = 2$ ,  $r = 3$  e  $r = 4$ ).

Ordem	$N_N$	EMQ etano	EMQ propano	EMQ pentano	EMQ validação	$E_{et}$	$E_{pr}$	$E_{pe}$
2	5	6,91e-06	3,16e-10	5,06e-07	7,42e-06	-1,47	0,33	-0,48
2	10	6,54e-06	3,17e-10	4,55e-07	7,00e-06	-1,51	0,32	-1,15
2	15	6,60e-06	4,86e-10	4,46e-07	7,04e-06	-1,26	0,43	-0,07
3	5	6,39e-06	9,21e-10	5,15e-07	6,90e-06	-1,29	0,73	-0,13
3	10	8,61e-06	6,59e-10	4,57e-07	9,07e-06	-1,82	0,57	-0,27
3	15	7,79e-06	6,02e-10	6,14e-07	8,40e-06	-1,66	0,55	-0,06
4	5	3,21e-06	1,21e-09	5,75e-07	3,79e-06	-0,95	0,77	0,60
4	10	2,57e-06	1,30e-09	4,37e-07	3,01e-06	-0,39	0,84	0,24
4	15	2,90e-06	1,10e-09	5,19e-07	3,43e-06	-0,40	0,76	0,81

A partir da análise da tabela percebe-se que a melhor rede foi a rede com 10 neurônios e  $r = 4$ . As figuras 4.1 a 4.3 mostram a inferência realizada por essa rede para a validação com maior EMQ.

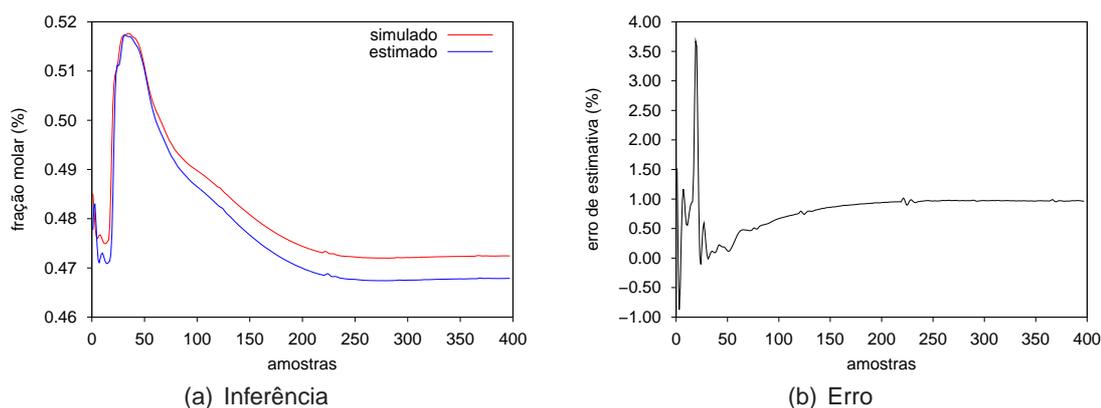
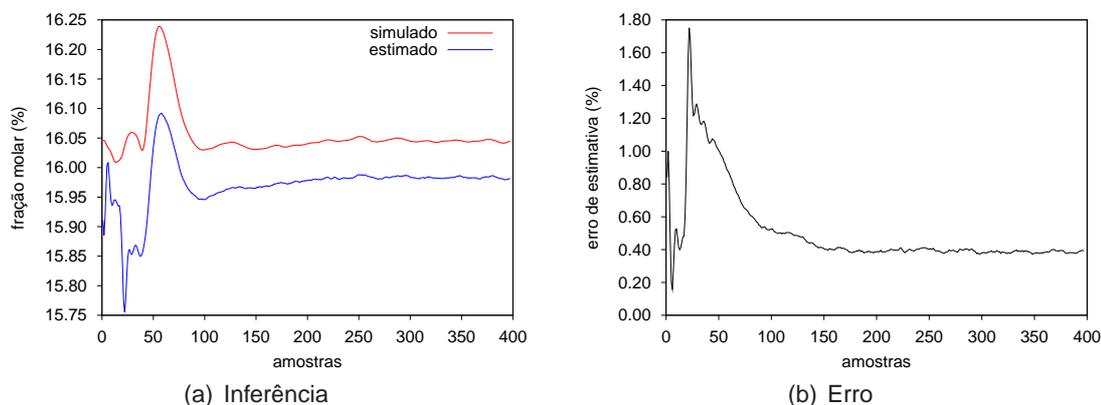
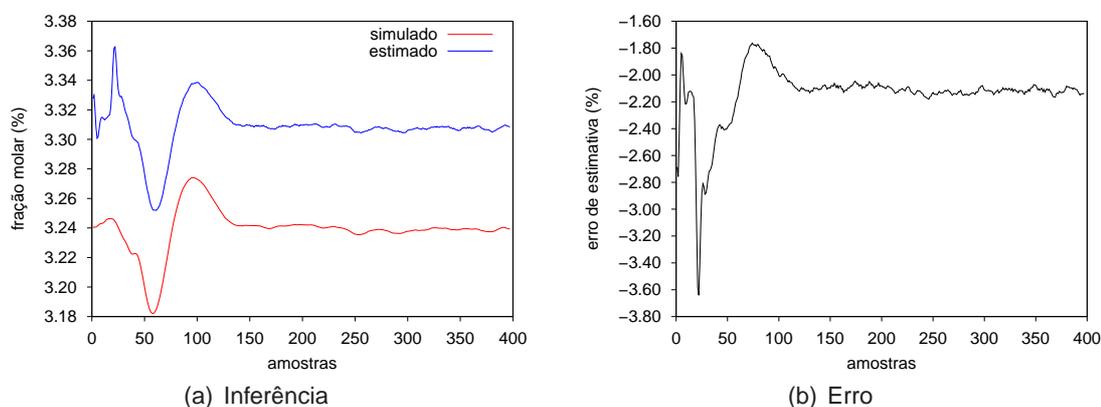


Figura 4.1: Estrutura 1 – Validação de propano ( $r = 4$ ).

Figura 4.2: Estrutura 1 – Validação de etano ( $r = 4$ ).Figura 4.3: Estrutura 1 – Validação de pentano ( $r = 4$ ).

