



APLICAÇÃO DE MAPAS AUTO-ORGANIZÁVEIS EM ESPECTROSCOPIA DE FLUORESCÊNCIA

Andrea Cabral Farias, Jorge Otávio Trierweiler, Marcelo Farenzena

Grupo de Intensificação, Modelagem, Simulação, Controle e Otimização de Processos (GIMSCOP)
Departamento de Engenharia Química, Universidade Federal do Rio Grande do Sul (UFRGS)
R. Eng. Luis Englert, s/n. Campus Central. CEP: 90040-040 - Porto Alegre - RS - BRASIL,
E-MAIL: {dea, jorge, farenz}@enq.ufrgs.br

Resumo Para controlar e modelar efetivamente um processo de fermentação é importante monitorar o processo continuamente. Do ponto de vista operacional, o controle de processos implica na necessidade de caracterização dos estados do sistema e, quando aplicado a bioprocessos, o grande problema encontra-se justamente nesta caracterização. O objetivo deste trabalho é aplicar a técnica de Mapas auto-organizáveis (SOM) para monitorar um processo fermentativo através de dados de espectroscopia de fluorescência 2D. Para isto foram utilizados dados reais já publicados em SOLLE *et al.*(2003) e RANZAN *et al.*(2011). A técnica foi aplicada para analisar estes dados experimentais e foi adicionada uma etapa de aprendizado supervisionado para estimar os estados do processo. Os resultados mostram que é possível visualizar claramente as diferentes etapas do processo, confirmando a viabilidade de utilização do SOM (construído somente a partir nos espectros de fluorescência) para monitoramento do processo. Além disto, foi avaliada a obtenção de estimativas das variáveis do processo, após a aplicação de uma etapa supervisionada. Os resultados obtidos na estimação foram promissores.

Palavras-chave: Mapas Auto Organizáveis de Kohonen, Espectroscopia de Fluorescência. Reconhecimento de Padrões

1. Introdução

Para controlar e modelar efetivamente um processo de fermentação é importante monitorar o processo continuamente (RHEE *et al.*, 2005). Do ponto de vista operacional, o controle de processos implica na necessidade de caracterização dos estados do sistema, sendo esse um dos grandes problemas relacionado aos bioprocessos (RANZAN *et al.*, 2011).

Os sensores ópticos começam a ser utilizados como instrumentos de medição em biotecnologia por permitir um monitoramento em linha não invasivo, não destrutivo e com medições rápidas. Dentre estes, o sensor de fluorescência torna possível o monitoramento de componentes intracelulares. Os sensores de espectrofluorescência bidimensional permitem amostrar uma grande faixa de pares excitação-emissão simultaneamente e têm sido utilizados para monitorar diferentes processos de fermentação. (SOLLE *et al.*, 2003; RHEE *et al.*, 2005)

A espectroscopia de fluorescência proporciona medições de diferentes intensidades de fluorescência com muitas combinações de pares comprimentos de onda de excitação-emissão. Embora os conjuntos de dados de fluorescência espectral incluam uma série de informações úteis sobre o estado fisiológico das células e do desempenho dos processos biológicos, não é fácil lidar com uma grande quantidade de dados espectrais utilizando os métodos estatísticos tradicionais. Alguns algoritmos de

estatística avançada, como análise de componentes principais (*Principal component analysis*, PCA) e regressão por mínimos quadrados parciais (*Partial Least Square regression*, PLS) podem ser usados para reduzir a grande quantidade de dados espectrais e extrair informações importantes visando o monitoramento e a obtenção de um modelo de processo. Mas o PCA e o PLS são lineares na natureza e processos de fermentação podem apresentar relações não-lineares entre os parâmetros do processo e os espectros de fluorescência, prejudicando a utilização das técnicas lineares.

