



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
ENG07053 - TRABALHO DE DIPLOMAÇÃO EM ENGENHARIA
QUÍMICA



Avaliação de regiões operacionais de uma coluna de separação de propeno utilizando *self-organizing maps*

Autor: José Ricardo Furlanetto de Azambuja

Orientador: Marcelo Farenzena

Co-Orientadora: Andrea Cabral Farias

Porto Alegre, dezembro de 11

Sumário

Sumário	ii
Agradecimentos	iii
Resumo	iv
Lista de Figuras	v
Lista de Tabelas	vi
Lista de Símbolos	vii
Lista de Abreviaturas e Siglas	viii
1 Introdução	1
2 Revisão Bibliográfica	3
2.1 Redes Neurais Artificiais	3
2.2 Mapas auto-organizáveis	5
2.2.1 Inicialização dos pesos sinápticos	8
2.2.2 Processo de competição	9
2.2.3 Processo de cooperação	9
2.2.4 Processo adaptativo	10
3 Materiais e Métodos	12
3.1 Aspen Plus	12
3.2 MATLAB	13
3.2.1 Algoritmo para pré-processamento de dados	13
3.2.2 SOM Toolbox	13
4 Estudo de Caso	15
4.1 Unidade de Processamento de Gás Liquefeito de Petróleo	15
4.1.1 Coluna despropanizadora	15
4.1.2 Coluna desetanizadora	15
4.1.3 Coluna de separação propano e propeno	17
4.2 Construção da simulação em Aspen Plus	18
5 Resultados	20
5.1 Simulação em Aspen Plus	20
5.2 Matrizes-U	21
5.2.1 Matrizes-U para todos os dados	21
5.2.2 Matrizes-U após pré-processamento dos dados	23
5.3 Trajetórias aplicadas ao mapa escolhido	25
5.4 Predição da concentração de propeno no topo	27
6 Conclusões e trabalhos futuros	31
6.1 Trabalhos futuros	32
7 Referências	33

Agradecimentos

Gostaria de agradecer primeiramente e principalmente aos meus pais, José Luiz e Taís, e ao meu irmão, José Rodrigo, pelo apoio incondicional e por me propiciarem as condições necessárias para que o estudo sempre fosse a minha principal preocupação. Certamente seus ensinamentos sempre fizeram parte da minha formação.

Gostaria de agradecer também ao meu orientador professor Dr. Marcelo Farenzena pela oportunidade de participar do seu grupo de pesquisa e pelas numerosas horas de conselhos, ajudas e sugestões sem as quais este trabalho não teria sido possível. Gostaria de agradecer à minha co-orientadora Andrea Cabral Farias pelas valiosas contribuições e sugestões que foram imprescindíveis para o desenvolvimento desta dissertação.

Finalmente, gostaria de agradecer aos meus amigos, tanto os que me acompanharam em salas de aula quanto os que preservaram a amizade fora delas. Aos que me acompanharam nas salas de aula da UFRGS agradeço principalmente à RAÇA.

Resumo

A constante melhora e aumento do número de instrumentos de medição e bancos de dados de variáveis de planta têm-se mostrado como uma dificuldade no que se refere à interpretação dos mesmos. Dentro desse contexto, é apresentado um estudo sobre a técnica de self-organizing maps, ou mapas auto-organizáveis, aplicada em uma unidade de processamento de gás liquefeito de petróleo. A simulação das colunas desetanizadora e de separação propano e propeno foram construídas no software Aspen Plus a partir de dados de projeto fornecidos por uma refinaria nacional. A coluna de separação propano e propeno foi o foco do trabalho de identificação e avaliação das regiões operacionais e foi realizada a predição da concentração de propeno na corrente de topo através de uma etapa supervisionada, posterior ao treinamento do mapa. Tanto a identificação das regiões operacionais quanto a predição de variáveis apresentaram resultados muito satisfatórios. A identificação das regiões operacionais se fez através da marcação das trajetórias crescentes de variáveis operacionais da planta frente à concentração de propeno na corrente de topo. O coeficiente de correlação (R^2) para a predição realizada com o mapa treinado com tamanho grande foi de 0,9999.

Lista de Figuras

Figura 2.1: Representação básica de um neurônio. Extraído de Haykin (1999).....	3
Figura 2.2: Rede neural artificial com uma única camada. Extraído de Haykin (1999).	4
Figura 2.3: Rede neural artificial de múltiplas camadas. Extraído de Haykin (1999).....	4
Figura 2.4: Estrutura de uma rede neural artificial com aprendizagem não supervisionada.	4
Figura 2.5: Estrutura de uma rede neural artificial com aprendizagem supervisionada.	5
Figura 2.6: Exemplo de matriz-U. Fonte: Ng & Srinivasan (2008a).....	6
Figura 2.7: Exemplo de matriz-U combinada com BMU's. Fonte: Ng & Srinivasan (2008a).7	
Figura 2.8: Mapas topológicos hexagonal e retangular.	8
Figura 2.9: Função de vizinhança gaussiana.	10
Figura 4.1: Fluxograma da Unidade de Processamento de Gás Liquefeito de Petróleo.....	16
Figura 4.2: Fluxograma da simulação realizada em Aspen Plus.	18
Figura 5.1: Matrizes-U para diferentes tamanhos de mapa. (a) pequeno, (b) normal e (c) grande.	21
Figura 5.2: Matriz-U do mapa de tamanho 20x5.	22
Figura 5.3: Visualização da Matriz-U em 3 dimensões para o mapa de tamanho 20x5.	22
Figura 5.4: Matrizes-U para $T = 0,02$	23
Figura 5.5: Matrizes-U para $T = 0,04$	24
Figura 5.6: Matrizes-U para $T = 0,1$	24
Figura 5.7: Visualização da Matriz-U em 3 dimensões com aplicação do algoritmo, T igual a 0,02 e mapa de tamanho grande.....	25
Figura 5.8: (a) trajetória crescente de composição de propeno no topo e (b) trajetória crescente da fração da corrente direcionada para produto de topo.	26
Figura 5.9: (a) trajetória crescente da composição de propeno no topo e (b) trajetória crescente da vazão de <i>hot bypass</i>	26
Figura 5.10: trajetória crescentes (a) e (c) da composição de propeno no topo, (b) da pressão de descarga do compressor e (d) da carga de alimentação.	27
Figura 5.11: Resultado da predição para o mapa treinado com tamanho pequeno.....	28
Figura 5.12: Resultado da predição para o mapa treinado com tamanho normal.....	28
Figura 5.13: Resultado da predição para o mapa treinado com tamanho grande.....	29
Figura 5.14: Boxplot para os três diferentes tamanhos de mapas.	30

Lista de Tabelas

Tabela 5.1: Comparação dos dados de carga total, concentração molar de propeno e temperatura da simulação realizada em Aspen com os dados de projeto.	20
---	----

Lista de Símbolos

$w_{k,j}$ – peso sináptico do neurônio j para o dado de entrada k

$d_{j,i}$ – distância entre o neurônio j e o neurônio i

σ – largura efetiva da vizinhança topológica

$h_{j,i}$ – função de vizinhança topológica centrada no neurônio i para o neurônio j

n – tempo discreto

τ – constante de tempo

η – parâmetro taxa de aprendizagem

T - valor arbitrário para comparação entre dados

Lista de Abreviaturas e Siglas

SOM – Self-Organizing Maps

VLSI – Very-Large-Scale Implementation

BMU – Best Matching Unit

PCA – Principal Component Analysis

FCC – Fluid Catalytic Cracking

GLP – Gás Liquefeito de Petróleo

1 Introdução

Atualmente, devido ao avanço em sensores e tecnologias de banco de dados, na operação de uma planta muitas variáveis são medidas e registradas, e interpretá-las acaba sendo uma tarefa de alta complexidade. Com o propósito de facilitar a visualização destas variáveis surgem as técnicas de redução da dimensionalidade, que têm como objetivo representá-las de uma forma mais simples de ser interpretada. Outro problema recorrente na indústria são as operações multiestado, ou seja, a operação da planta pode transitar em diferentes regiões operacionais. Geralmente, regiões de transição entre estados representam operações em que a planta não está operando em sua capacidade ótima. Sendo assim, a redução da dimensionalidade e a identificação das diferentes regiões da planta se tornam muito importantes para operabilidade e lucratividade industrial.

No presente trabalho, será aplicada uma técnica que se propõe a anular esses dois grandes problemas. A técnica em questão é a de mapas auto-organizáveis, ou *self-organizing maps*, que é um tipo de rede neural de treinamento não supervisionado. Outro problema enfrentado nesse trabalho será que, às vezes, variáveis de planta tomam tempo para serem medidas, como de concentrações de correntes, por exemplo. Será proposto, dessa maneira, uma etapa posterior a de mapeamento de regiões operacionais a fim de prever variáveis de planta, através de um treinamento supervisionado.

Os mapas auto-organizáveis, não importando com quantas variáveis são alimentados, apresentarão sempre uma saída em duas ou três dimensões, facilitando assim a visualização de padrões operacionais. Tais mapas se diferem das outras redes neurais artificiais pelas suas etapas de competição e de cooperação. A etapa de competição se baseia na determinação do neurônio que mais se assemelha a um dado conjunto de variáveis alimentado, chamado de neurônio vencedor. Por sua vez, a etapa de cooperação atualiza com pesos diferentes neurônios mais próximos ao neurônio vencedor.

A visualização do mapa treinado é através da chamada matriz-U, que relaciona a distância entre neurônios vizinhos por uma escala de cores. Outra alternativa de visualização é marcar na matriz-U do mapa treinado os neurônios vencedores de um conjunto de amostras ordenado de forma crescente ou decrescente em relação a uma variável. Assim sendo, é possível ver a trajetória que determinada variável percorre na matriz-U.

Para uma boa aplicação da técnica de mapas auto-organizáveis se torna fundamental um vasto banco de dados de diferentes regiões de operação da planta. De modo a se obterem os dados necessários foi construída e validada com dados de projeto a simulação em Aspen Plus de uma unidade petroquímica em funcionamento. A unidade em questão pertence a uma refinaria nacional. De posse da simulação foram variadas diversas características operacionais, gerado o banco de dados e treinados os mapas para posterior identificação de regiões operacionais. A etapa seguinte foi o treinamento supervisionado do mapa auto-organizável, que é na sua concepção não-supervisionado, para a predição da concentração de propeno na corrente de topo da última torre de destilação da unidade que é responsável pela separação entre propano e propeno.

O presente trabalho está dividido em cinco seções. Na primeira parte é apresentada uma revisão bibliográfica sobre redes neurais artificiais e sobre mapas auto-organizáveis. Na segunda seção é apresentada a metodologia do trabalho e como os programas MATLAB e Aspen Plus foram utilizados para a obtenção dos objetivos do trabalho. Na seção seguinte foi descrito o caso de estudo do presente trabalho para uma melhor compreensão do problema a ser enfrentado. Na quarta parte são apresentados os resultados da simulação e dos mapas treinados na forma de visualização de diversas matrizes-U variando o tamanho do mapa e a utilização ou não do algoritmo de eliminação de amostras com dados similares, assim como as trajetórias crescentes de variáveis de operação da planta e a predição da composição de propeno na corrente de topo da coluna de separação propano e propeno. Na última seção estão apresentadas as conclusões obtidas ao longo do trabalho.

2 Revisão Bibliográfica

2.1 Redes Neurais Artificiais

O conceito de redes neurais artificiais surgiu na tentativa de recriar as conexões de um cérebro biológico utilizando ferramentas de um computador digital convencional. A partir disso, começou-se a imitar as interações entre os neurônios de um cérebro a fim de recriar esse mesmo padrão e organização em termos de processamento de dados. A grande vantagem de um sistema biológico é a experiência adquirida e a adaptabilidade ao meio ambiente em que ele está inserido, e ao estender-se aos computadores digitais esse conceito foi chamado de etapa de aprendizagem.

Segundo Haykin (1999), os maiores benefícios de uma rede neural artificial são: não-linearidade, mapeamento de entrada-saída, adaptabilidade, resposta a evidências, informação contextual, tolerância a falhas, implementação em VLSI (*Very-Large-Scale Implementation*), uniformidade de análise e projeto e analogia neurobiológica. VLSI consiste em criar circuitos integrados combinando milhares de transistores em um único chip, aumentando o poder computacional enormemente.

A unidade básica de uma rede neural artificial assim como a de uma rede neural biológica é o neurônio. No caso da artificial o neurônio é alimentado pelos dados de todos neurônios da camada anterior multiplicado por um valor chamado de peso sináptico. O peso sináptico, denominado por w , é atualizado conforme a alimentação de novos dados de entrada e varia de 0 a 1. Os valores, depois de serem multiplicados pelo peso sináptico, são somados e alimentados no neurônio, que por sua vez tem uma função de ativação escolhida pelo usuário, normalmente uma função tangencial hiperbólica. Outros exemplos de função de ativação são função linear, degrau e rampa. A representação de um neurônio típico é mostrada na Figura 2.1.