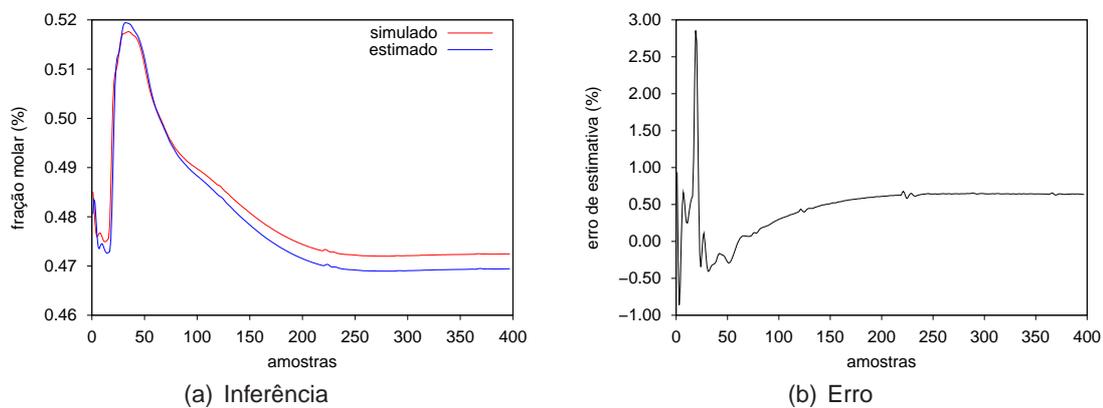
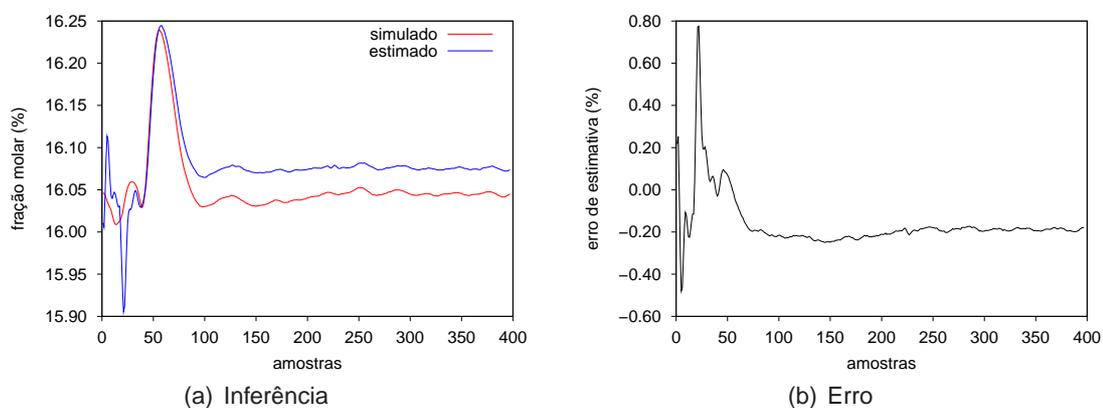
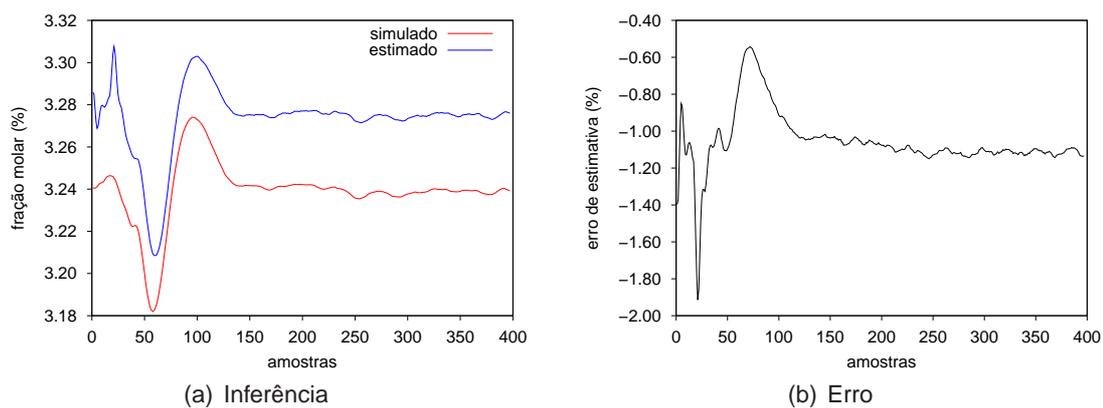
#### 4.6.2 Estrutura 2

A tabela 4.4 mostra os EMQs de validação para as melhores redes da estrutura 2.

Tabela 4.4: EMQs Estrutura 2 ( $r = 4$ ).

Ordem	$N_N$	EMQ etano	EMQ propano	EMQ pentano	EMQ validação	$E_{et}$	$E_{pr}$	$E_{pe}$
4	8	2,10e-06	1,21e-09	3,19e-07	2,42e-06	-0,56	0,80	0,69
4	12	1,98e-06	8,13e-10	2,69e-07	2,25e-06	-0,56	0,58	0,56
4	16	2,13e-06	1,40e-09	3,51e-07	2,49e-06	-0,37	0,86	0,84

A partir da análise da tabela percebe-se que a melhor rede foi a rede com 12 neurônios. As figuras 4.4 a 4.6 mostram a inferência realizada por essa rede para a validação com maior EMQ.

Figura 4.4: Estrutura 2 – Validação de propano ( $r = 4$ ).Figura 4.5: Estrutura 2 – Validação de etano ( $r = 4$ ).Figura 4.6: Estrutura 2 – Validação de pentano ( $r = 4$ ).

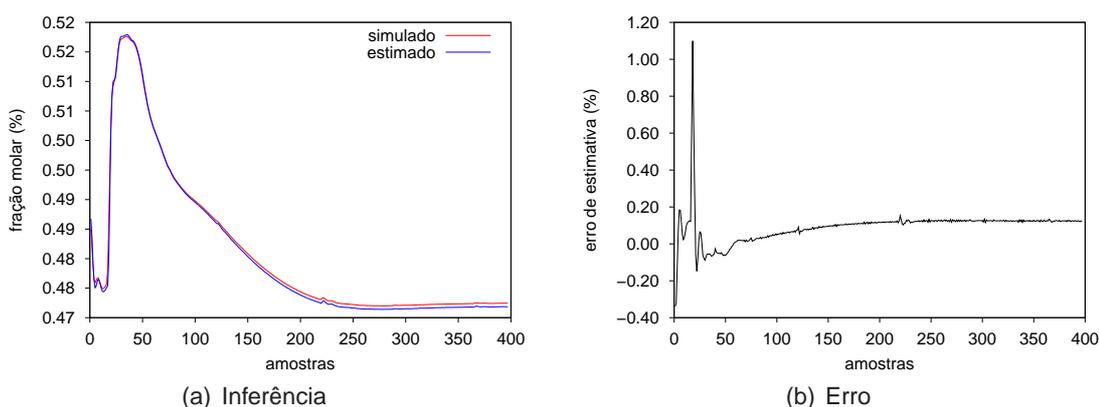
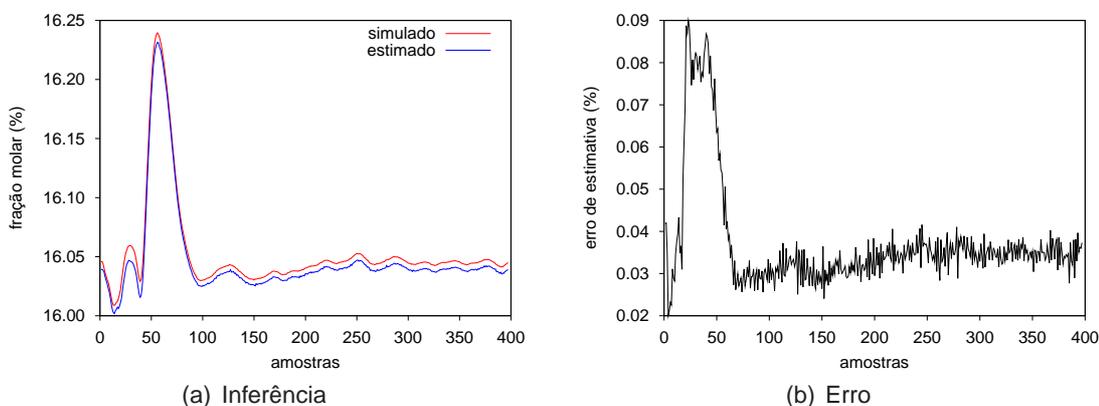
### 4.6.3 Estrutura 3