Para lidar com este problema pode-se fazer uso dos Mapas Auto-Organizáveis de Kohonen (*Self Organizing Maps* – SOM, KOHONEN, 1998). O SOM é uma rede neural com aprendizado não supervisionado e é considerado o substituto não linear do PCA. Esta técnica tem sido empregada com sucesso para analisar grandes conjuntos de dados em diferentes áreas, como por exemplo: monitoramento de processos e visualização (YEW & SRINIVASAN, 2008 (a) e (b)), detecção de falhas (BHUSHAN & ROMAGNOLI, 2008). Ainda é possível adicionar uma etapa de aprendizagem supervisionada ao algoritmo e então estimar variáveis de processos complexos (ABONYI *et al.*, 2003; ALHONIEMI *et al.*, 1999). Neste caso, o mapa auto-organizável é usado para particionar o espaço de entrada e então é realizada uma modelagem local. Desta forma ainda é possível utilizar a técnica para predição de séries temporais (TRYBA & GOSER, 1991).

O SOM é considerado, ainda, uma excelente ferramenta para fazer o mapeamento não linear de dados em espectroscopia. Trabalhos prévios utilizam a técnica em conjunto com espectroscopia de ressonância magnética (NMR) para agrupar (*clustering*) o espectro (LOW *et al.*, 2004), ou ainda, combinado com espectroscopia de mobilidade iônica para monitoramento em linha e visualização de dados de fermentação (KOLEHMAINEN *et al.*, 2004). Para interpretação de processos de fermentação, RHEE *et al.*, 2005 utilizou o SOM para classificação de diferentes processos de fermentação baseados nos espectros de fluorescência 2D.

O objetivo deste trabalho é utilizar o Mapa auto-organizável para monitorar o processo fermentativo de glicose com a levedura *Sccharomyces cerevisiae* através de dados de espectroscopia de fluorescência 2D. Para isto foram utilizados dados reais já publicados em (SOLLE *et al.*, 2003 e RANZAN *et al.*, 2011). Num segundo momento, de maneira supervisionada, são estimados parâmetros do processo. Os resultados mostram que é possível visualizar claramente as diferentes etapas do processo e, após uma etapa supervisionada, obter excelentes estimativas das variáveis do processo.

2. Mapas auto-organizáveis (SOM)

O mapa auto-organizável é uma rede neural não supervisionada que foi proposta por KOHONEN (1998). Esta é capaz de projetar dados de grandes dimensões em uma grade bidimensional e pode servir como ferramenta de visualização. Auto-organização quer dizer que a rede orienta e adaptativamente assume a forma que pode melhor descrever os vetores de entrada. SOM aplica uma regressão não paramétrica de vetores discretos e ordenados a uma distribuição de vetores de características de entrada. Um número finito de vetores de referência são adaptativamente colocados no espaço do sinal de entrada para aproximar tais sinais.

Considere um conjunto de dados X que contem I amostras, cada uma N -dimensional. O conjunto X é uma matriz com duas dimensões de tamanho $I \times N$, com a i -ésima amostra sendo $x_i = \{x_{i1}, \dots, x_{in}, \dots, x_{iN}\}$. O SOM é uma coleção ordenada de neurônios. Cada neurônio tem um vetor de referência associado $m_j \in \mathfrak{R}^n$ então, $m_j = \{m_{j1}, \dots, m_{jn}, \dots, m_{jN}\}$. Considerando que um SOM, descrito com $m_j = \{m_{j1}, \dots, m_{j1}, \dots, m_{jN}\}^T$, com J neurônios, que deve ser treinado para representar e visualizar X . Isso envolve o cálculo do vetor de referência para cada neurônio. Inicialmente, fazendo com que cada m_j seja associado com um vetor aleatório do domínio de X . Quando a amostra $x_i \in X$ é apresentada para o treinamento do SOM, o neurônio onde o vetor de referência tem a menor diferença de x_i é identificado e definido como o vencedor ou a unidade com melhor combinação (*BMU* – *best matching unit*) para aquela entrada.

$$b_i = \arg \min_j \|x_i - m_j\| \quad \forall j \in [1, J] \quad (1)$$

A distância $\| \dots \|$ entre x_i e m_j é medida usando a distância euclidiana, mas outras métricas podem também ser usadas.