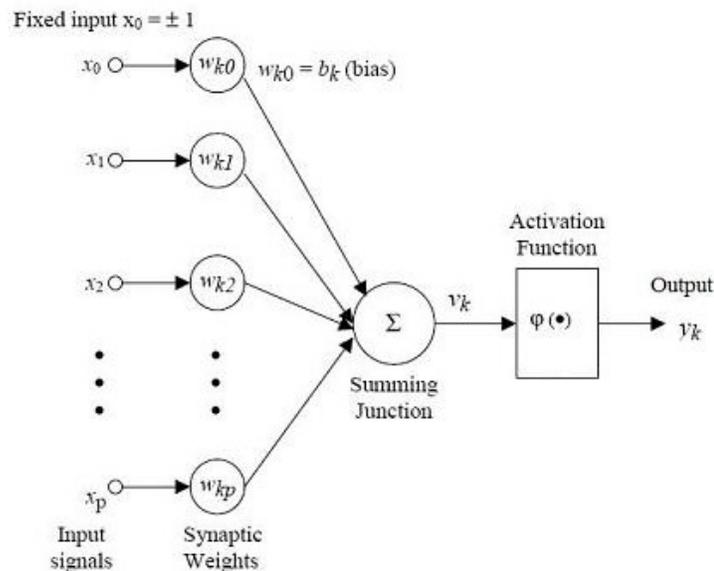


Figura 2.1: Representação básica de um neurônio. Extraído de Haykin (1999).

Os neurônios são dispostos em uma rede neural artificial em camadas de entrada, de saída e ocultas. Quando uma rede artificial só apresenta uma camada é chamada de *rede de camada única*, assim os dados de entrada são alimentados direto na camada de saída, como pode ser visto na Figura 2.2. No caso de haver camadas ocultas de neurônios, ela é

chamada de *rede com múltiplas camadas*, e torna-se importante quando é desejado que a rede extraia relações altamente não-lineares entre os dados e consiga fazer uma conexão adequada entre a camada de entrada de dados e a camada de saída. Uma rede com múltiplas camadas típica pode ser vista na Figura 2.3.

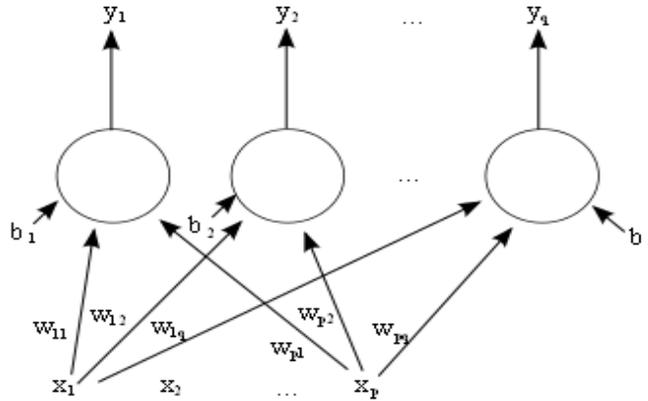


Figura 2.2: Rede neural artificial com uma única camada. Extraído de Haykin (1999).

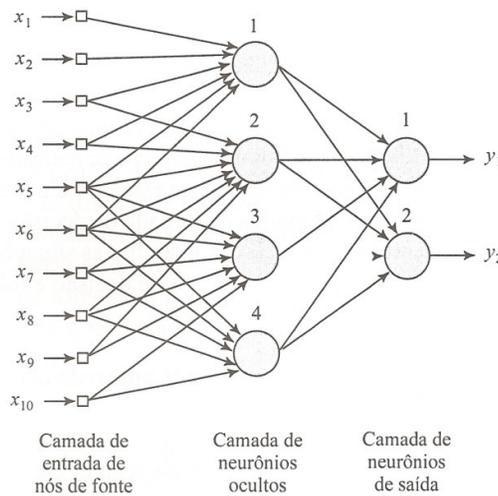


Figura 2.3: Rede neural artificial de múltiplas camadas. Extraído de Haykin (1999).

A etapa de aprendizagem pode ser feita de forma supervisionada ou não supervisionada. Na aprendizagem não supervisionada não há a resposta desejada para os dados de entrada alimentados ao sistema. Sendo assim, os pesos sinápticos dos neurônios são atualizados conforme a rede consegue detectar padrões do ambiente, levando em conta a similaridade dos dados na camada de entrada. A estrutura dessa rede pode ser vista na Figura 2.4. Nota-se que a saída do neurônio no sistema de aprendizagem é realimentado diretamente, sem passar por um valor base desejado.

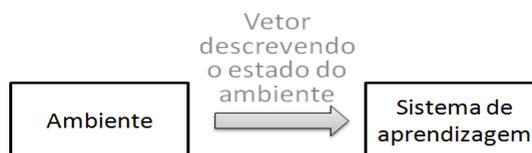


Figura 2.4: Estrutura de uma rede neural artificial com aprendizagem não supervisionada.

Caso o método escolhido seja supervisionado é preciso ter uma resposta do sistema desejada para cada conjunto de dados de entrada. Esse conhecimento sobre o sistema é chamado de *professor*, e a diferença entre o valor de saída calculado pela rede neural e o valor desejado é realimentado na rede a fim de que os pesos sinápticos dos neurônios sejam atualizados e atinjam o valor mais próximo possível do que se quer. Ao longo do tempo, quando os pesos sinápticos já tiverem sido iterados e os neurônios tenham aprendido as características da camada de entrada, é possível prescindir o *professor*. Um exemplo de uma estrutura básica de uma rede supervisionada pode ser visto na Figura 2.5.

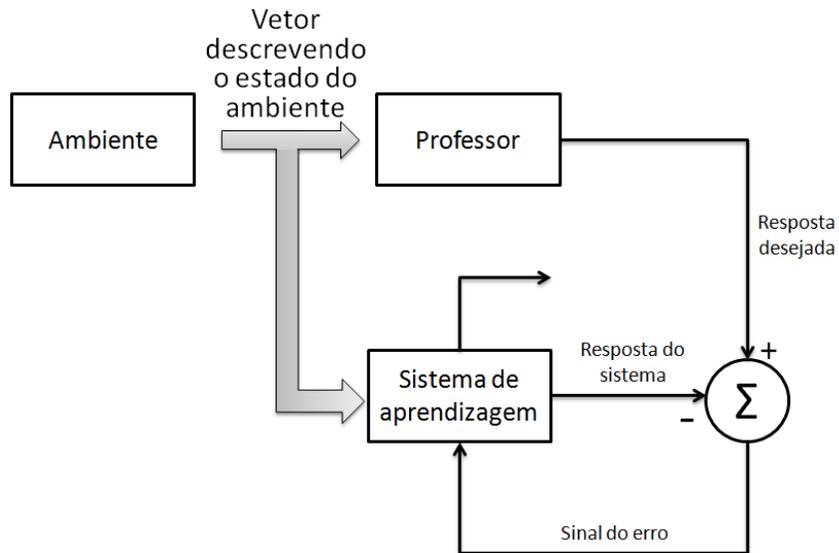


Figura 2.5: Estrutura de uma rede neural artificial com aprendizagem supervisionada.

2.2 Mapas auto-organizáveis

Os mapas auto-organizáveis, ou self-organizing maps (SOM), foram desenvolvidos no final da década de 1980 por Teuvo Kohonen e são um tipo de rede neural artificial não supervisionada, ou seja, não é preciso o conhecimento do ambiente em que foram gerados os dados a serem alimentados na rede. Como resultado, é gerado um mapa bidimensional que relaciona as distâncias entre os neurônios do mapa já treinado, permitindo assim uma redução da dimensionalidade dos dados de entrada. Relações entre esses dados são facilmente vistas no mapa, enquanto que nos dados originais seriam mais difíceis quanto maior fosse o número de variáveis.

Outras características importantes dos mapas auto-organizáveis, além da redução da dimensionalidade, são a sua não-linearidade, a técnica de aprendizagem competitiva e a identificação de agrupamentos ou clusterização. Em relação a sua não-linearidade, é visto como uma generalização da análise dos componentes principais ou PCA. A técnica de aprendizagem competitiva se baseia principalmente em não atualizar todos os neurônios da rede, para isso é definido um “neurônio vencedor”, que é o que mais se assemelha a determinado conjunto de dados de entrada e só ele e seus vizinhos são atualizados. A identificação de agrupamentos vem da interpretação do mapa gerado, e são facilmente vistos quando há uma região de neurônios com pequena distância entre eles.

Uma desvantagem inerente ao uso de mapas auto-organizáveis é que o mapa gerado pela técnica, sem que se conheça muito bem a origem dos dados de entrada, não é muito

informativo, ou seja, é preciso que alguém com conhecimento prévio reconheça os padrões apresentados no gráfico. Além disso, para a predição de dados de saída, é necessário uma etapa supervisionada posterior ao treinamento dos neurônios pelo algoritmo de SOM, com a vinculação dos dados de entrada com um valor esperado de saída.

O mapa treinado pelo algoritmo de SOM pode ser visto de duas maneiras diferentes que são a matriz-U e a visualização do “best matching unit” (BMU) de determinado conjunto de dados de entrada. A matriz-U relaciona a distância entre os neurônios da grade e a sua vizinhança, sendo fácil desse jeito visualizar agrupamento de padrões de entrada, representados por regiões com pequena distância entre os neurônios adjacentes e regiões de transição, representadas por grandes distâncias. Na Figura 2.6 pode ser visto uma matriz-U retirada de Ng & Srinivasan (2008a) onde três clusters, apontados por C1, C2 e C3 são facilmente identificados.

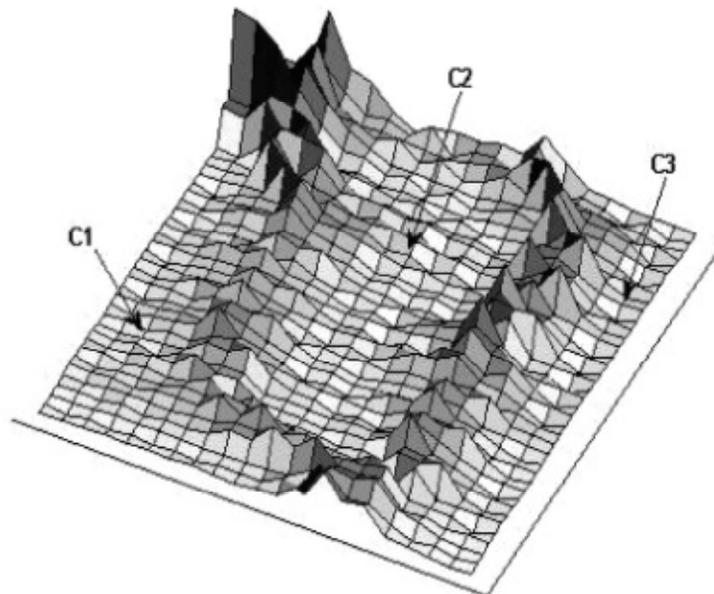


Figura 2.6: Exemplo de matriz-U. Fonte: Ng & Srinivasan (2008a).

Uma desvantagem de utilizar a matriz-U é que nenhuma noção temporal pode ser retirada dela, somente adicionando simultaneamente as informações dos neurônios vencedores. Para tanto, toma-se uma sequência temporal de dados e determina-se onde estão os neurônios vencedores para cada conjunto de dados. Assim sendo, pode-se ver com facilidade a trajetória que os dados percorreram no mapa topológico treinado por SOM durante o tempo de amostra. Na Figura 2.7 se vê a trajetória em azul percorrida por um conjunto de dados de entrada simulado pelos autores organizados temporalmente no mesmo mapa da Figura 2.6.

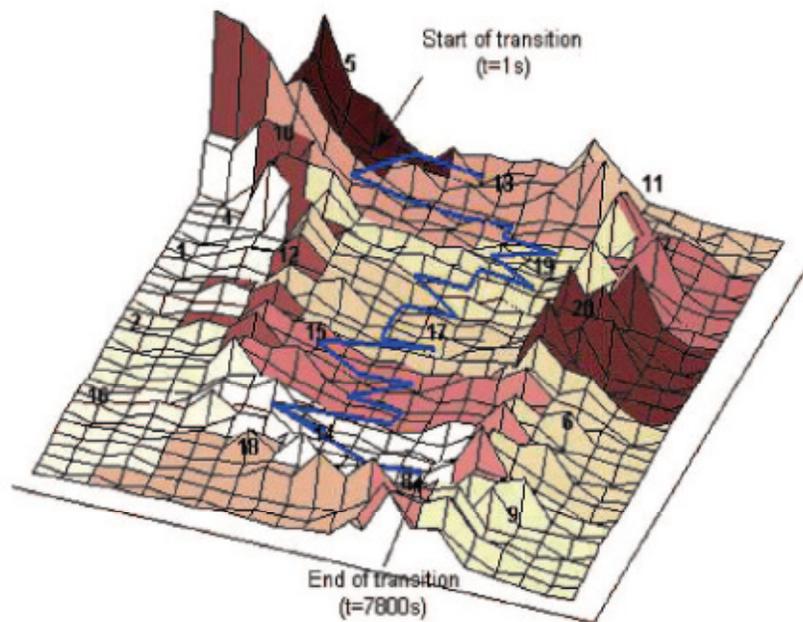


Figura 2.7: Exemplo de matriz-U combinada com BMU's. Fonte: Ng & Srinivasan (2008a).