A tabela 4.5 mostra os EMQs de validação para as melhores redes da estrutura 3.

Tabela 4.5: EMQs Estrutura 3 – Ordem 2, 3 e 4 ( $r = 2$ ,  $r = 3$  e  $r = 4$ ).

Ordem	$N_N$	EMQ etano	EMQ propano	EMQ pentano	EMQ validação	$E_{et}$	$E_{pr}$	$E_{pe}$
2	10	3,73e-08	6,70e-11	5,57e-09	4,29e-08	-0,04	0,18	-0,05
2	13	5,95e-08	3,88e-11	3,48e-09	6,30e-08	-0,06	0,09	0,00
2	16	5,56e-08	3,48e-11	4,17e-09	5,98e-08	-0,07	0,01	-0,05
3	14	3,72e-08	3,94e-11	3,77e-09	4,11e-08	-0,07	0,09	0,08
3	18	5,74e-08	3,48e-11	6,04e-09	6,35e-08	-0,03	0,07	0,10
3	22	2,64e-08	2,82e-11	3,42e-09	2,98e-08	-0,02	0,04	-0,07
4	15	3,47e-08	2,70e-11	2,94e-09	3,77e-08	0,02	0,06	-0,09
4	20	3,47e-08	3,44e-11	1,80e-09	3,65e-08	0,00	0,06	0,02
4	25	2,48e-08	2,16e-11	8,79e-09	3,36e-08	-0,02	0,00	0,11

A partir da análise da tabela percebe-se que a melhor rede foi a rede com 20 neurônios e  $r = 4$ . As figuras 4.7 a 4.9 mostram a inferência realizada por essa rede para a validação com maior EMQ.

Figura 4.7: Estrutura 3 – Validação de propano ( $r = 4$ ).Figura 4.8: Estrutura 3 – Validação de etano ( $r = 4$ ).

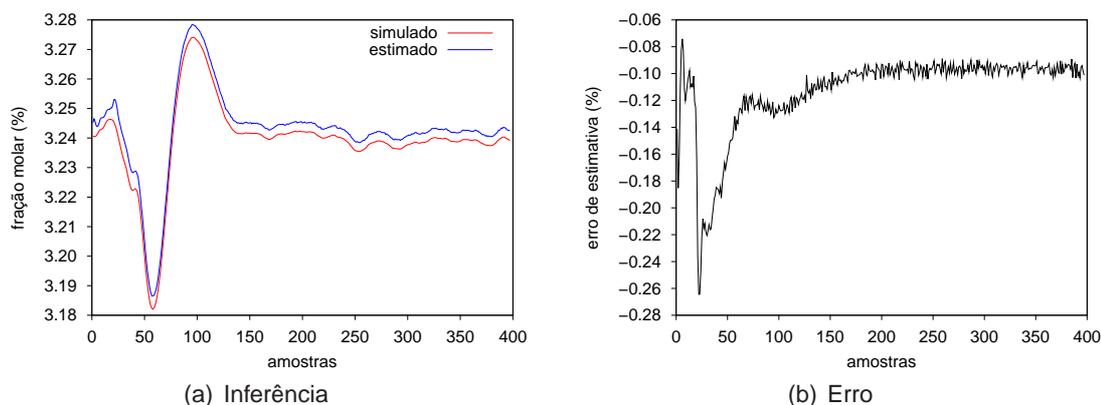


Figura 4.9: Estrutura 3 – Validação de pentano ( $r = 4$ ).

#### 4.6.4 Estrutura 4

Como pode ser observado na tabela 4.2 a estrutura 4 foi treinada para dois valores diferentes de  $\gamma$  (equação 3.11). A tabela 4.6 mostra os EMQs de validação para as melhores redes dessa estrutura com  $\gamma = 0,84$  e  $\gamma = 0,75$ .

Tabela 4.6: EMQs Estrutura 4 ( $r = 4$ ).

$\gamma$	$N_N$	EMQ etano	EMQ propano	EMQ pentano	EMQ validação	$E_{et}$	$E_{pr}$	$E_{pe}$
0,84	28	1,76e-08	1,54e-11	1,53e-09	1,91e-08	0,00	0,03	-0,08
0,84	32	1,48e-08	1,92e-11	9,05e-10	1,57e-08	0,02	0,07	0,03
0,84	36	3,32e-08	1,80e-11	3,78e-09	3,70e-08	0,04	0,06	-0,06
0,75	24	3,08e-08	1,26e-11	2,43e-09	3,32e-08	-0,01	0,02	-0,01
0,75	28	3,38e-08	1,78e-11	2,63e-09	3,64e-08	0,02	0,03	0,01
0,75	32	5,27e-08	1,34e-11	2,18e-09	5,49e-08	0,04	-0,02	-0,04

A partir da análise da tabela percebe-se que a melhor rede foi a rede com 28 neurônios, para  $\gamma = 0,84$  e com 24 neurônios, para  $\gamma = 0,75$ . As figuras 4.10 a 4.12 mostram a inferência realizada para a rede com  $\gamma = 0,84$ , enquanto que as figuras 4.13 a 4.15 mostram a inferência realizada para a rede com  $\gamma = 0,75$ .