Os neurônios em M^{SOM} são geralmente colocados em uma grade bidimensional. Tomando a localização do j -

ésimo neurônio como sendo r_j , onde $r_j \in \mathfrak{R}^2$. A métrica de distância pode ser definida na grade bidimensional e todos os neurônios dentro de certa distância do j -ésimo neurônio podem ser considerados como seus vizinhos topológicos N_j na grade. Este conceito de vizinhança do neurônio é o elemento-chave diferenciador de características do SOM e é responsável por suas propriedades unidas (descritas abaixo). Durante o treinamento, quando cada amostra $r_j \in X$ é apresentada, o vetor de referência da BMU, m_{b_i} , assim como sua vizinhança topológica na grade, são atualizados movendo-os para perto da amostra de treinamento x_i . Na sua forma mais simples, a regra de aprendizado do SOM na t -ésima iteração é dada por:

$$m_j(t+1) = m_j(t) + \alpha(t) h_{b_i j}(t) \|x_i(t) - m_j(t)\| \quad (2)$$

Onde $\alpha(t)$ é o fator de taxa de aprendizagem e $h_{b_i j}(t)$ é a função vizinhança centrada na BMU b_i mas definida sobre todos os neurônios em M^{SOM} . A função de vizinhança gaussiana e a mais comumente utilizada e dada por:

$$h_{b_i j}(t) = \exp\left(-\frac{\|r_{b_i} - r_j\|^2}{2\sigma^2(t)}\right) \quad (3)$$

Onde $\sigma(t)$ é a largura da vizinhança. A largura da vizinhança varia de iteração a iteração mudando $\alpha(t)$. Grandes valores de σ e α são usados inicialmente e usualmente decrescem monotonicamente com t . Para garantir a convergência, é necessário que a medida que o treinamento prossegue e $t \rightarrow \infty$, $\alpha(t) \rightarrow 0$ e $\sigma(t)$ se aproxima de um valor pequeno (tipicamente 1). Outras variantes da função de vizinhança, assim como os algoritmos de treinamento, têm sido propostos na literatura (ZUPAN *et al.*, 1999).

A atualização do vetor de referência dos neurônios da vizinhança, juntamente com a da BMU, dão uma característica de preservação da topologia para o SOM (i.e, a vizinhança dos neurônios são ativadas e aprendem a partir da entrada de treino x_i) e então adquirem vetores de referência similares. Portanto, as unidades de vizinhança são mais semelhantes entre si e os neurônios próximos, nos mapas SOM treinados apresentam similares amostras de entradas. Após o treino, pode-se usar o mapa para a classificação de novas amostras. Neste caso, o índice do BMU pode ser considerado com o índice de classe. Para qualquer ponto de teste x_i , uma classe pode ser atribuída encontrando a sua BMU.

A principal aplicação do SOM é para visualização de dados com muitas dimensões. Um SOM treinado pode ser pensado como um mapeamento de $X \in \mathfrak{R}^n$ em duas dimensões. Os neurônios em um SOM treinado não são igualmente distribuídos ao longo de todo o espaço de entrada \mathfrak{R}^n , mais neurônios são designados para regiões com mais amostras em X (alta densidade) e alguns poucos para regiões de baixa densidade no \mathfrak{R}^n . Portanto, uma maneira de visualizar *clusters* em X é pela distância entre o neurônio e a sua vizinhança.

$$D_{jj'} = \|m_j - m_{j'}\| \quad j' \in N_j \quad (4)$$

A *matriz de distância unificada* (U-matrix) visualiza o SOM diferenciando o limite entre cada par de neurônios por uma cor que é proporcional a sua distância $D_{jj'}$. Alternativamente, a distância média entre o neurônio e sua vizinhança pode ser usada para colorir o neurônio.

$$D_j = \frac{1}{|N_j|} \sum_{j'} \|m_j - m_{j'}\| \quad j' \in N_j \quad (5)$$

Onde $|N_j|$ é o tamanho da vizinhança. As bordas do cluster são então indicadas por montanhas de grandes distâncias que separam vales de neurônios similares. Existe ainda outra maneira de visualização. Nessa, os *clusters* podem ser rotulados e diretamente usados como uma exibição tridimensional para descrever uma nova amostra x_i do espaço X. Essa é particularmente útil para identificar classes (*clusters*) de uma nova amostra. A *BMU* correspondente a este novo x_i é muitas vezes referida como seu *hit*. Se novas amostras estão regularmente disponíveis a partir de um processo online, o estado do processo em cada período de tempo pode também ser visualizado como uma sequência de hits e esta é a forma de visualização de séries temporais.