A técnica de mapas auto-organizáveis vem sendo bastante estudada e utilizada, e o campo de aplicação para ela é muito amplo. Algumas aplicações são listadas a seguir:

- Em Ng & Srinivasan (2008a) foi sugerida uma nova implementação para melhor visualização de operações multiestado de uma unidade de craqueamento de petróleo. Essa implementação se baseia em excluir pontos do estado estacionário que sejam muito parecidos, podendo visualizar, assim, os estados de transição mais claramente. A técnica foi testada em uma destilação de etanol-água em escala de laboratório e também na operação de um reator de craqueamento de uma refinaria para detecção de falhas operacionais;
- Em Ng & Srinivasan (2008b) a mesma implementação de exclusão de pontos semelhantes utilizada em Ng & Srinivasan (2008a) foi aplicada para um acompanhamento online na detecção de falhas. A técnica foi aplicada na partida de uma planta de destilação em escala de laboratório e em um caso conhecido como "*Tennessee Eastman (TE) industrial challenge problem*";
- Em Ouzounoglou et al. (2010) foi implementada a técnica de self organizing maps para detectar a existência ou não de um correspondente a uma impressão digital fornecida no banco de dados;
- Em Aguado et al. (2008) a técnica de SOM e de PCA foram aplicadas para a análise de dados multidimensionais na remoção de fósforo biológico de águas residuais. As duas técnicas apresentaram resultados muito bons em relação à interpretação dos dados, porém foi apontado que o SOM teria vantagem sobre o PCA por ter uma estratégia apta a lidar com problemas não-lineares e também por lidar diretamente com as variáveis originais;
- Em Fuertes et al. (2010) foi utilizada a técnica de SOM para visualizar operações de multiestado. Propôs-se um método que identifica os clusters

presentes na operação e a probabilidade de transição entre eles usando a trajetória seguida pelos dados alimentados ao sistema no mapa 2D. Um novo método de detecção e identificação de falhas também foi proposto. Os métodos foram aplicados na supervisão de uma planta real com 26 variáveis;

- Em Domínguez et al. (2007) foi proposto um acompanhamento em tempo real de visualizações estacionárias como plano de componentes e mapas de distâncias entre neurônios e dinâmicas como a trajetória temporal no mapa dos dados de uma planta piloto. O objetivo do trabalho era detectar possíveis condições anormais de operação.

O mapa topológico do SOM para o caso bidimensional possui duas opções de inicialização: hexagonal ou retangular. Na Figura 2.8 estão representadas as duas alternativas e é fácil ver que no caso retangular cada neurônio tem quatro vizinhos diretos, enquanto que no caso hexagonal cada neurônio tem seis vizinhos diretos. A inicialização mais comum é a hexagonal.

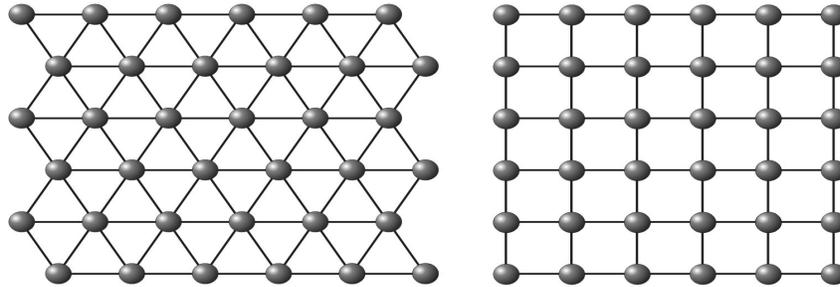


Figura 2.8: Mapas topológicos hexagonal e retangular.

O algoritmo de treinamento do mapa auto organizável começa pela inicialização dos pesos sinápticos dos neurônios, que pode ser estimada pela análise da representação de dados ou de forma aleatória, a mais utilizada. Segundo Haykin (1999) uma vez que a grade tenha sido apropriadamente inicializada, há três processos essenciais envolvidos na formação do mapa auto-organizável: competição, cooperação e adaptação sináptica. A seguir é resumido como cada processo é feito.

2.2.1 Inicialização dos pesos sinápticos

A inicialização mais utilizada, a aleatória, é também a mais simples. Nela nenhum estado de organização é considerado e os pesos sinápticos dos neurônios são gerados por um gerador de números aleatórios que satisfaça a condição

$$\sum_{k=1}^K w_{kj} = 1 \text{ para todo } j \quad (2.1)$$

Sendo j o sub-índice do neurônio na grade e k o sub-índice que o relaciona ao seu dado de entrada, lembrando que existe um peso sináptico para cada dado que é alimentado ao neurônio.

2.2.2 Processo de competição

Uma característica que diferencia o SOM de redes neurais artificiais comuns é a tática de competição entre os neurônios. Dado um conjunto de dados de entrada, o neurônio que tiver a maior ativação àquele conjunto é declarado o neurônio vencedor, e só ele sofre a adaptação sináptica por completa. Para declarar qual neurônio é o vencedor precisa-se de um algoritmo, que é detalhado a seguir.

Sendo m a dimensão dos dados de entrada, seleciona-se um conjunto de dados representado por x , sendo

$$X = [x_1, x_2, \dots, x_m]^T \quad (2.2)$$

A dimensão do vetor de pesos sinápticos do neurônio é a mesma dos dados de entrada e considerando que existam k neurônios na grade, o vetor peso sináptico do neurônio j é representado por

$$W = [w_{j1}, w_{j2}, \dots, w_{jm}]^T, \quad j = 1, 2, \dots, k \quad (2.3)$$

Para declarar-se então o neurônio vencedor pode-se tomar como parâmetro o produto interno $\langle [w_j]^T, X \rangle$ para todo j de 1 a k , sendo o neurônio com maior ativação a certo conjunto de dados de entrada o que tenha o maior produto interno. Porém, segundo Haykin (1999) o critério do melhor casamento, baseado na maximização do produto interno $\langle [w_j]^T, X \rangle$ é matematicamente equivalente a minimizar a distância euclidiana entre os vetores X e W . Sendo essa uma representação computacionalmente mais fácil é só determinar para todos os neurônios o neurônio $i(x)$ que satisfaça a condição representada por

$$i(x) = \operatorname{argmin} \|x_j - w_j\|, j = 1, 2, \dots, k \quad (2.4)$$

2.2.3 Processo de cooperação

O processo de cooperação é de fundamental importância na técnica de SOM e funciona para atualizar o peso sináptico não só do neurônio vencedor, e sim de uma vizinhança ao redor do mesmo. Quando o neurônio vencedor é determinado ele fica definido como o centro de uma vizinhança e vai ser atualizado por completo. À medida que se afasta dele o fator de atualização é decrescido até chegar a zero para neurônios muito distantes.

A função que descreve a vizinhança pode ser estabelecida conforme se quiser desde que satisfaçam as condições de decrescer conforme a distância até o neurônio vencedor aumente e seja simétrica em relação ao centro da vizinhança. Uma escolha típica para a função de vizinhança é a função gaussiana que é representada por

$$h_{j,i} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right) \quad (2.5)$$

Onde $d_{j,i}$ é a distância entre o neurônio j e o neurônio vencedor i e σ é a largura efetiva da vizinhança topológica como ilustrada na Figura 2.9.

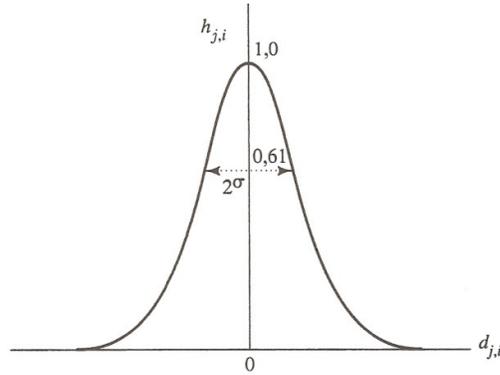


Figura 2.9: Função de vizinhança gaussiana.

Uma característica necessária para a função de vizinhança é que ela restrinja cada vez mais quantos neurônios serão atualizados conforme o número de iterações. Na prática, para a função gaussiana o valor de σ deve ser reduzido, ou seja, a função deve ficar mais estreita. Um método para se fazer isso é descrito por Ritter & Kohonen (1989), alterando o valor da largura efetiva da vizinhança topológica em função do tempo, definido como

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_1}\right) \quad n = 0, 1, 2, \dots, \quad (2.6)$$

Sendo σ_0 o valor de σ na inicialização do algoritmo, n o tempo discreto e τ_1 uma constante de tempo.

Desse jeito a função vizinhança que atende todos esses requisitos torna-se

$$h_{j,i}(n) = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2(n)}\right) \quad (2.7)$$

2.2.4 Processo adaptativo

No último processo de atualização dos pesos sinápticos dos neurônios tem-se a sua efetiva modificação. O valor w_j no tempo $(n+1)$ é dado, como Kohonen (1982)

$$w_j(n+1) = w_j(n) + \eta(n)h_{j,i}(n)[x - w_j(n)] \quad (2.8)$$

Onde $\eta(n)$ é chamado de taxa de aprendizagem e decresce exponencialmente com o tempo discreto n conforme

$$\eta(n) = \eta_0 \exp\left(-\frac{n}{\tau_2}\right), \quad n = 0, 1, 2, \dots, \quad (2.9)$$

Sendo η_0 o valor inicial para o parâmetro taxa de aprendizagem e τ_2 uma constante de tempo.

Pode-se dividir o processo adaptativo em duas distintas fases: fase de auto-organização ou de ordenação e fase de convergência. A fase de ordenação ocorre nas primeiras iterações e começa com os pesos sinápticos inicializados, então é necessário que eles sejam ordenados e se comece a identificação dos padrões dos dados de entrada.

Segundo Kohonen (1997a) essa fase pode exigir 1000 iterações ou mais e algumas escolhas para os parâmetros que se mostraram adequadas é fazer com que a taxa de aprendizagem decresça gradualmente com o tempo, mas permaneça em um valor acima de 0,01 e a função de vizinhança inclua inicialmente quase todos os neurônios da grade e diminua para no final incluir apenas alguns neurônios vizinhos em torno do neurônio vencedor. Esses objetivos são alcançados escolhendo-se $\eta_0=0,1$, $\tau_2=1000$ e $\tau_1 = \frac{1000}{\log \sigma_0}$.

Na fase de convergência ocorre a sintonia fina dos neurônios da grade e a representação efetiva dos dados de entrada. Ainda segundo Kohonen (1997a) o número de iterações deve ser no mínimo de 500 vezes o número de neurônios na grade. O parâmetro da taxa de aprendizagem deve se manter em um valor na ordem de 0,01 e não deve-se permitir que chegue a zero, pois o mapa pode alcançar um estado que não se atualize mais nem represente os dados de entrada. A função de vizinhança deve conter apenas os vizinhos mais próximos do neurônio vencedor, ou às vezes somente ele próprio.

O algoritmo do SOM se resume, então, de acordo com Haykin (1999), em 5 etapas:

1. Inicialização: escolha dos pesos sinápticos de todos os neurônios para o tempo discreto $n=0$. Usualmente a inicialização é feita de modo aleatório;
2. Amostragem: retirada de um conjunto de dados do espaço de entrada e alimentação do mesmo no processo de competição;
3. Determinação do neurônio vencedor: determinar qual neurônio tem a maior ativação para o conjunto de dados alimentados pela menor distância euclidiana, essa etapa corresponde ao processo de competição;
4. Atualização: os pesos sinápticos do neurônio vencedor e de seus vizinhos são atualizados conforme a equação 2.8. Corresponde ao processo adaptativo;
5. Continuação: volta ao passo da amostragem até que não sejam observadas mudanças significativas no mapa.

3 Materiais e Métodos

Como não se dispunha de dados operacionais contemplando uma ampla região de operação, optou-se pela construção de uma simulação estática utilizando o software Aspen Plus e obter os dados necessários para implementar a técnica de redes neurais e de self-organizing maps. O algoritmo referente aos mapas auto-organizáveis já está disponível em MATLAB, então foi escolhido este programa para processar as variáveis de entrada. A seguir será descrito como foi usado cada programa.

3.1 Aspen Plus

O software Aspen Plus oferece um amplo banco de dados no que se refere a componentes puros e equilíbrio de fases entre eles, além de poder incluir no fluxograma colunas de destilação, compressores e trocadores de calor. Sendo assim mostrou-se uma boa ferramenta para simular qualquer planta que fosse necessária. Adicionalmente, é um programa que gera os dados referentes à simulação facilmente, característica fundamental para uma boa utilização de mapas auto-organizáveis, já que os mesmos requerem um vasto conjunto de dados de entrada.