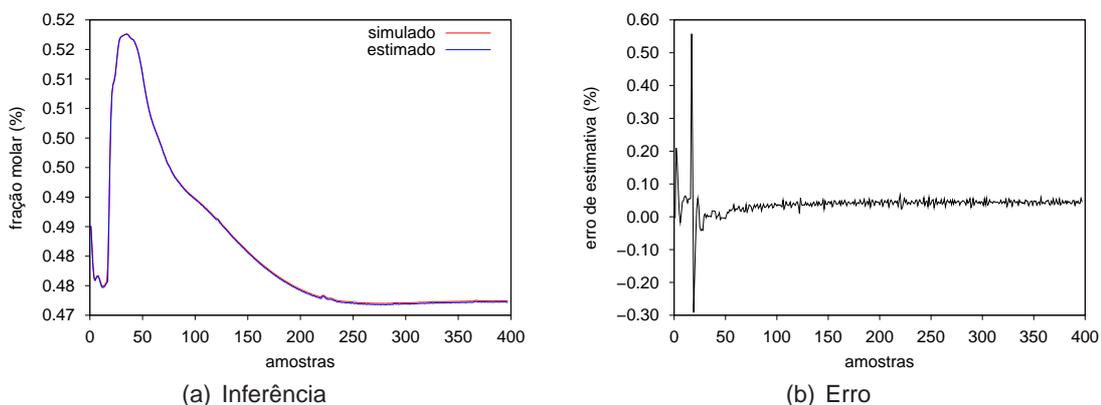


Figura 4.10: Estrutura 4 – Validação de propano ( $r = 4, \gamma = 0,84$ ).

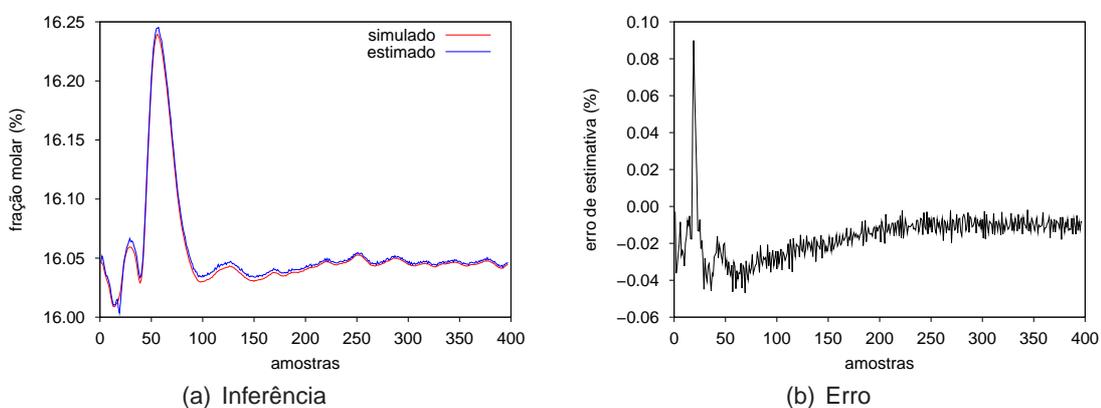


Figura 4.11: Estrutura 4 – Validação de etano ( $r = 4, \gamma = 0,84$ ).

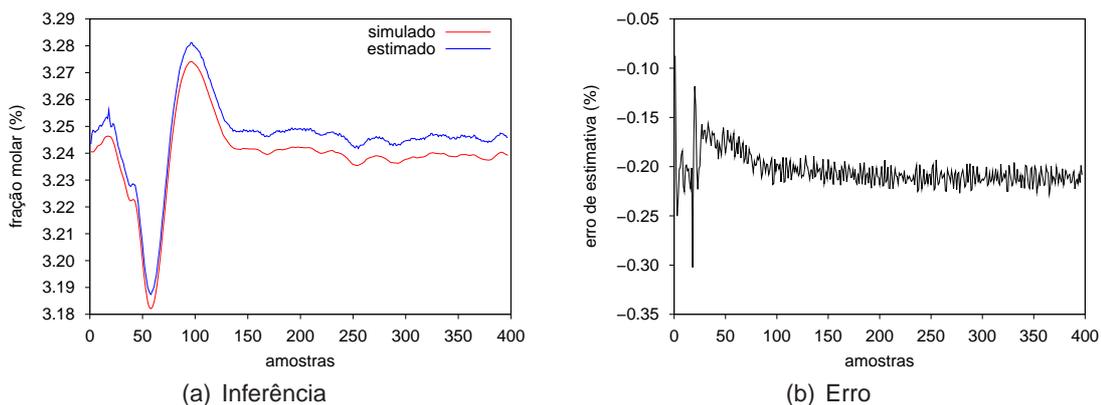


Figura 4.12: Estrutura 4 – Validação de pentano ( $r = 4, \gamma = 0,84$ ).

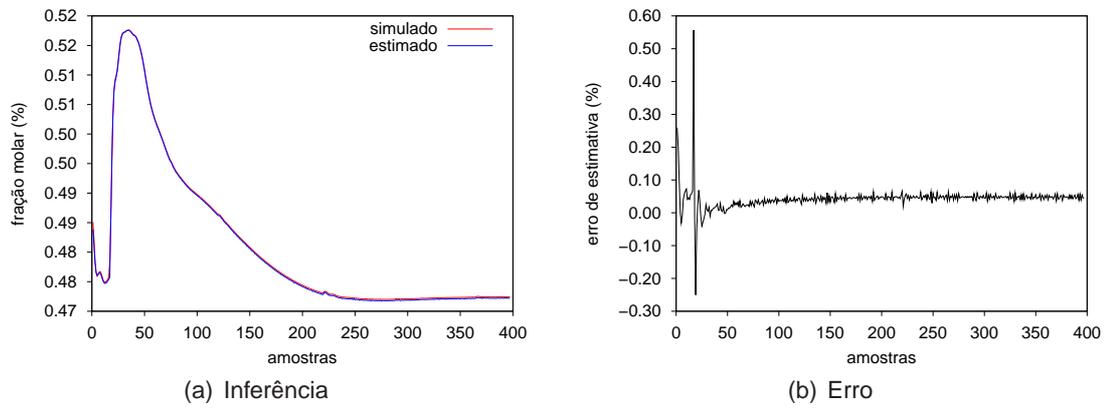


Figura 4.13: Estrutura 4 – Validação de propano ( $r = 4, \gamma = 0,75$ ).

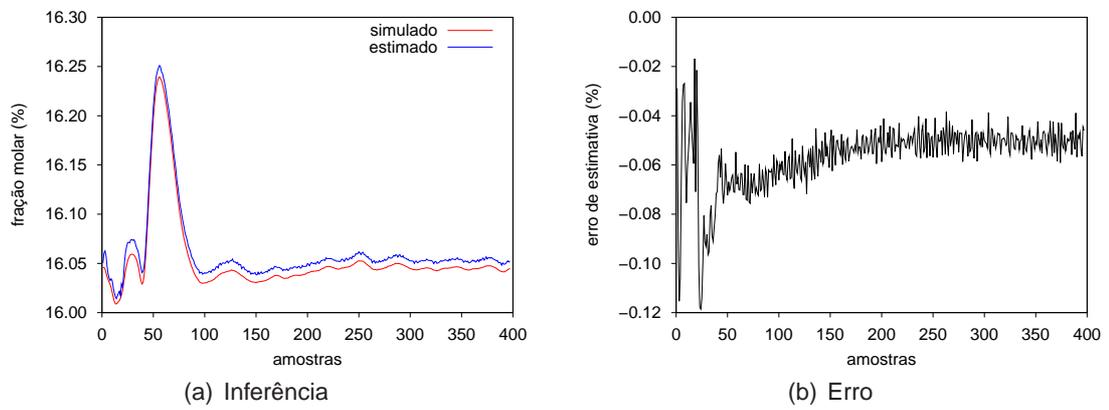


Figura 4.14: Estrutura 4 – Validação de etano ( $r = 4, \gamma = 0,75$ ).