Neste trabalho foi utilizado um pacote de programas para *Matlab*, fornecido gratuitamente por KOHONEN (2001), conhecido como *somtoolbox*. A escolha deste pacote se deu por sua popularidade entre os pesquisadores da área. Foram realizados diferentes testes variando os mais diversos parâmetros e que não serão apresentados neste trabalho. Os parâmetros utilizados para a geração dos mapas foram obtidos a partir de sugestões encontradas em KOHONEN (2001) e ZUPAN (1999). Muitos deles já fazem parte dos *defaults* encontrados no *somtoolbox*.

3. Espectroscopia de Fluorescência Bidimensional

A espectroscopia de fluorescência, ou espectrofluorometria, analisa a fluorescência de uma amostra como resultado de um processo emissão/re-emissão de luz de baixa energia, normalmente luz no comprimento do visível a ultravioleta. O princípio dos espectrofluorômetros se baseia na re-emissão de luz com mudança espectral, que algumas espécies químicas sofrem. A luz re-emitida é proporcional a concentração de espécies químicas do meio analisado e possui comprimento de onda igual ou superior ao de emissão. (SOLLE *et al.*, 2003).

Os dados experimentais utilizados neste trabalho foram fornecidos pelo professor B. Hitzmann, do Institut für Technische Chemie, University of Hannover e utilizados nos trabalhos de SOLLE *et al.* (2003) e RANZAN *et al.*, 2011(b). Utilizou-se a levedura *S. cerevisiae* H620, crescendo em biorreator de 1,5 L à temperatura constante de 30 °C e pH 5,5 e meio Schatzmann com glicose. Os espectros fluorescentes foram coletado a cada 6 minutos pelo BioView-espectrômetro (Delta Light & Optics, Dinamarca) como descrito por STÄRK *et al.* (2002). Cada espectro é constituído de 150 pares de comprimento de onda de excitação/emissão, composto com 15 filtros na região de 270 à 550 nm para excitação e 15 filtros na região de 310 à 590 nm para emissão com largura de banda de 20 nm cada e coletado equidistante 20 nm.

Foram utilizados dados de 2 cultivos, onde 201 espectros foram coletados e, desta forma, avaliou-se 30150 pontos individuais no decorrer do processo. As concentrações de glicose, etanol e biomassa no decorrer dos processos fermentativos foram obtidas a partir de um modelo dinâmico. Descrições detalhadas da estrutura de modelo utilizada, parâmetros ajustados e metodologia empregada na simulação dos ensaios fermentativos podem ser obtidos em RANZAN *et al.*, 2011(b). Vale ressaltar que o processo fermentativo apresenta duas etapas de interesse. Na primeira, a fonte de carbono é a glicose e ocorre a produção de etanol. Quando a glicose se extingue, o etanol passa a ser a fonte de carbono e é consumido.

4. Resultados e Discussões

Para gerar o Mapa Auto-Organizável, os dados de cada espectro foram transformados em um vetor unidimensional, onde cada ponto corresponde a um par de intensidades de excitação-emissão. O conjunto dos espectros pode ser visto como uma representação de uma série temporal. Para melhor visualização, a evolução dos estados é demonstrada na Figura 1.