Para a realização da simulação é preciso escolher o pacote termodinâmico que o Aspen Plus irá utilizar para o cálculo das interações entre os diversos componentes de alimentação da planta. O pacote termodinâmico escolhido para a simulação foi o Peng-Robinson, que apresenta bons resultados quando os componentes presentes são de origem petroquímica. Esse pacote não é aconselhado quando se tem a presença de água, o que acontecia nos dados fornecidos de simulação. A fim de contornar esse problema, foi retirada a alimentação de água da primeira coluna de destilação, responsável por separar os compostos C2 dos C3, já que a totalidade da água é retirada nessa coluna e o objetivo de estudo do trabalho era a coluna de separação de propeno e propano, sendo assim, a água não atrapalhou os resultados da simulação.

Para a simulação das colunas de destilação foi utilizado o bloco “RadFrac”, que pode simular todo tipo de fracionamento líquido-vapor em múltiplos estágios. Para o refeedor da coluna de separação propano e propeno utilizou-se o bloco “HeatX”, que é usado para a troca de calor entre duas correntes de alimentação. Nesse bloco é possível especificar somente a área de troca térmica e as temperaturas de saída são calculadas tomando valores para o coeficiente de troca térmica global. Para o compressor utilizou-se o bloco “Compr”, escolhendo um modelo isentrópico e especificando a pressão de descarga. O vaso de topo foi simulado utilizando o bloco de nome “Flash2”, usado quando se tem duas correntes de saída, para o mesmo foram especificadas a pressão e a temperatura.

Após concluir-se a simulação foram feitas diversas análises de sensibilidade da planta variando tanto características dentro do processo, como a quantidade de corrente desviada para o *hot bypass*, pressão de saída do compressor assim como a carga de alimentação. A variação dessas características foi analisada coletando informações importantes e assim gerando um banco de dados de operação da planta importantíssimo para o desenvolvimento do trabalho.

De posse desses dados foi possível seguir para a próxima etapa, realizada em MATLAB.

3.2 MATLAB

Para a aplicação da técnica das redes neurais e de mapas auto-organizáveis escolheu-se trabalhar com o programa MATLAB, que se mostra uma ferramenta muito útil quando quer se trabalhar com matrizes, que é o caso em questão. No MATLAB foram feitos dois processos distintos: pré-processamento dos dados por eliminar amostras que seriam muito semelhantes às outras e a aplicação da técnica de self-organizing maps em si, para posterior escolha do que mais se adequava ao objetivo proposto. Esta seção está dividida nas duas partes já mencionadas.

3.2.1 Algoritmo para pré-processamento de dados

O treinamento de um mapa por self-organizing maps é influenciado somente pelos dados que são alimentados ao mesmo, assim sendo, a escolha dos dados a serem alimentados é de fundamental importância. A remoção de dados muito semelhantes torna-se importante quando o banco de dados fornecido para o treinamento possui muitas amostras da mesma região. Desse jeito a região que possui mais dados acaba sendo muito representada e a existência de outra região com menos dados, porém não necessariamente menos importante, é mascarada.

Em MATLAB foi criado um algoritmo que conseguisse filtrar os dados de entrada em dados que fossem realmente diferentes entre si, a fim de acabar com o problema de um exagero de dados da mesma região. Para tanto, cada conjunto de dados era comparado com os dados da nova matriz filtrada através da equação a seguir

$$|x_{i,j} - y_{k,j}| > T \text{ para } i = 1, 2, \dots, n, k = 1, 2, \dots, m \text{ e } j = 1, 2, \dots, l \quad (3.1)$$

Onde T é um valor arbitrário, $x_{i,j}$ é o dado alimentado, $y_{k,j}$ é o dado da nova matriz já filtrada, l é o número de variáveis e n e m são o número de amostras do conjunto de dados alimentados e da matriz filtrada, respectivamente.

No algoritmo criado o primeiro dado alimentado era automaticamente transferido para a matriz filtrada e então começava a rotina de analisar se a próxima amostra satisfazia a equação de número 3.1. Se a equação 3.1 fosse satisfeita para pelo menos um valor de j em todas as m amostras da nova matriz a amostra de dados alimentada era considerada apta a fazer parte da matriz filtrada. Desse jeito, foi possível pré-processar os dados de forma eficiente e avaliar qual valor de T produzia um mapa de melhor qualidade facilmente.

É importante frisar que esse algoritmo só tem finalidade se for utilizado em dados normalizados, já que o valor de T não muda para as variáveis. Logo, se fosse aplicado em dados não normalizados, haveria distorções devido às diferentes ordens de grandeza das variáveis. A normalização, que no caso foi pela variância, foi feita dentro do mesmo toolbox em que se aplicou a técnica de self-organizing maps.

3.2.2 SOM Toolbox

O toolbox utilizado, chamado de “SOM Toolbox”, foi desenvolvido pela mesma equipe de pesquisadores que trabalha com quem formulou toda a base matemática para a evolução da técnica de mapas auto-organizáveis. O “SOM Toolbox” é uma ferramenta muito versátil e de fácil utilização, sendo possível tanto pré-processar como processar os dados e visualizar os resultados de forma muito rápida.

A biblioteca implementada pelo toolbox tem funções para normalizar os dados de entrada, o que é muito importante, já que as grandezas entre as variáveis podem diferir consideravelmente. Outras características podem ser alteradas, como o número de neurônios treinados pelo mapa, treinamento sequencial ou em batelada, inicialização do mapa topológico hexagonal ou retangular, função da vizinhança utilizada, inicialização dos pesos sinápticos randômica ou linear, entre outras. Também é possível treinar o mapa em duas etapas: primeiro a etapa de convergência com todos os dados e depois escolher um conjunto para a etapa de sintonia fina, assim representando-o de forma mais precisa.

Após o treinamento do mapa de neurônios por self-organizing maps usou-se uma etapa supervisionada de treinamento, ou seja, com um professor. Para cada conjunto de dados foi fornecida uma resposta desejada do sistema, ou seja, cada conjunto de variáveis tinha um “rótulo” referente a uma variável de saída. No caso, foi estabelecido como variável de saída a concentração molar de propeno no produto de topo da segunda coluna de destilação.

4 Estudo de Caso

A metodologia estabelecida no Capítulo 3 será aplicada a uma unidade de processamento de gás liquefeito de petróleo em funcionamento pertencente à PETROBRAS. O presente capítulo está dividido em duas partes: a primeira parte descrevendo a unidade que foi estudada e as suas características e a segunda parte sobre a construção da simulação da unidade em questão no software Aspen Plus. Parte do escopo do trabalho será a construção da simulação estática da coluna desetanizadora e da coluna de separação de propano e propeno e a predição de variáveis de saída da última coluna, que tem como objetivo separar uma alimentação composta de propano e propeno e outros contaminantes. Os dados, tanto de características do sistema como composições das correntes, foram fornecidos e utilizados para a construção da simulação em Aspen Plus.

4.1 Unidade de Processamento de Gás Liquefeito de Petróleo

A alimentação da unidade consiste de hidrocarbonetos de dois a seis carbonos e basicamente de três colunas de destilação que serão melhores descritas a seguir. O fluxograma geral da unidade está representado na Figura 4.1.

4.1.1 Coluna despropanizadora

A primeira coluna da unidade recebe a alimentação com hidrocarbonetos de C2 a C6 e água e tem a finalidade de separar os compostos com mais de quatro carbonos do restante. Esta coluna consegue atingir uma remoção de praticamente 99,9% de indesejados, deixando passar apenas resquícios de isobutano, isobuteno e 1-buteno junto com o etano, o propano, o propeno e água. Essa coluna não foi simulada em Aspen, somente foram utilizadas as propriedades e a composição do produto de topo para a segunda coluna. A coluna despropanizadora está representada por D-1 na Figura 4.1 e a alimentação, refluxo, produto de topo e produto de fundo são as correntes 1, 3, 5 e 6 respectivamente. A corrente 5 é a alimentação da segunda coluna, a desetanizadora, a qual será objeto de estudo desse trabalho, juntamente com a coluna de separação propano e propeno.

A corrente de topo ainda é passada por um tratamento de lavagem cáustica, um vaso de lavagem com água e um vaso de decantação para remoção de impurezas presentes como sulfeto de hidrogênio, representados por V-1, V-2 e V-3.

4.1.2 Coluna desetanizadora

A segunda coluna de destilação da unidade recebe o produto de topo da coluna despropanizadora e tem a finalidade de remover o etano e a água para a próxima etapa. A remoção de etano é de cerca de 99,8% da alimentação e a água é totalmente removida. O produto de topo da coluna é mandado para a unidade de FCC (Fluid Catalytic Cracking) contendo etano e também propano e propeno em baixas quantidades. O produto de fundo, contendo principalmente propano e propeno é enviado para a coluna de separação dos dois. A coluna deetanizadora foi simulada em Aspen Plus com os dados de simulação fornecidos pelo respectivo projeto.

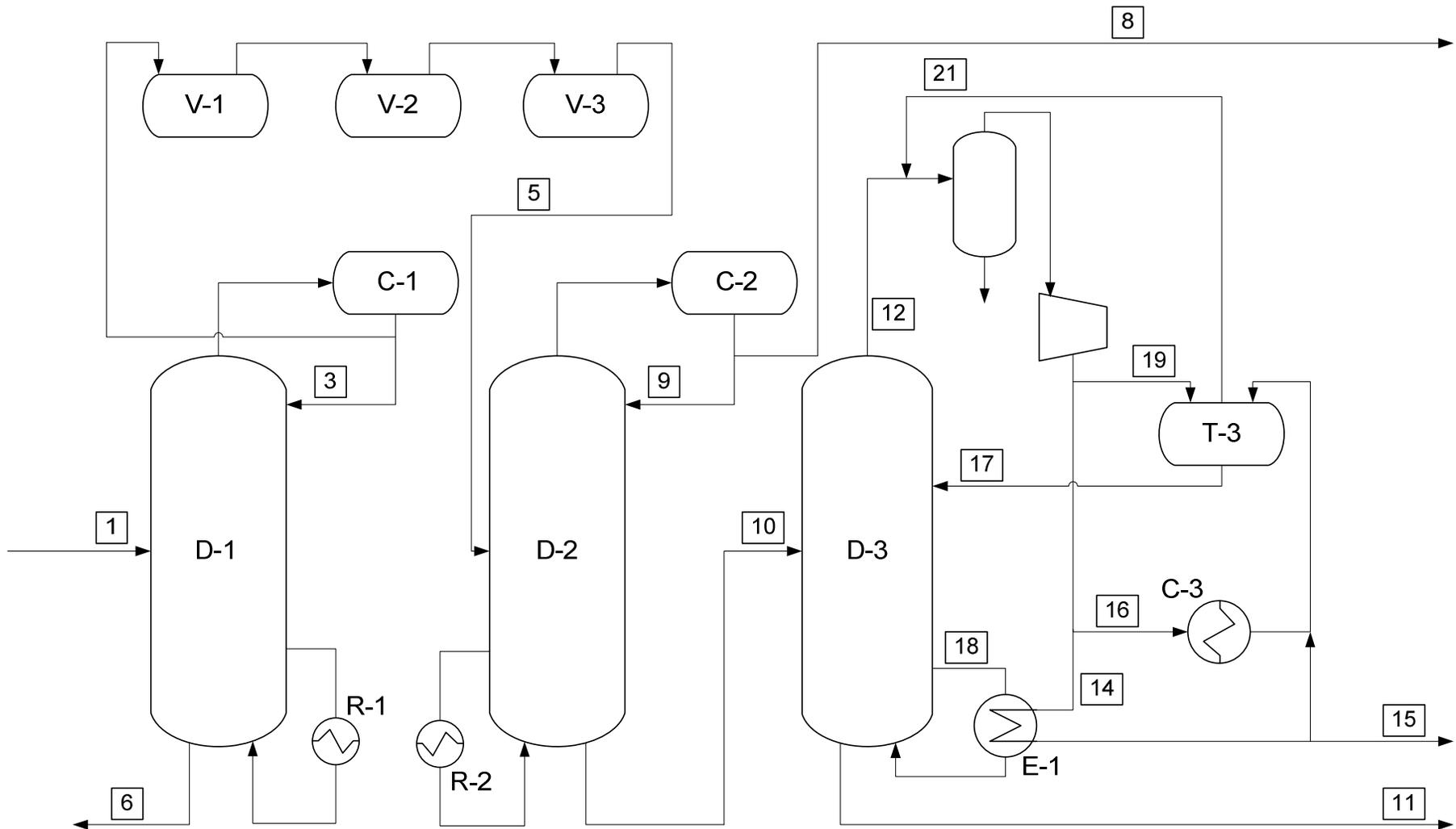


Figura 4.1: Fluxograma da Unidade de Processamento de Gás Liquefeito de Petróleo.

Na Figura 4.1 a coluna está representada pelo nome de D-2 e as correntes de alimentação, refluxo, produto de topo e produto de fundo representadas pelas correntes de número 5, 9, 8 e 10 respectivamente. A corrente 10 é a alimentação da terceira coluna.