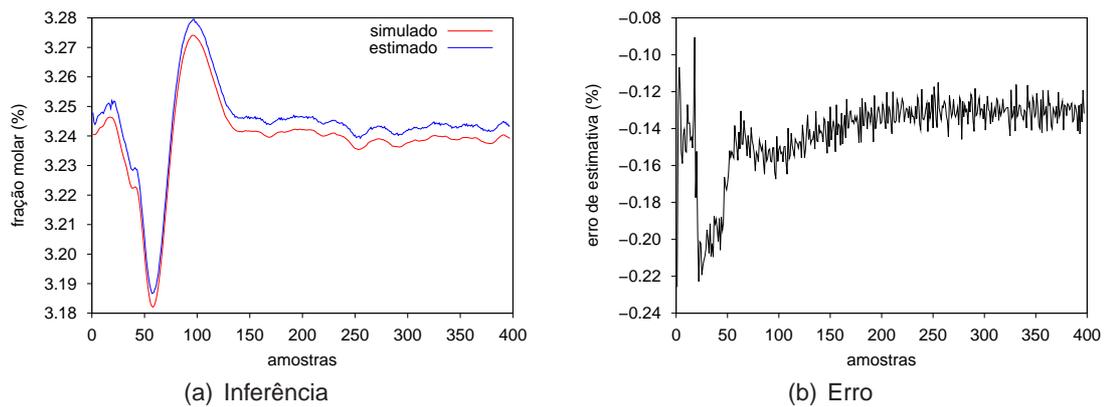


Figura 4.15: Estrutura 4 – Validação de pentano ( $r = 4, \gamma = 0,75$ ).

## 4.7 Análise das RNAs de quarta ordem

Considerando os valores expostos nas tabelas 4.3 a 4.6 pode-se concluir que as estruturas que melhor representam a dinâmica do processo são as estruturas de quarta ordem. A tabela 4.7 mostra as melhores redes dessa ordem. Nessa tabela o valor entre parênteses representa a porcentagem limite do módulo ACP ( $\gamma \times 100\%$ ).

Tabela 4.7: Melhores redes ( $r = 4$ ).

Estrutura	EMQ etano	EMQ propano	EMQ pentano	EMQ validação	$E_{et}$	$E_{pr}$	$E_{pe}$
1 (95%)	2,57e-06	1,30e-09	4,37e-07	3,01e-06	-0,39	0,84	0,24
2 (95%)	1,98e-06	8,13e-10	2,69e-07	2,25e-06	-0,56	0,58	0,56
3 (95%)	3,47e-08	3,44e-11	1,80e-09	3,65e-08	0,00	0,06	0,02
4 (84%)	1,76e-08	1,54e-11	1,53e-09	1,91e-08	0,00	0,03	-0,08
4 (75%)	3,08e-08	1,26e-11	2,43e-09	3,32e-08	-0,01	0,02	-0,01

Na tabela 4.8 é exposto um resumo das características do bloco RNA para cada uma dessas estruturas. Nessa tabela  $k$  representa o número de componentes principais selecionadas pelo módulo ACP,  $N_{No}$  o número de neurônios da camada oculta e  $T_T$  o tempo médio de treinamento em segundos.

Tabela 4.8: Características das RNAs ( $r = 4$ ).

Modelo	$k$	Número de entradas	$N_{No}$	Número de conexões	$T_T$
Sem ACP	0	44	46	2162	106,8016
Estrutura 1 (95%)	10	10	10	130	9,6958
Estrutura 2 (95%)	9 e 3	12	12	150	13,0291
Estrutura 3 (95%)	8	20	20	460	28,8024
Estrutura 4 (84%)	5	32	28	980	60,9591
Estrutura 4 (75%)	4	28	24	744	42,8787

## 4.8 Melhor estrutura e validação do sistema

A partir da análise das tabelas 4.7 e 4.8 pode-se concluir que as duas melhores estruturas propostas foram a estrutura 4 com  $\gamma = 0,75$  e a estrutura 3. O aumento de 40% do número de entradas e de 20% do número de neurônios da camada oculta, da estrutura 4 ( $\gamma = 0,75$ ) com relação a estrutura 3, refletiram em um aumento de 61,74% do número de conexões e 48,87% do tempo de treinamento. Por outro lado, ao considerar os EMQs expostos pela tabela 4.7, conclui-se que a melhor estrutura proposta é a estrutura 4.

Assim sendo, para confirmar a funcionalidade do sistema de inferência é necessário acoplá-lo ao sistema simulado, formado pelas colunas deetanizadora e debutanizadora, e

comparar as frações molares das variáveis primárias fornecidas pelo *software* simulador e pelo sistema de inferência proposto.

Na primeira análise realizada optou-se por alterar os valores do *set point* dos controladores de temperatura TIC-100 e TIC-102-2. Essa operação é considerada uma operação rotineira realizada por operadores humanos, ou por alguma estratégia de controle, uma vez que esses controladores influenciam direta ou indiretamente todo o processo.

Os *set points* desses controladores foram mantidos até a ducentésima amostra. A partir ducentésima primeira amostra esses valores foram alterados formando um degrau negativo. As figuras 4.16 a 4.18 apresentam uma comparação entre as variáveis primárias estimadas e as variáveis primárias fornecidas pelo *software* HYSYS®.

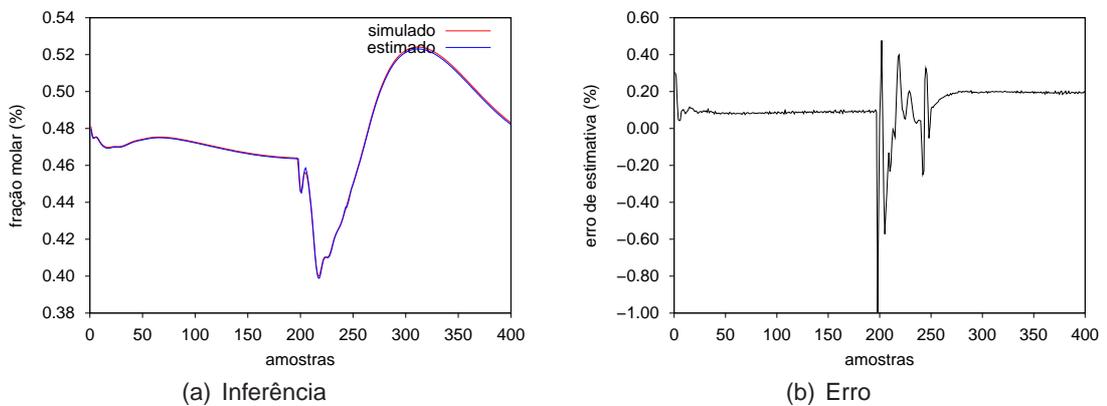


Figura 4.16: Validação de propano (Degrau negativo, Estrutura 4,  $r = 4$  e  $\gamma = 0,75$ ).