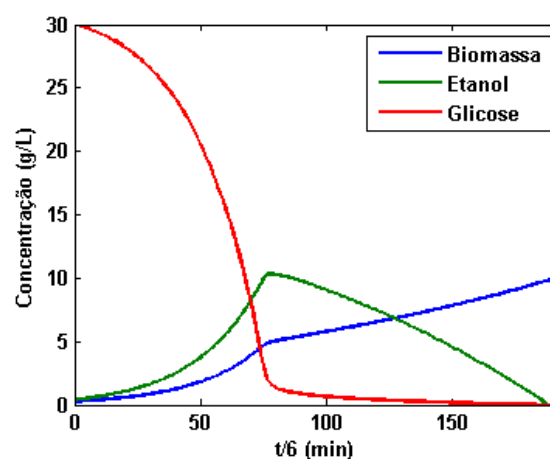


Figura 1. Estados do processo em função do tempo

Já na Figura 2 são apresentados três espectros obtidos em momentos diferentes. O primeiro, coletado em $t/6=20$ min. (120 min.), e refere-se à etapa de consumo de glicose e produção de etanol. O segundo, amostrado em $t/6=156$ min. (468 min.), refere-se ao ponto com maior concentração de etanol e mínima de glicose. Já o terceiro, amostrado em $t/6=360$ min. (1080 min.) refere-se ao final da reação, com pequena concentração de etanol e a maior de biomassa. Observa-se, como esperado, a correlação existente entre as concentrações e a intensidade do espectro em diferentes regiões. No procedimento de análise do espectro, normalmente é realizada uma etapa de calibração, muito dispendiosa tanto em tempo quanto em recurso.

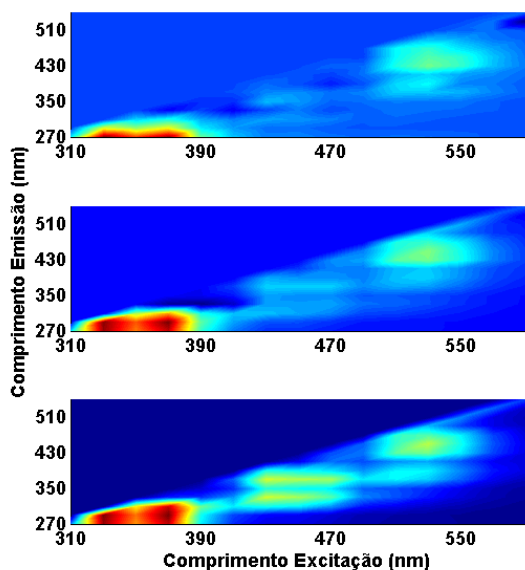


Figura 2. Espectros de fluorescência obtidos em $t=[120\ 468\ 1080]$. Mapa de cores cresce do azul para o vermelho.

E é justamente para que a etapa de calibração não necessite ser realizada que se justifica a utilização o mapa auto-organizável. Ao invés de calibrar todas as regiões do espectro com todos os componentes possíveis em todas as etapas da reação, busca-se calibrar somente um mapa bidimensional utilizando somente as variáveis de interesse.

Para isto, foram realizados os seguintes procedimentos:

1. Os dados de cada espectros foram transformados em um vetor unidimensional
2. Os vetores (conjunto de espectros) foram normalizados de maneira a se obter uma variância igual a 1.
3. Os vetores foram utilizados para o treinamento do mapa auto-organizável. Este foi inicializado linearmente, treinado de acordo com os parâmetros já explicitados e foi obtido o Mapa.
4. Os dados são projetados no mapa para encontrar alguma relação com os clusters obtidos.