4.1.3 Coluna de separação propano e propeno

A terceira coluna da unidade é a mais complexa de todas, com mais de 150 pratos de destilação e peculiaridades operacionais que a distinguem das outras colunas pertencentes à unidade. A alimentação é proveniente do produto de fundo da coluna desetanizadora e é composta de propano e propeno em grandes quantidades, totalizando quase 99,9% do total, e etano, isobutano, isobuteno e 1-buteno em baixíssimas quantidades. A recuperação de propeno no produto de topo é de 98,8% enquanto que a de propano no produto de fundo é de 98,4%.

O tamanho da coluna a princípio parece exagerado, porém quando comparam-se as propriedades físicas e químicas de propano e propeno se torna evidente que é necessário uma coluna de tamanha proporção para que se tenha uma separação efetiva, já que a destilação simples utilizada baseia-se na diferença entre as temperaturas de ebulição dos componentes. A viabilidade econômica da coluna também pode ser questionada, porém a diferença considerável de preço entre o propano, que retorna a GLP, e o propeno, que é matéria-prima fundamental da indústria de polímeros, torna esse balanço favorável à separação entre os dois, mesmo que para isso seja necessária uma coluna de proporções tão grandes.

A coluna está representada na Figura 4.1 pela legenda D-3, e sua alimentação é a corrente 10. A saída de topo, representada pela corrente 12, se junta com a corrente de vapor da saída do vaso de topo, representada pela corrente 21, e entra em um vaso para retirada de líquido. A saída de vapor desse vaso segue para um compressor, que irá aumentar tanto a pressão como a temperatura dessa corrente. O uso de um vaso para a retirada de líquido tem importância visto que a alimentação do compressor tem que ser composta somente por vapor.

O vapor comprimido da saída do compressor é dividido em duas correntes: a corrente 16, que vai para o condensador refrigerado com água, e a corrente 14, que segue para o refeedor da coluna. No refeedor da coluna há um trocador de calor, representado pelo nome de E-1, entre a corrente 14 e a corrente 18, que é recirculada para dentro da coluna após ser evaporada. Essa estratégia de operação, chamada de bomba de calor, dispensa o uso de utilidades no refeedor da coluna, diminuindo os custos de operação.

A corrente 14, após a passagem pelo trocador de calor, é dividida mais uma vez em duas correntes: a corrente 15 é resfriada e retirada como produto de topo rico em propeno e o resto junta-se com a corrente 16 que vem do condensador e vai para o vaso de topo. Do vaso de topo sai a corrente já mencionada 21 que se junta com a saída de topo da coluna e também a corrente 17 que retorna à coluna como refluxo.

Outra característica dessa coluna de destilação é a possibilidade de uma divisão da corrente de saída do compressor direto para o vaso de topo, numerada pela corrente 19. Essa estratégia, chamada de *hot bypass*, é uma forma de controlar a pressão dentro do vaso de topo de forma mais eficiente.

Do fundo da coluna saem duas correntes: a 18 que volta como recirculação proveniente do refulvedor e a 11, que é o produto de fundo rico em propano, é bombeada como GLP.

As variáveis controladas da coluna são a pressão de topo e os níveis da coluna, do condensador e do refulvedor. As principais variáveis manipuladas são as vazões de refluxo, de produto de topo e fundo.

4.2 Construção da simulação em Aspen Plus

Para a aplicação da técnica de self-organizing maps é necessário que se tenha um amplo banco de dados contendo ampla região operacional da unidade estudada. Para isso, foi construída a simulação da planta em Aspen plus. Desse jeito, seria fácil manipular variáveis da planta e gerar um conjunto de amostras significativo. Na Figura 4.2 se vê o fluxograma da simulação realizada. O pacote termodinâmico escolhido, como já foi mencionado, é o Peng-Robinson (PR) visto que o simulador possui os parâmetros necessários para a mistura estudada, rica em hidrocarbonetos.

Na simulação construíram-se apenas as colunas deetanizadora e de separação propano e propeno, já que o foco do trabalho é estudar a segunda. A coluna deetanizadora foi a de mais fácil implementação, já que seus refluxos foram feitos internamente no seu bloco. A coluna de separação propano e propeno apresentou maiores dificuldades na construção e na convergência pela necessidade de estabelecer os reciclos externamente ao bloco já implementado em Aspen Plus.

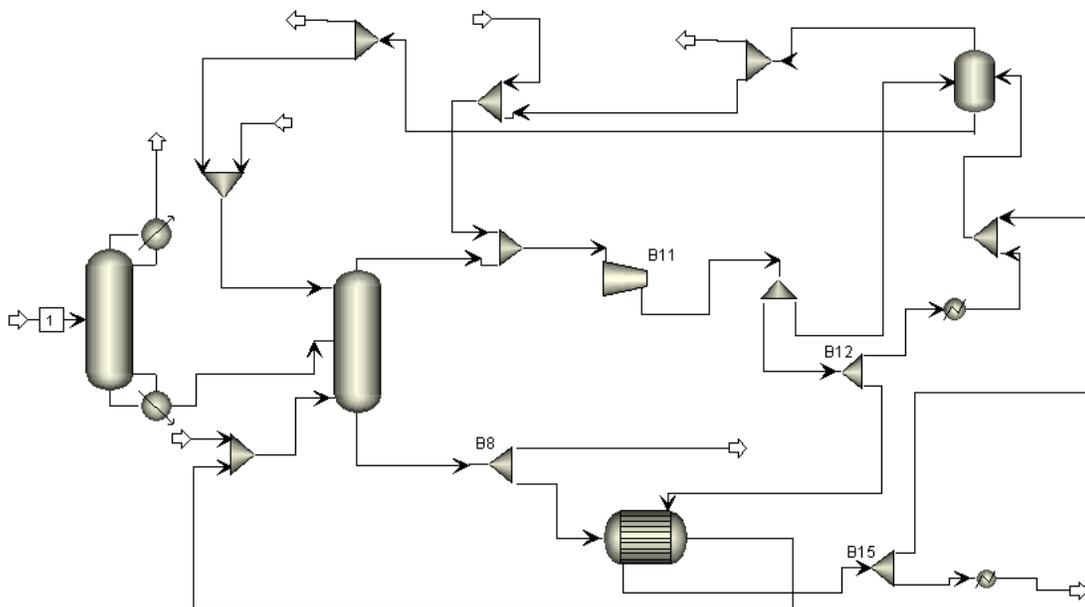


Figura 4.2: Fluxograma da simulação realizada em Aspen Plus.

De posse da simulação foi possível, através da ferramenta de análise de sensibilidade, variar características de simulação da planta e obter os dados necessários para aplicação da técnica. Para as variáveis manipuladas variou-se a pressão de saída do compressor, representado na Figura 4.2 pelo bloco B11, a saída de produto de topo variando o valor do bloco B15, a vazão de *hot bypass*, alterando o valor de split do bloco que recebe a

descarga do compressor e também a carga de alimentação da coluna deetanizadora representada pela corrente 1.

Para as variáveis de saída utilizou-se a carga total da planta, o split do bloco B15, as composições de propano e propeno no produto de topo e de fundo, as vazões molares de produto de topo e de fundo, a vazão da corrente que vai para o condensador, a vazão de *hot bypass* e de refluxo, a pressão de saída do compressor, o calor trocado no refeedor da coluna e as temperaturas dos pratos 20, 40, 80, 100, 120 e 140 da coluna de separação propano e propeno, resultando em 19 variáveis.

Com todas essas variações criou-se um banco de dados com 1243 amostras das 19 variáveis, alterando as variáveis manipuladas até os seus limites de convergência na simulação. A pressão de saída do compressor, representado pelo bloco B11, foi variada de 33% a menos até 20% a mais que a folha de dados de projeto. No bloco B15, em que foi variada a fração de corrente direcionada a produto de topo, foi alterado o valor de 0,05 até 0,15, sendo que o dado de projeto é de 0,08. A vazão de *hot bypass*, igualada a zero no projeto, foi variada de 0 a 5% de corrente direcionada do bloco que recebe a descarga do compressor e a carga de alimentação foi variada, em relação aos dados de projeto, de 40% menos até 30% mais no que se refere à vazão total.

5 Resultados

Os resultados deste trabalho serão divididos em três partes. Na primeira parte serão apresentados os resultados da simulação feita em Aspen Plus e os erros quando comparados os dados de simulação aos dados de projeto. Na segunda serão mostradas as matrizes-U para os mapas treinados pelos dados de simulação tanto com como sem a aplicação do algoritmo de remoção de dados semelhantes assim como as trajetórias crescentes de composição e de variáveis de operação da planta. Na terceira parte serão mostrados os resultados da etapa supervisionada para a predição de composição de propeno no produto de topo.

5.1 Simulação em Aspen Plus

A simulação em Aspen Plus foi realizada com os dados de projeto fornecidos para as colunas desetanizadora e de separação propano e propeno. Na Tabela 5.1 encontram-se os erros relativos entre a simulação realizada e os dados de projeto para a carga total e concentração molar de propeno nas principais correntes fornecidos pela refinaria.

Tabela 5.1: Comparação dos dados de carga total, concentração molar de propeno e temperatura da simulação realizada em Aspen com os dados de projeto.

Correntes	Erro relativo (%) entre a carga total de projeto e simulação	Erro relativo (%) entre a concentração molar de propeno de projeto e simulação	Erro relativo (%) entre a temperatura de projeto e de simulação
8	0,211	0,013	0,041
10	0,024	0,001	0,187
11	0,506	6,945	0,289
12	1,474	1,381	0,303
14	1,693	1,600	0,672
15	0,007	0,098	0,639
16	0,000	0,091	0,076
17	1,574	1,481	0,311
21	1,574	1,507	3,337

Como visto na Tabela 5.1 a simulação gerou dados com erros muito baixos em relação aos dados de projeto. O maior erro, na corrente de número 11 da concentração de propeno, se explica porque essa corrente é a de propano concentrado, portanto, a concentração de propeno é muito baixa.

5.2 Matrizes-U

Para visualização do mapa treinado por SOM o mais comum é utilizar-se a matriz-U. A matriz-U é calculada pela equação 5.1 e é preenchida com uma escala de cor para as distancias calculadas. O centro de cada hexágono é um neurônio e regiões de cor referentes a baixas distâncias podem ser interpretadas como clusters.

$$D = \sum_{m \in NN(n)} d(w(n) - w(m)) \quad (5.1)$$

Onde D é o valor mostrado na matriz-U, $d(w(n) - w(m))$ é a distância entre o vetor peso sináptico do neurônio n e do neurônio m e $NN(n)$ o conjunto de neurônios vizinhos a n .

A seguir serão mostradas diversas matrizes para diferentes tamanhos de mapas e também para diferentes valores para o algoritmo de pré-processamento de dados e também trajetórias crescentes de composição para o mapa escolhido.

5.2.1 Matrizes-U para todos os dados

Com os dados obtidos a partir da simulação, foi treinado o mapa com o SOM Toolbox para diferentes tamanhos. O tamanho do mapa é automaticamente calculado pelo *toolbox* pelos auto valores associados à matriz do conjunto de dados de entrada, e então esse valor foi alterado para ver o que produzia a melhor matriz-U. Na Figura 5.1 são observadas as matrizes de tamanho pequeno, normal e grande do mapa. O mapa de tamanho pequeno tem aproximadamente um quarto do número de neurônios do mapa de tamanho normal, enquanto que o mapa de tamanho grande tem quatro vezes o número de neurônios.

O número de neurônios calculados para o mapa de tamanho normal pelos auto valores associados à matriz do conjunto de dados de entrada foi de 20x9. Sendo assim, os mapas de tamanho pequeno e grande têm tamanhos de 9x5 e 40x18.

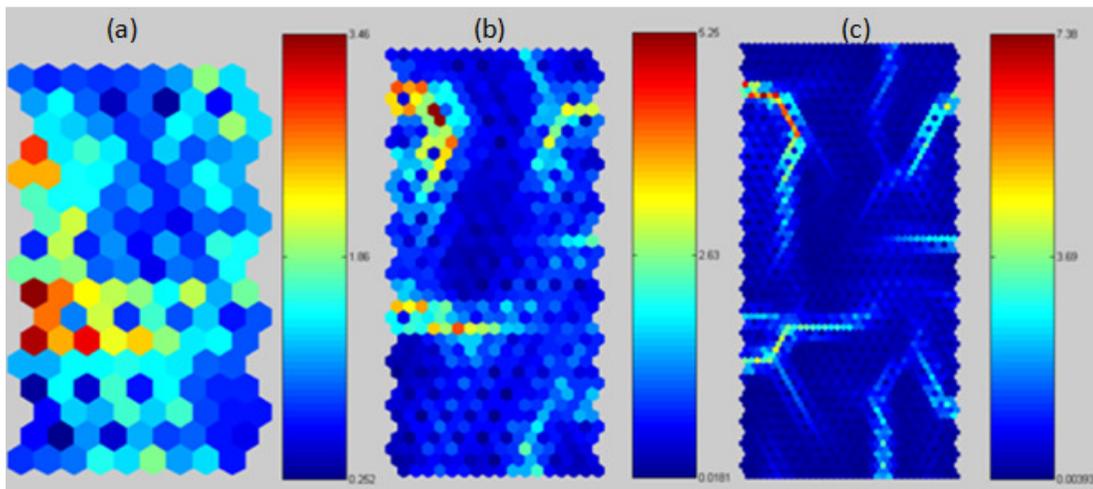


Figura 5.1: Matrizes-U para diferentes tamanhos de mapa. (a) pequeno, (b) normal e (c) grande.