Na figura 4.16 percebe-se que a fração molar de propano no gás residual diminui rapidamente e aumenta lentamente logo em seguida. Isso ocorre devido a diminuição da vazão de óleo térmico ocasionada pelo fechamento da válvula VLV-101, que é controlada pelos controladores FIC-102 e TIC-100 em cascata. A sequência de ações faz com que a temperatura e, conseqüentemente, a quantidade de propano que sai no topo da coluna deetanizadora diminuam.

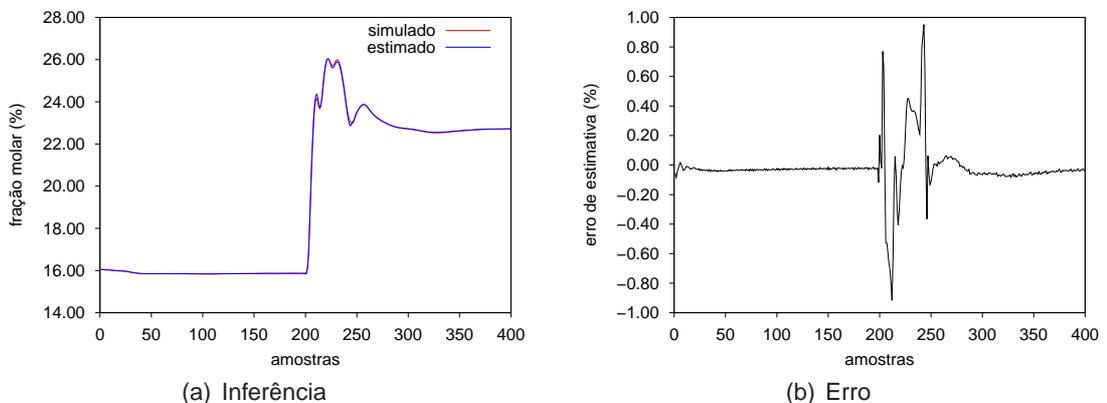


Figura 4.17: Validação de etano (Degrau negativo, Estrutura 4,  $r = 4$  e  $\gamma = 0,75$ ).

Na figura 4.17 percebe-se um aumento na quantidade de etano no GLP. Esse aumento ocorre devido a maior quantidade de etano que deixa de ser vaporizado a partir do esfriamento da coluna deetanizadora. Esse etano em excesso fica retido no LGN que alimenta a coluna debutanizadora, sendo então completamente vaporizado.

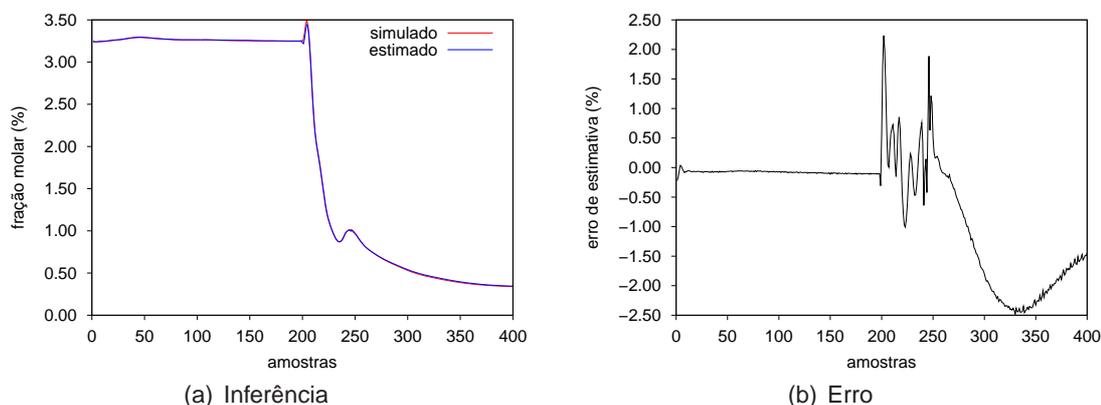


Figura 4.18: Validação de pentano (Degrau negativo, Estrutura 4,  $r = 4$  e  $\gamma = 0,75$ ).

Na figura 4.18 percebe-se a diminuição da fração molar de pentano no GLP. Isso ocorre devido a diminuição do *set point* do controlador TIC-102-2, o que faz com que a temperatura e, conseqüentemente, a quantidade de pentano produzido no topo da coluna debutanizadora diminuam.

Percebe-se que nos três casos o sistema de inferência conseguiu acompanhar satisfatoriamente a dinâmica do sistema simulado. O erro de estimativa, em valores absolutos, não chega a ultrapassar 1% para as duas primeiras inferências e 2,5% para a última.

Na segunda análise feita, buscou-se testar a robustez do sistema através da variação aleatória de todos os *set points* dos controladores da tabela 3.1. Esse caso, apesar de representar uma situação improvável quando as colunas estão operando normalmente, é ideal para avaliar o sistema de inferência. As figuras 4.19 a 4.21 apresentam uma comparação entre as variáveis primárias estimadas e as variáveis primárias fornecidas pelo software HYSYS®.

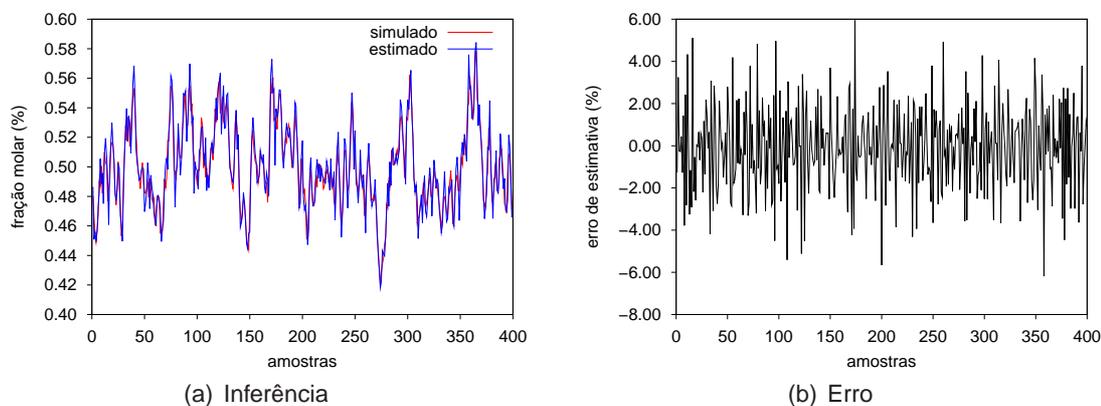


Figura 4.19: Validação de propano (PRBS, Estrutura 4,  $r = 4$  e  $\gamma = 0,75$ ).