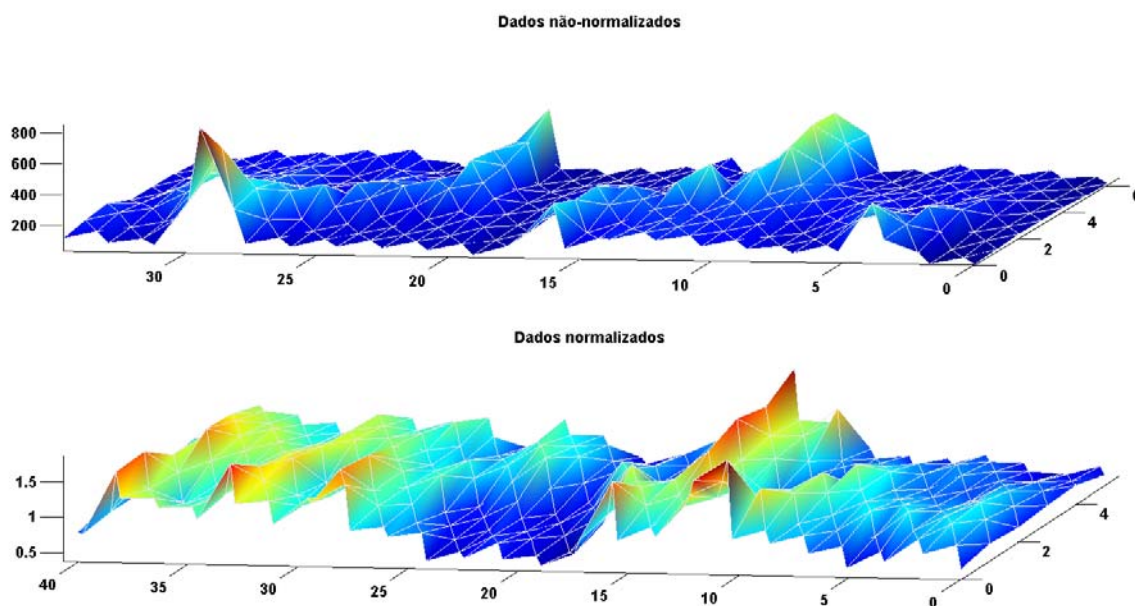


Figura 3. Mapas Auto-Organizáveis obtidos com os diferentes dados

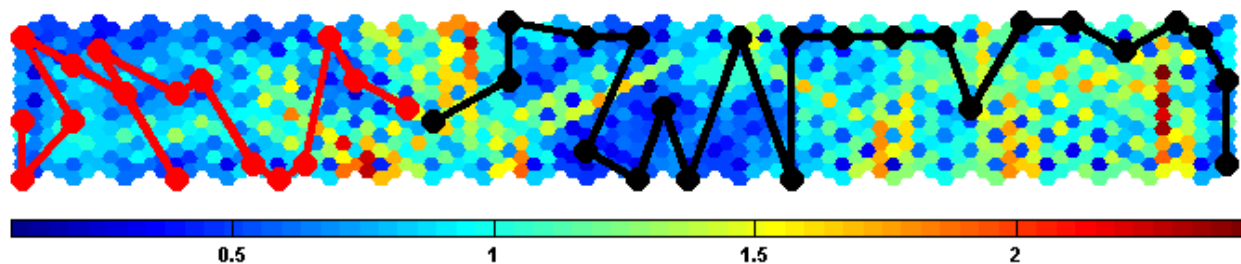


Figura 4. Mapa Auto-Organizável obtido com os dados normalizados e séries temporais. (vermelho $t=1$ (E):78 (D) min, preto $t=79$:193min)

Na Figura 3 são apresentadas duas versões do mapa tridimensional. A primeira foi obtida com os dados não normalizados e a segunda com a normalização dos dados. Nas duas versões os clusters estão localizados nas

regiões com pouca distância entre os neurônios (tons azulados) e as bordas entre os clusters são as regiões com afastamento mais pronunciado entre neurônios vizinhos. Além disto, verifica-se que a normalização apresenta um

papel importante visto que o segundo mapa é bem mais informativo que o primeiro.

Na Figura 4 é demonstrada a versão bidimensional dos dois mapas e o segundo mapa apresenta a calibração pela série temporal. A linha vermelha corresponde à etapa da produção de etanol e a linha preta corresponde ao consumo do etanol. Nas duas figuras os clusters estão localizados nas regiões com pouca distância entre os neurônios (tons azulados) e as bordas entre os clusters são as regiões com afastamento mais pronunciado entre neurônios vizinhos. Isto pode ser visto pela quantidade de regiões com pequenas distâncias (azuis). Nessas regiões os dados são considerados idênticos mas com algum ruído (se fossem idênticos não teriam um neurônio específico para representá-los no mapa). Segundo YEW et al. (2008 a), isto pode acontecer se os dados não forem normalizados ou ainda se uma condição operacional, por exemplo um estacionário do processo, for mais bem representado que as outras condições. Esse não é o caso neste trabalho, mas é

interessante salientar esta semelhança de comportamento.