Pela Figura 5.1 é visto claramente como o tamanho do mapa tem importância fundamental para a posterior caracterização do mesmo. O mapa pequeno não consegue separar adequadamente as regiões e acaba juntando diferentes regiões de operação no mesmo lugar, impossibilitando a análise. O mapa normal já começa a delimitar melhor as regiões de operação e o mapa grande acaba tendo uma separação exagerada das regiões, criando clusters muito grandes, representados pelas regiões azuis. A seguir foi feito um mapa com tamanho 20x5, representado na Figura 5.2, para melhorar a resolução do mapa de tamanho normal.

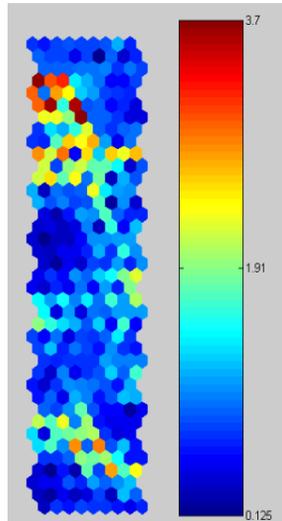


Figura 5.2: Matriz-U do mapa de tamanho 20x5.

A seguir, na Figura 5.3 pode ser vista a visualização em 3D para a matriz-U, dessa vez com uma componente vertical representando as distâncias entre o neurônio e os seus neurônios vizinhos.

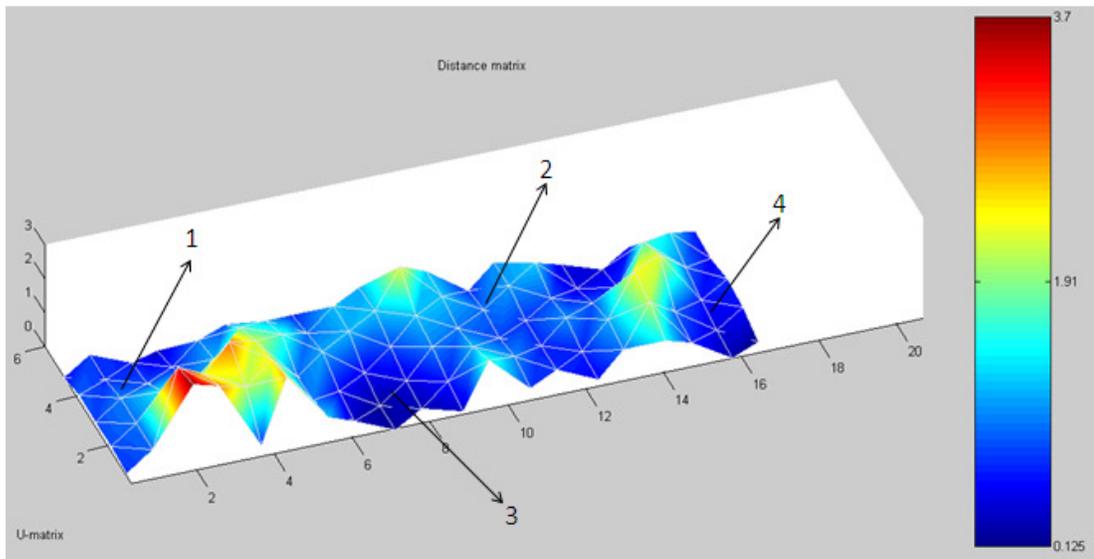


Figura 5.3: Visualização da Matriz-U em 3 dimensões para o mapa de tamanho 20x5.

É possível ver quatro regiões bem distintas na matriz em 3D representada na Figura 5.3. As regiões 1 e 4 estão bem separadas das restantes, enquanto que as regiões de 2 e 3

têm uma separação mais suave entre elas. Para a posterior classificação de cada uma das regiões é necessário traçar as trajetórias dos neurônios vencedores em ordem crescente ou decrescente de uma das variáveis. Dentre todos os mapas, o de tamanho 20x5 foi o que conseguiu separar melhor as regiões e foi escolhido para a visualização das trajetórias.

5.2.2 Matrizes-U após pré-processamento dos dados

Na criação do algoritmo para remoção de dados similares deixou-se o valor arbitrário podendo ser escolhido pelo usuário, e claramente a escolha dele tem implicações diretas na qualidade do mapa treinado. A seguir são mostradas diversas matrizes-U resultantes dos mapas treinados utilizando diferentes valores de T . As matrizes são mostradas nos tamanhos normal e grande.

Quanto maior o valor de T mais restrita é a nova matriz de dados filtrada. De um banco de dados inicial de 1243 amostras, a matriz filtrada para os valores de T 0,02, 0,04 e 0,1 tem, respectivamente, em torno de 740, 540 e 260 amostras. As Figuras 5.4, 5.5 e 5.6 mostram as matrizes-U de tamanho normal e grande para valores de T iguais a 0,02, 0,04 e 0,1, respectivamente.

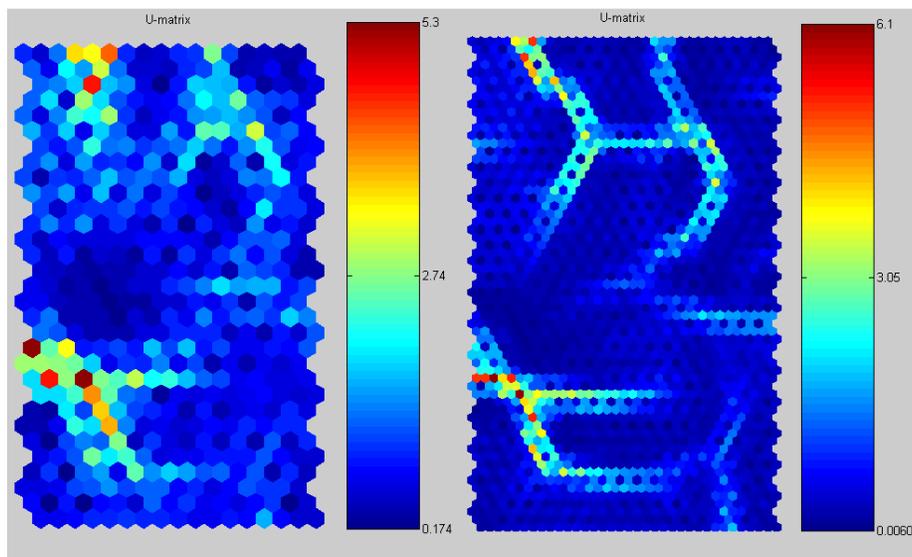


Figura 5.4: Matrizes-U para $T = 0,02$.

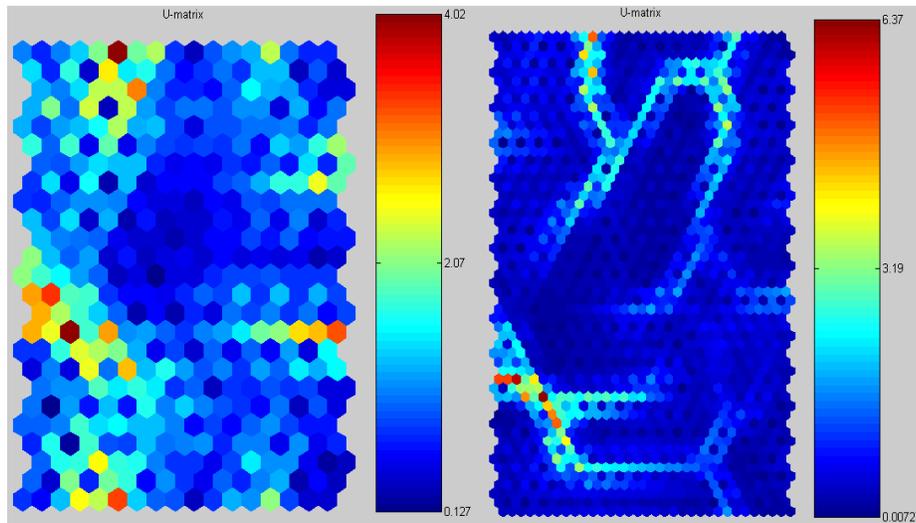


Figura 5.5: Matrizes-U para $T = 0,04$.

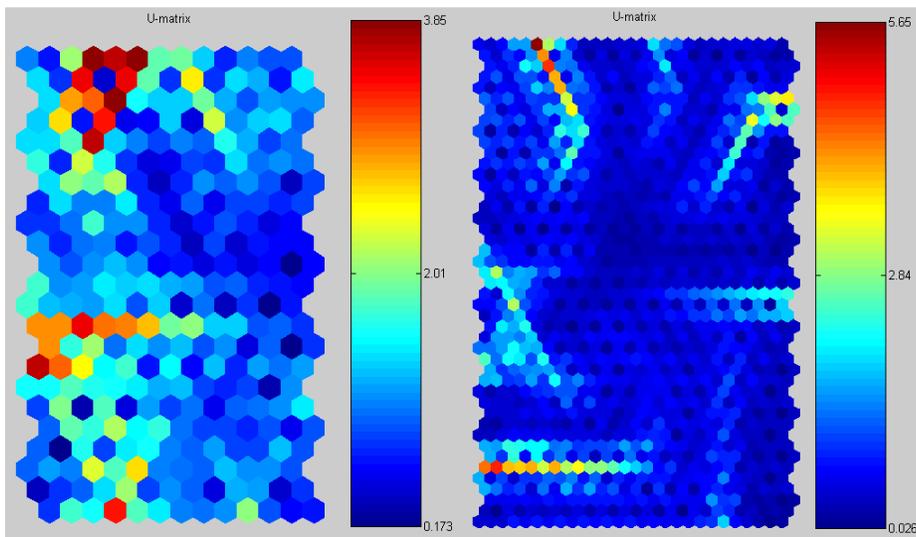


Figura 5.6: Matrizes-U para $T = 0,1$.

É visto, claramente, que quanto maior o valor de T cada vez menos regiões operacionais são representadas no mapa devido ao cada vez menor número de dados alimentados para o treinamento. O mapa treinado com o valor de T igual a 0,02 foi o que apresentou a melhor separação entre as regiões e a visualização da matriz-U em 3D é representada na Figura 5.7. As seis principais regiões operacionais que o mapa apresentou estão numeradas no gráfico.

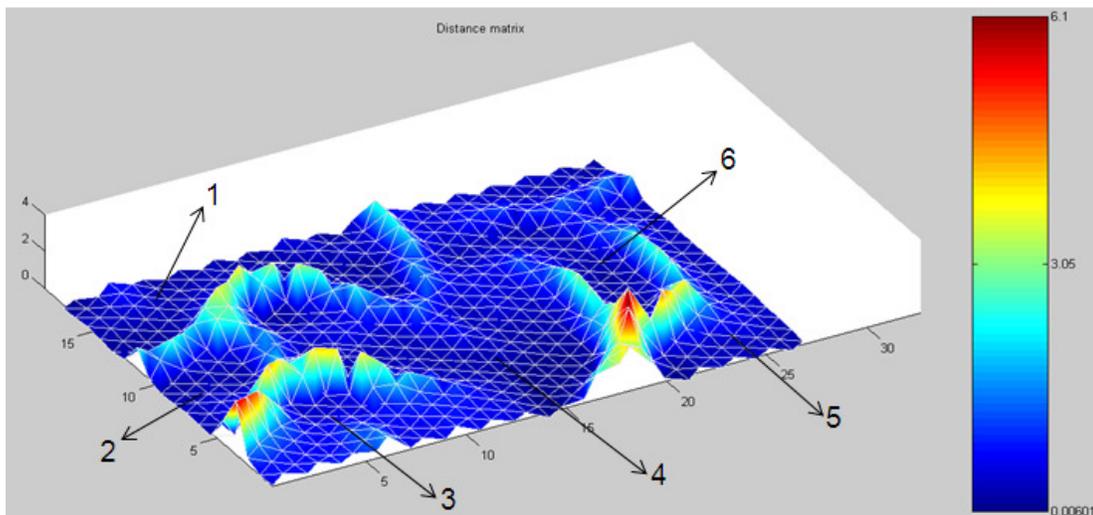


Figura 5.7: Visualização da Matriz-U em 3 dimensões com aplicação do algoritmo, T igual a 0,02 e mapa de tamanho grande.