Observa-se aqui que as variações bruscas do sistema simulado fizeram com que o erro de estimativa do sistema de inferência aumenta-se bastante quando comparado com o resultado anterior.

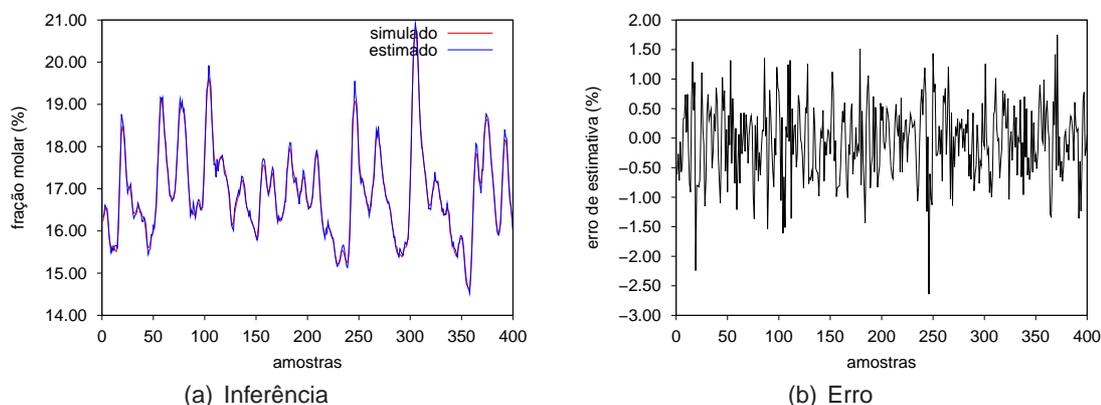


Figura 4.20: Validação de etano (PRBS, Estrutura 4,  $r = 4$  e  $\gamma = 0,75$ ).

A figura 4.20 mostra que as variações bruscas não influenciaram diretamente a variação da fração molar de etano no GLP.

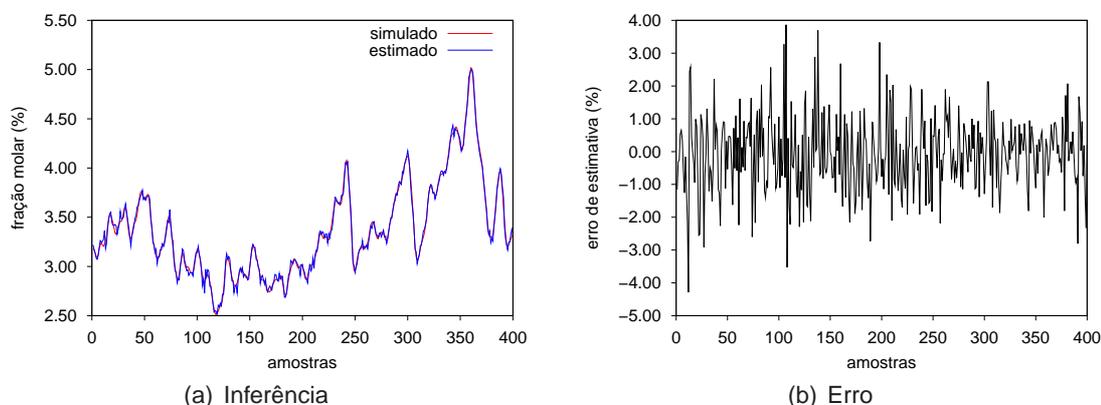


Figura 4.21: Validação de pentano (PRBS, Estrutura 4,  $r = 4$  e  $\gamma = 0,75$ ).

Assim como para a fração molar de etano, as variações bruscas também não influenciaram diretamente a variação da fração molar de pentano no GLP.

De maneira similar aos casos das figuras 4.16 a 4.18, pode-se perceber que o sistema de inferência conseguiu acompanhar a dinâmica do sistema simulado. Apesar de se tratar de uma situação incomum na prática, os erros de estimativa não ultrapassam 7% para o propano, 3% para o etano e 5% para o pentano.

# CONCLUSÕES

---

## 5.1 Conclusões

O presente trabalho foi desenvolvido com o intuito de atender a necessidade industrial da medição “instantânea” das frações molares de determinados componentes presentes nos produtos de topo das colunas de destilação da UPGN-II situada em Guamaré-RN. Para isso, a medição realizada devia fornecer valores de forma confiável e precisa e ter um baixo custo de manutenção e implementação.

Assim sendo, foi proposto um sistema baseado em técnicas de inteligência artificial, capaz de realizar a inferência de propano no gás residual de qualidade industrial, produto de topo da coluna deetanizadora, e de etano e pentano no GLP, produto de topo da coluna debutanizadora.

O sistema em questão é composto pelos módulos ACP e RNA que foram realimentados com os valores fornecidos pela simulação, não havendo risco de perda de estabilidade o que, conseqüentemente, faz com que não haja necessidade de um módulo de correção.

A estimativa realizada pelo sistema proposto conseguiu atender satisfatoriamente as expectativas, permitindo que possam ser desenvolvidas estratégias de controle avançado de forma a melhorar o desempenho do processo.

A partir dos resultados expostos no capítulo 4 percebe-se que, para a estimativa realizada, a redução do número de regressores das variáveis secundárias não afeta tanto quanto a redução do número de regressores das variáveis primárias. O motivo disso acontecer ainda é uma questão em aberto.

Além disso, observa-se também que, ao comparar a estrutura 3 com a estrutura 4 ( $\gamma = 0,75$ ), o aumento do número de neurônios da camada oculta refletiu em um aumento do número de conexões e, conseqüentemente, na complexidade da rede neural de inferência e no tempo de treinamento. Por outro lado, esse aumento poderá vir a beneficiar um módulo de correção que venha a ser implementado.

Dessa forma, fica comprovada a eficiência da utilização de redes neurais de múltiplas

camadas para a identificação de dinâmicas em sistemas complexos e não-lineares como o de uma UPGN.

## 5.2 Perspectivas

Além das estruturas 1, 2, 3 e 4 propostas neste trabalho, existem diversas outras possibilidades de agregação dos módulos ACP e RNA que podem vir a ser testadas. Pode-se testar ainda se o aumento da ordem do modelo fará com que diminua o erro de estimativa do sistema de inferência. Lembrando sempre que o aumento da ordem do modelo influi diretamente no aumento da complexidade do módulo RNA, seja por gerar mais componentes principais ou por aumentar o número de entradas da rede neural de inferência.