Ainda na Figura 4, fica claro que este mapa pode ser utilizado para monitorar o processo sem que nenhuma análise off-line (i.e., concentrações de Glicose, Etanol e Biomassa) seja informada. Dado um espectro de fluorescência, pode-se encontrar a região no mapa correspondente a alguma etapa da reação, por exemplo, verificar se o etanol está sendo produzido (até o último ponto em vermelho à direita) ou utilizado como fonte de carbono (na trajetória ao longo da linha preta). E para isto, após a etapa de treinamento, é somente necessário armazenar uma matriz bidimensional, correspondente ao peso dos neurônios na grade (neste caso o tamanho do mapa é de 47x6). Mas cabe salientar que é necessário o conhecimento qualitativo do processo, para se poder associar as diferentes regiões encontradas a etapas esperadas para o processo fermentativo.

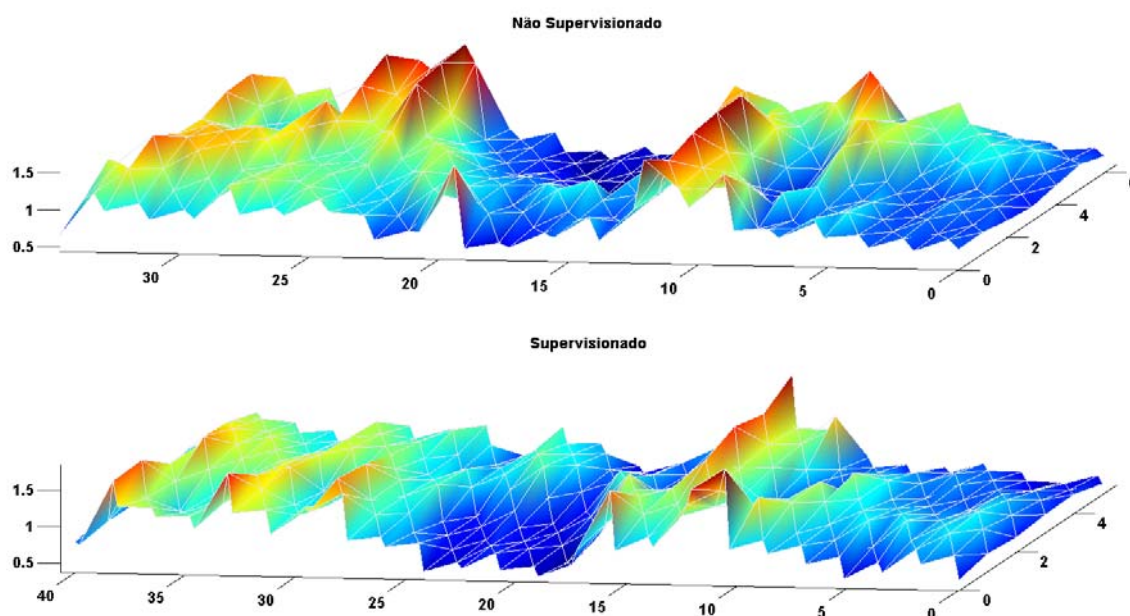


Figura 5. Mapas obtidos de maneira não-supervisionada e supervisionada.

Além da utilização do mapa auto-organizável como ferramenta para monitoramento, ainda existem algumas outras propriedades interessantes da técnica.

Segundo KOHONEN (2001), como o mapa mantém a topologia dos dados de entrada garantindo um caráter não linear ao mapa treinado, é possível, obter bons modelos locais para estados de um processo, por exemplo. Existem diferentes maneiras para que isso seja realizado e todas baseadas numa etapa adicional de aprendizado supervisionado. Nessa, parte dos dados é utilizada para treinar o mapa e é necessário apresentar os estados do processo como variáveis adicionais (etapa supervisionada). A segunda parte dos dados é utilizada para validação.