5.3 Trajetórias aplicadas ao mapa escolhido

Como os mapas auto-organizáveis não permitem, a princípio, visualizar a evolução de alguma variável no mapa, é preciso ordenar a variável que se quer estudar, de forma crescente ou decrescente e, para cada amostra calcular o neurônio vencedor associado. Posteriormente, marcam-se na matriz-U os neurônios vencedores e traça-se a trajetória entre os pontos. Na simulação, para a obtenção dos dados, variaram-se 4 características da planta: vazão do produto de topo, pressão de descarga do compressor, vazão de *hot bypass* e carga total de alimentação. Obtiveram-se 499 dados para a variação da fração da corrente direcionada para o produto, 99 para a vazão de *hot bypass*, 145 para a pressão de descarga do compressor e 500 para a carga de alimentação. Para cada um desses bancos de dados foram feitas as trajetórias crescentes da variação respectiva e da concentração de propeno no produto de topo, portanto, cada par de matriz-U apresentada no decorrer do trabalho foi alimentada com dados diferentes.

Na Figura 5.8 tem-se a matriz-U com a trajetória crescente de composição de propeno no topo e também a trajetória crescente da fração direcionada ao produto de topo no bloco B15 da Figura 4.2. O menor valor está marcado com a cor azul, assim vê-se claramente que a composição e o valor da fração seguem caminhos com sentidos opostos, porém com a mesma tendência.

Desse jeito conclui-se que a fração da corrente de alimentação do bloco que segue como produto tem relação inversamente proporcional com a composição de propeno, ou seja, quanto menor a vazão redirecionada para o produto de topo maior a concentração de propeno dessa corrente. Também conclui-se que na matriz-U os valores de composição baixa estão em uma região do mapa, enquanto que os de composição alta estão em outra região oposta. Como os mapas são alimentados pelos mesmos dados, a escala de distância é a mesma para os dois. A coloração do mapa na escala de cinza foi escolhida para melhor visualização da trajetória.

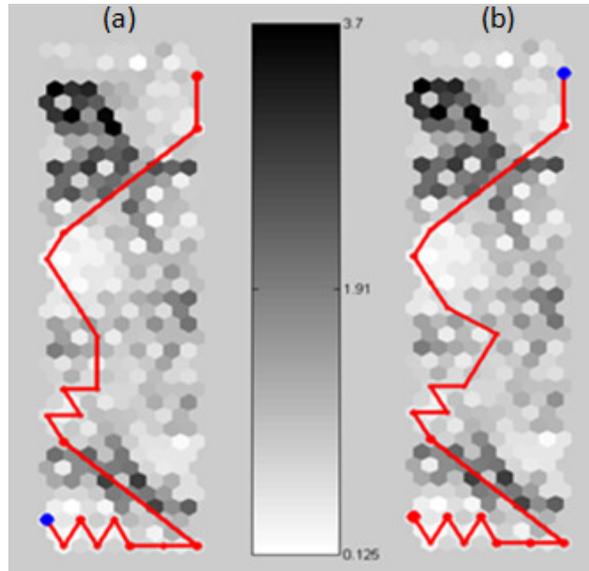


Figura 5.8: (a) trajetória crescente de composição de propeno no topo e (b) trajetória crescente da fração da corrente direcionada para produto de topo.

Um comportamento semelhante ao apresentado na Figura 5.8 é mostrado na Figura 5.9, que tem as trajetórias crescentes da vazão de *hot bypass* e da concentração de propeno no topo. Vê-se claramente que ambas as trajetórias possuem mesma tendência, porém agora as duas são diretamente proporcionais, o que pode ser visto pelos dois neurônios marcados de azul referentes aos menores valores de vazão de *hot bypass* ou concentração. Assim conclui-se que, para os dados analisados, quanto maior o *hot bypass*, maior a concentração de propeno no topo.

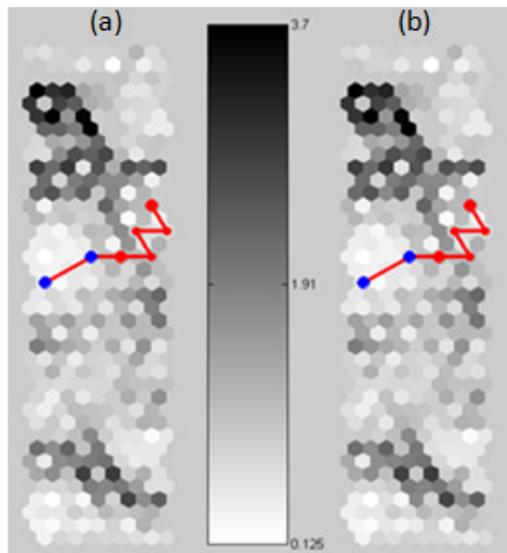


Figura 5.9: (a) trajetória crescente da composição de propeno no topo e (b) trajetória crescente da vazão de *hot bypass*.

Nas Figuras 5.10 são apresentadas as trajetórias crescentes de composição de propeno no topo frente à pressão de descarga do compressor e carga de alimentação. Vê-se que para os dados analisados a pressão de descarga do compressor segue a mesma tendência até certo ponto e é inversamente proporcional à composição de propeno no

topo, ou seja, quanto menor a pressão de descarga do compressor, maiores as concentrações de propeno no topo para as amostras analisadas.

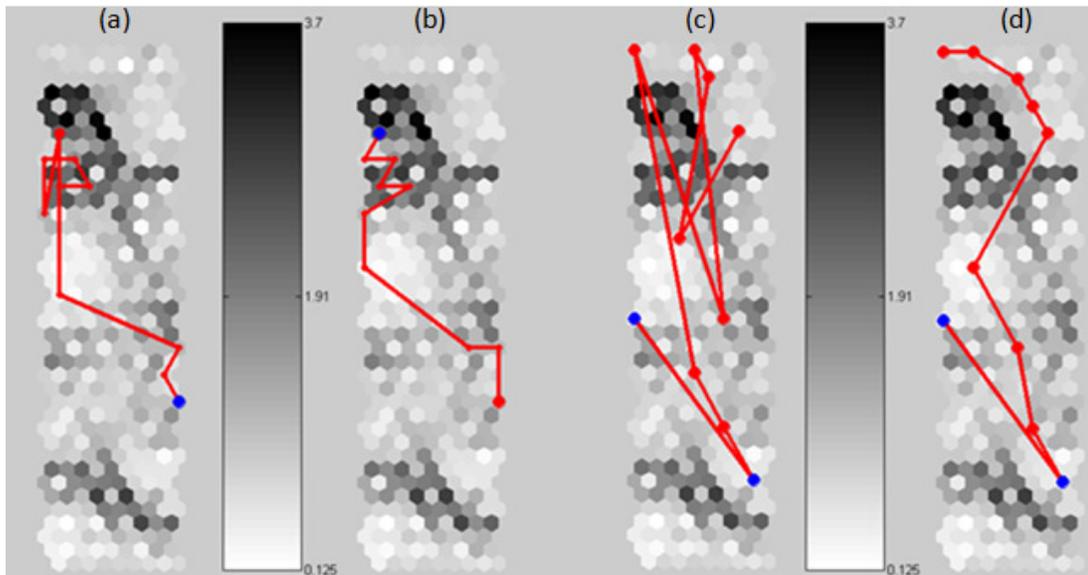


Figura 5.10: trajetória crescentes (a) e (c) da composição de propeno no topo, (b) da pressão de descarga do compressor e (d) da carga de alimentação.

A carga de alimentação segue uma trajetória crescente completamente diferente da trajetória crescente de composição de propeno no topo para os mesmos dados analisados, como pode ser visto na Figuras 5.10(c) e 5.10(d). Esse comportamento leva a crer que a variação da carga total de alimentação da coluna não tem influência direta com a composição de propeno no produto.

5.4 Predição da concentração de propeno no topo

Para a predição da concentração de propeno no topo é preciso fazer uma etapa supervisionada no mapa auto organizável, que é não supervisionado. A técnica do supervisionado consiste em fornecer na etapa de treinamento um valor desejado para cada amostra. Posteriormente ao treinamento, é alimentado um novo valor de amostras sem o valor que quer se prever e calculado o BMU (Best Matching Unit). Cada neurônio tem um valor associado a ele, e quando calcula-se o BMU de uma amostra tem-se o neurônio que mais se assemelha à amostra. Assim, o valor associado ao neurônio vencedor é tido como a resposta.

A etapa de treinamento dos mapas foi realizada com os mesmos dados de simulação estática para os três diferentes tamanhos de mapa. Para a escolha dos dados que seriam usados somente para treinamento ordenou-se a matriz do banco de dados aleatoriamente e tomaram-se as 1000 primeiras amostras. As 243 amostras restantes foram alimentadas na etapa supervisionada sem a concentração de propeno no topo e realizou-se a predição da variável. Com o valor predito pela técnica, o mesmo foi comparado com o dado obtido na simulação estática da planta.

A seguir, serão mostrados os resultados quando se prediz as composições de propeno na corrente de topo utilizando mapas de tamanho pequeno, normal e grande. Por ter mais neurônios, o mapa de tamanho grande é previsto para ter os melhores resultados. Como a obtenção do valor predito é posterior à determinação do neurônio vencedor, se

houver mais neurônios no mapa, mais valores de resposta podem ser obtidos. O número de neurônios treinados pelos mapas de tamanho pequeno, normal e grande são 45, 180 e 720, respectivamente.

Nas Figuras 5.11, 5.12 e 5.13 estão apresentadas no gráfico as concentrações previstas pelo tamanho de mapa respectivo em um eixo e as concentrações da mesma amostra tirada da simulação no outro eixo. Foi feita a predição das mesmas 243 amostras para os três diferentes tamanhos de mapas e a reta de equação $y = x$, que representa onde o erro é nulo, também está representada no gráfico.

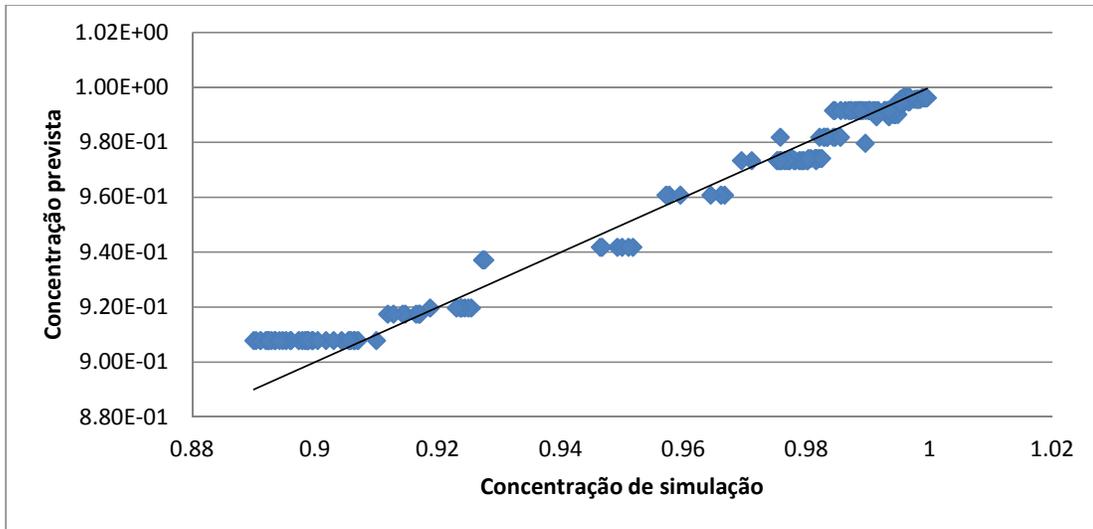


Figura 5.11: Resultado da predição para o mapa treinado com tamanho pequeno.

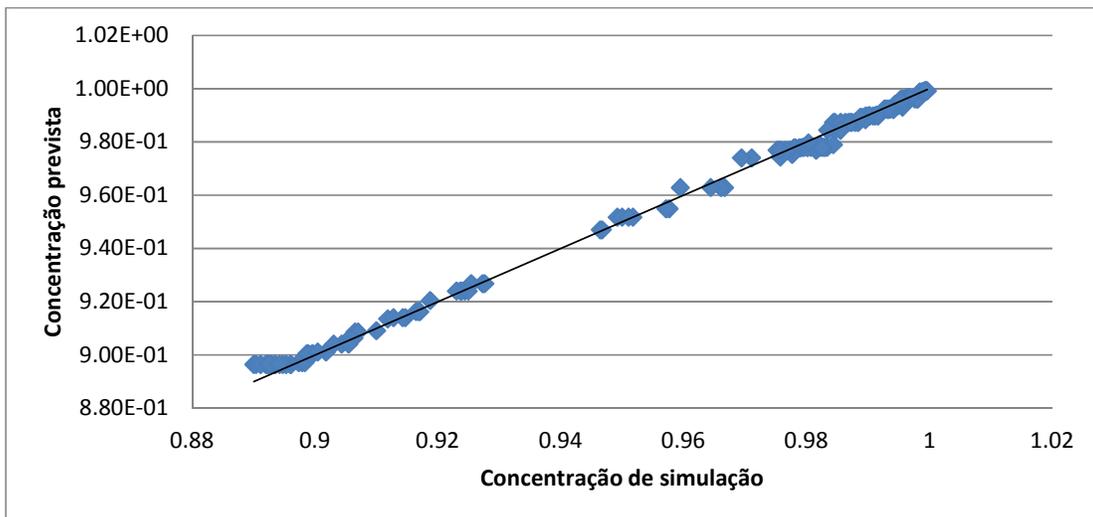


Figura 5.12: Resultado da predição para o mapa treinado com tamanho normal.