Outra possibilidade é fazer com que a configuração do sistema seja validada com os dados coletados de uma planta real, verificando a evolução comportamental do sensor virtual e a sua adequação às condições de operação. Para isso deverá ser implementado o módulo de correção de erro de maneira que o sistema, ao ser alimentado com as medições realizadas pelo cromatógrafo, realize a correção automática dos pesos sinápticos da rede neural de inferência, reduzindo assim a possibilidade de desestabilização.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

- Abdullah, Zalizawati, Norashid Aziz & Zainal Ahmad (2007), 'Nonlinear modelling application in distillation column', *Chemical Product and Process Modeling* **2**(12).
- Arruda, Lúcia Valéria Ramos, Flávio Neves Júnior & Sílvio Fávaro (2003), 'Identificação genética de modelos por pólos e zeros baseada no compromisso entre os erros de polarização e variância', *Revista Controle & Automação* **14**(2).
- Bahrtiya, Sharad & James R. Whiteley (2000), 'Development of inferential measurements using neural networks', *School of Chemical Engineering* **40**, 307–323.
- Baratti, Roberto, Alberto Bertucco, Alessandro Da Rold & Massimo Morbidelli (1998), 'A composition estimator for multicomponent distillation columns – development and experimental test on ternary mixtures', *Chemical Engineering Science* **53**(20), 3601–3612.
- Bawazeer, K. & Ali Zilouchian (1997), 'Prediction of products quality parameters of a crude fractionation section of an oil refinery using neural networks', *International Conference on Neural Networks* **1**(9-12), 157–162.
- Bo, C. M., J. Li, S. Zhang, C. Y. Sun & Y. R. Wang (2003), The application of neural network soft sensor technology to an advanced control system of distillation operation, em 'Proceedings of the International Joint Conference on Neural Networks', Vol. 2, pp. 1054–1058.
- Cybenko, George (1989), 'Approximation by superpositions of a sigmoidal function', *Mathematics of Control, Signals, and Systems (MCSS)* **2**(4), 303–314.
- Fortuna, Luigi, Pietro Giannone, Salvatore Graziani & Maria Gabriella Xibilia (2007), 'Virtual instruments based on stacked neural networks to improve product quality monitoring in a refinery', *IEEE Transactions instrumentation and Measurement* **56**(1), 95–101.

- Fortuna, Luigi, Salvatore Graziani & Maria Gabriella Xibilia (2005), 'Soft sensors for product quality monitoring in debutanizer distillation columns', *Control Engineering Practice* **13**, 499–508.
- Haykin, Simon (2000), *Redes neurais: princípios e prática*, Bookman – 2ª Edição.
- Hotelling, Harold (1933), 'Analysis of a complex of statistical variables into principal components', *Journal of Psychology* **24**, 417–441.
- Jolliffe, Ian (2002), *Principal Component Analysis*, Springer – 2ª Edição.
- Joseph, Babu (1999), A tutorial on inferential control and its applications, em 'Proceedings of the American Control Conference', San Diego, California.
- Júnior, José Medeiros Araújo (2007), Técnicas de inteligência artificial para a geração dinâmica de *set points* para uma coluna de destilação, Dissertação de mestrado, Universidade Federal do Rio Grande do Norte.
- Kano, Manabu, Natthaphon Showchaiya, Shinji Hasebe & Iori Hashimoto (2001), Inferential control of distillation compositions: selection of model and control configuration, Cheju Island, Korea, pp. 347–352.
- Lima, Marcos Ferreira (2001), Projeto e implementação de um sistema baseado em redes neurais para inferência da composição de pentano em uma torre de destilação, Dissertação de mestrado, Universidade Federal do Rio Grande do Norte.
- Linhares, Leandro Luttiane Silva (2009), Sistema de inferência baseado em redes neurais para controle de plantas de processamento de gás natural, Dissertação de mestrado, Universidade Federal do Rio Grande do Norte.
- Linhares, Leandro Luttiane Silva, Diogo Leite Rebouças & Fábio Meneghetti Ugulino de Araújo e André Laurindo Maitelli (2008), Estimação da fração molar de etano no produto de fundo de uma coluna deetanizadora a partir de um sistema neural de inferência, Salvador, Bahia.
- Lucena, Pedro Berretta (2005), Análise de um controlador baseado no jacobiano estimado da planta através de uma rede neural, Dissertação de mestrado, Universidade Federal do Rio Grande do Norte.
- Luyben, W. L. (1990), *Process Modeling, Simulation and Control for Chemical Engineers*, McGraw-Hill – 2ª Edição.
- Mejdell, Thor & Sigurd Skogestad (1991), 'Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression', *Industrial & Engineering Chemistry Research* **30**(12), 2543–2555.
- Mingoti, Sueli Aparecida (2005), *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*, Editora UFMG.

- Nørgaard, Magnus, Ole Ravn, Niels K. Poulsen & Lars K. Hansen (2000), *Neural Networks for Modelling and Control of Dynamic Systems*, Springer-Verlag.
- Oisiovici, Ronia Marques & Sandra Lúcia Cruz (2000), 'State estimation of batch distillation columns using an extended kalman filter', *Chemical Engineering Science* **55**, 4667–4680.
- Pearson, Karl (1901), 'On lines and planes of closest fit to systems of points in space', *Philosophical Magazine* **2**(6), 559–572.
- Pulis, Anna Maria (2008), *Soft Sensor Design for Distillation Columns*, Tese de doutorado, Università degli Studi di Cagliari.
- Thomas, José Eduardo, Attilio Alberto Triggia, Carlos Alberto Correia, Clodoveu Verrotti Filho, José Augusto Daniel Xavier, José Carlos Vieira Machado, José Erasmo de Souza Filho, José Luiz de Paula, Nereu Carlos Milani Rossi, Nilton Emanuel Santos Pitombo, Paulo Cezar Vaz de Melo Gouveia, Renato de Souza Carvalho & Roberto Vinicius Barragan (2001), *Fundamentos de Engenharia de Petróleo*, Editora Interciência – 2ª Edição.
- Tronci, Stefania, Fabrizio Bezzo, Massimiliano Barolo & Roberto Baratti (2005), 'Geometric observer for a distillation column: Development and experimental testing', *Industrial & Engineering Chemistry Research* **44**(26), 9884–9893.
- Venkateswarlu, C. & S. Avantika (2001), 'Optimal state estimation of multicomponent batch distillation', *Chemical Engineering Science* **56**, 5771–5786.
- Warne, K., G Prasad, N. H. Siddique & L. P. Maguire (2004), 'Development of a hybrid pca-anfis measurement system for monitoring product quality in the coating industry', *IEEE Internation Conference on Systems, Man and Cybernetics* .
- Zamprogna, Eliana, Massimiliano Barolo & Dale E. Seborg (2004), 'Optimal selection of soft sensor inputs for batch distillation columns using principal component analysis', *Journal of Process Control* **15**, 39–52.
- Zhang, Jie (2001), Inferential feedback control of distillation composition based on pcr and pls models, em 'Proceedings of the American Control Conference', Vol. 2, pp. 1196–1201.