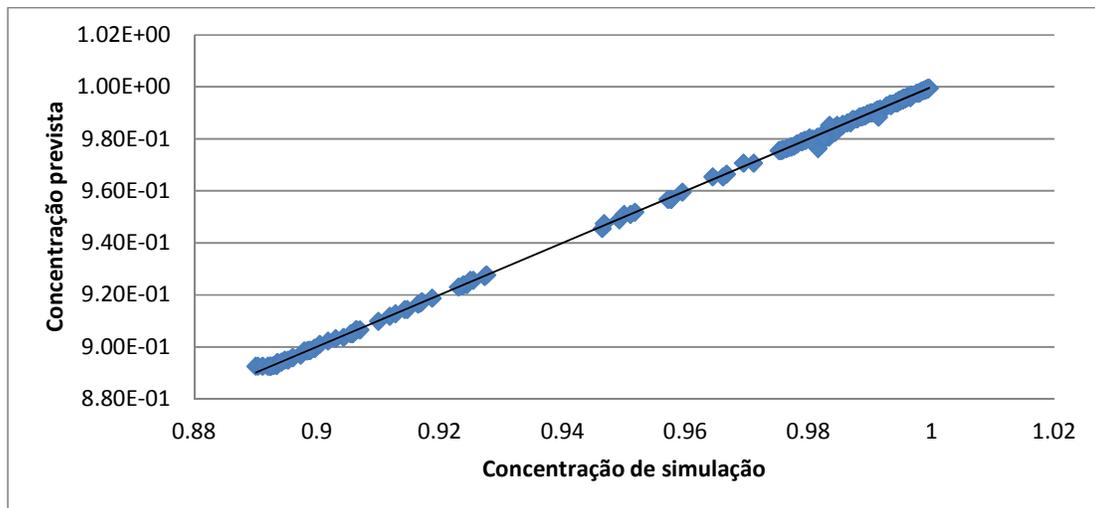


Figura 5.13: Resultado da predição para o mapa treinado com tamanho grande.

O mapa de tamanho pequeno para diversas faixas de concentrações de saída de propeno prevê o mesmo valor de saída, isso se deve ao reduzido número de neurônios treinados pelo mapa. Como visto anteriormente, cada neurônio tem um valor desejado associado, sendo assim o mapa de tamanho pequeno tem menos possibilidades de concentrações de saída resultando em maiores erros. Mesmo assim, o maior erro relativo calculado para esse mapa foi de 2%. O mapa de tamanho normal apresenta a mesma tendência de agrupamento de dados de saída do mapa pequeno, porém já começa a representar melhor algumas regiões de concentrações. Para o mapa de tamanho normal o maior erro relativo apresentado foi de 0,8%.

Como previsto, o mapa de tamanho grande apresentou os menores erros para a predição das amostras quando comparados com os outros tamanhos de mapas. O maior erro relativo apresentado por esse mapa foi de 0,5%, porém mais de 93% dos dados analisados apresentaram erros relativos menores que 0,1%. Pela análise dos desvios padrões dos três diferentes mapas é possível ver que o mapa de tamanho grande tem também a menor dispersão em torno da média, o que pode ser muito vantajoso. Os desvios padrões calculados são 0,57, 0,18 e 0,06 para os mapas de tamanho pequeno, normal e grande, respectivamente.

Foi também calculado o coeficiente de correlação entre os dados previstos e os de predição para os três diferentes tamanhos de mapa. Para o mapa de tamanho pequeno tem-se um valor de R^2 igual a 0,9924, para o mapa de tamanho normal tem-se R^2 igual a 0,9991 e para o mapa de tamanho grande o R^2 é igual a 0,9999. Esses valores para o coeficiente de correlação demonstram que tanto o mapa de tamanho normal quanto o de tamanho grande apresentam resultados excelentes para a predição da variável estudada.

O problema do agrupamento de concentrações, mais acentuado nos mapas de tamanho pequeno e normal, pode ser atenuado com uma melhora das amostras do banco de dados alimentado à rede neural para treinamento. Com a identificação das regiões em que o problema acontece é possível gerar mais dados nessa região e assim a mesma ser representada de forma mais satisfatória. A realização desse passo se faz conforme as necessidades do usuário, ou seja, é aconselhável treinar o mapa com mais dados da região que a planta opera normalmente e na qual será feita a etapa de predição de variáveis.

O boxplot é uma ferramenta estatística muito útil para a visualização gráfica de medidas de tendências centrais e de dispersão. No gráfico são mostradas as medianas, primeiro e terceiro quartis e os limites superiores e inferiores, calculados multiplicando a distância interquartílica por 1,5. O boxplot com os valores da concentração prevista pela técnica divididos pelos valores de simulação para os três tamanhos de mapas está representado na Figura 5.14.

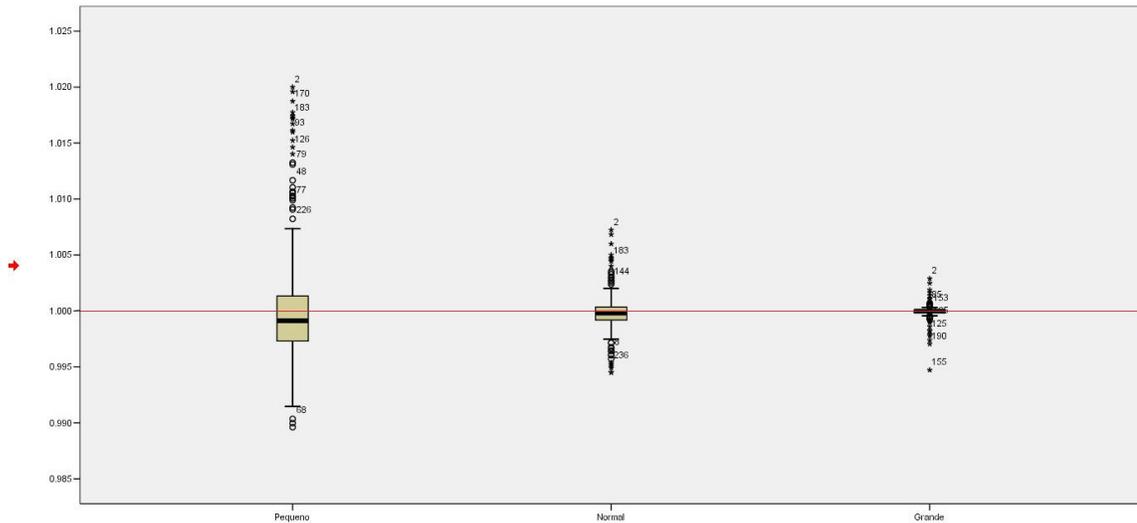


Figura 5.14: Boxplot para os três diferentes tamanhos de mapas.

Como visto anteriormente, o mapa de tamanho grande é o que mais se aproxima do valor ideal, representado pela linha vermelha e igual a unidade, uma vez que os dados alimentados para a construção do boxplot foram as concentrações previstas divididas pela concentração de simulação. Pela análise do boxplot é fácil ver também que o mapa de tamanho grande apresenta a menor dispersão dos dados em torno da média, seguido pelo mapa de tamanho normal e o mapa de tamanho pequeno.

A construção do boxplot também facilita a visualização dos pontos que estão mais afastados do valor ideal. Com a identificação desses pontos é possível alimentar mais dados dessa região na fase de treinamento do mapa para que a mesma fique mais representada. Como o mapa só depende dos dados alimentados, quanto melhor e mais aperfeiçoado for o banco de dados, menores os erros e maior a confiabilidade na definição das regiões operacionais da planta vão ser conquistados.

6 Conclusões e trabalhos futuros

A numerosa quantidade de variáveis medidas em planta e a quantidade de dados armazenados no banco de dados se apresentam como as grandes dificuldades em interpretar o funcionamento de uma operação. No que se refere a esses dois problemas, a técnica de mapas auto-organizáveis se mostrou uma ferramenta com resultados muito satisfatórios. Em uma etapa posterior de treinamento supervisionado do mapa foi predita a concentração de propeno no produto de topo da coluna de separação de propano e propeno.

No presente trabalho foi construída a simulação em Aspen Plus de duas colunas pertencentes a uma refinaria nacional. A partir dos dados obtidos de simulação estacionária variando características operacionais da unidade, construíram-se as matrizes-U tanto usando todos os dados como após um pré-processamento dos mesmos utilizando um algoritmo desenvolvido para remoção de dados similares. Posteriormente treinou-se o mapa de forma supervisionada e realizou-se a predição da concentração de propeno na corrente de topo comparando-a com os dados da simulação.

No que tange à construção da simulação, os erros obtidos, quando comparados com dados de projeto, foram muito baixos tanto para vazões totais como concentração de propeno e temperatura. Através da ferramenta de análise de sensibilidade variaram-se características operacionais da planta e coletaram-se dados de diversas medições. De posse do banco de dados construído foi possível passar para a próxima parte do trabalho: a aplicação dos mapas auto-organizáveis.

A etapa não-supervisionada de treinamento, apresentada na forma de matrizes-U e de trajetórias crescentes de variáveis, reduziu a dimensionalidade dos dados de entrada a mapas em duas ou três dimensões. A partir da análise dos mesmos foi possível delimitar regiões de operação da coluna de separação propano e propeno. O mapa de tamanho 20x5 foi o escolhido dentre todos para a visualização de trajetórias e relações entre características operacionais e a concentração de propeno no produto de topo.

A visualização das trajetórias foi feita de acordo com as quatro características de operação da coluna de separação propano e propeno variadas na simulação em Aspen Plus: pressão de saída do compressor, fração de corrente desviada para o produto de topo, vazão de *hot bypass* e vazão total de alimentação da coluna. Para cada conjunto de dados traçaram-se as trajetórias crescentes tanto da concentração de propeno no topo como da variação operacional correspondente. A partir dos pares de matrizes-U com as trajetórias traçadas foi possível visualizar a evolução de cada uma no mapa e estabelecer relações entre elas.

A predição de variáveis, realizada após um treinamento supervisionado dos dados de entrada, obteve resultados excelentes para a concentração de propeno no produto de topo da coluna de separação propano e propeno, com coeficientes de correlação entre os dados simulados e preditos muito próximos da unidade.

A técnica de self-organizing maps mostrou-se, desse jeito, uma técnica robusta para a prospecção de conhecimento através do processamento de bancos de dados. A interpretação dos mapas treinados e a filtragem dos dados alimentados são funções que cabem ao especialista, assim, o conhecimento da planta e a alteração dos parâmetros de treinamento do mapa têm fundamental importância no resultado final.

6.1 Trabalhos futuros

Como continuidade desse trabalho pode-se utilizar, ao invés de dados de simulação estacionária, dados reais da planta em operação. Também é possível estender a identificação de regiões operacionais e a predição de variáveis de interesse para as outras colunas da unidade. A melhora no banco de dados e a conseqüente melhor representação das regiões se apresentam como tarefas que devem estar sendo sempre aprimoradas e não podem ser cessadas.

7 Referências

- Aguado, D., Montoya, T., Borrás, L., Seco, A. & Ferrer, J. (2008). Using SOM and PCA for analysing and interpreting data from a P-removal SBR. *Engineering Applications of Artificial Intelligence*, **21**(6), 919-930.
- Domínguez, M., Fuertes, J. J., Reguera, P., Díaz, I. & Cuadrado, A. A. (2007). Internet-based remote supervision of industrial processes using self-organizing maps. *Engineering Applications of Artificial Intelligence*, **20**(6), 757-765.
- Fuertes, J. J., Domínguez, M., Reguera, P., Prada, M. A., Díaz, I. & Cuadrado, A. A. (2010). Visual dynamic model based on self-organizing maps for supervision and fault detection in industrial processes. *Engineering Applications of Artificial Intelligence*, **23**(1), 8-17.
- Haykin, S. S. (1999). *Neural networks : a comprehensive foundation. 2nd ed.* Prentice Hall, Upper Saddle River, N.J.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**, 43:59-69.
- Ng, Y. S. & Srinivasan, R. (2008a). Multivariate temporal data analysis using self-organizing Maps. 1. Training methodology for effective visualization of multistate operations. *Industrial and Engineering Chemistry Research*, **47**(20), 7744-7757.
- Ng, Y. S. & Srinivasan, R. (2008b). Multivariate temporal data analysis using self-organizing Maps. 2. Monitoring and diagnosis of multistate operations. *Industrial and Engineering Chemistry Research*, **47**(20), 7758-7771.
- Ouzounoglou, A. N., Economopoulos, T. L., Asvestas, P. A. & Matsopoulos, G. K. (2010). Fingerprint matching with self organizing maps. *IFMBE Proceedings*. 307-310.
- Ritter, H. & Kohonen, T. (1989). Self-organizing semantic maps. *Biological Cybernetics*, **61**(4), 241-254.