Nesse trabalho o mapa foi treinado com os dados dos espectros e as concentrações de glicose, etanol e biomassa obtidos no primeiro experimento. O objetivo é estimar as concentrações a partir dos espectros do primeiro

experimento. Essa é uma tarefa difícil de ser realizada com estatística convencional, por exemplo. Em RANZAN et al. (2011) é apresentado um extenso trabalho que trata deste assunto. Esta tarefa é realizada com base em análise de correlação entre subconjuntos de medidas dos espectros e as variáveis de estado do processo. Uma das principais dificuldades é realizar a busca exaustiva pelos subconjuntos representativos. Os resultados aqui apresentados têm por finalidade tratar do mesmo sujeito

Na Figura 5 é apresentada a comparação entre o mapa obtido na etapa anterior e o mapa obtido após a etapa de aprendizagem supervisionada. Pode-se ainda ver que a adição das variáveis gera uma inversão na borda do cluster localizada na região $15 < x < 10$. Mesmo assim é evidente a presença de 2 clusters bem definidos em ambos os mapas. A partir do mapa obtido de forma supervisionada foi realizada a estimação dos estados

referentes ao 2º experimento. Esta foi realizada pela busca *BMU* de cada espectro, que foi associada com um valor de concentração na etapa de treinamento. Quando necessário foi utilizada a relação entre as distâncias para fazer regressão. Os resultados obtidos na estimação das concentrações de etanol, glicose e biomassa são apresentados na Figura 6, assim com o erro absoluto de estimação. Para avaliar o ajuste foi calculado o coeficiente de correlação entre as curvas. Os valores obtidos para as concentrações de etanol, glicose e biomassa foram de: 0.9973, 0.9982 e 0.9993 respectivamente.

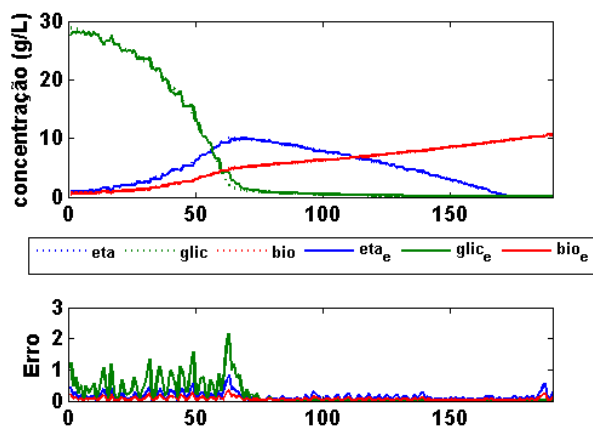


Figura 6. Estimação das concentrações de Etanol, Glicose e Biomassa e erro absoluto de estimação

Estes valores são tão bons quanto os obtidos por RANZAN (2011) com a diferença de não ser necessário realizar o procedimento de busca exaustiva para as melhores combinações entre variáveis, entretanto, com a necessidade de se utilizar todos os comprimentos de onda do espectro.

5. Conclusão

O objetivo deste trabalho foi avaliar a utilização dos Mapas Auto-organizáveis para monitorar um processo fermentativo através de dados de espectroscopia de fluorescência 2D. Utilizando dados reais, a construção dos mapas se mostrou simples e os resultados demonstrados indicaram que é possível monitorar o processo utilizando somente os espectros de fluorescência. Este é um ganho muito importante visto que a redução no número de variáveis é muito grande, mesmo para um processo com relações não lineares e as características não lineares da técnica garantem a consistência matemática do mapa. Esse resultado é significativo e possibilita que a técnica seja utilizada no futuro para processos mais complexos, como por exemplo, fermentação com meio não caracterizado.

Num segundo momento, de maneira supervisionada, foram estimados parâmetros do processo. Os resultados mostram, apesar da técnica não ter como objetivo a estimação de parâmetros, que é possível obter excelentes estimativas das variáveis do processo.

6. Referências

ABONYI J., NEMETH S., VINCZE C., ARVA P., Process analysis and product quality estimation by Self-Organizing Maps with an application to polyethylene production, *Computers in Industry*, 52, (3), 221-234, 2003

ALHONIEMI E., HOLLMÉN J., SIMULA O., VESANTO J., Process Monitoring and Modeling using the Self-Organizing Map, *Integrated Comp.-Aided Engng*, 6 (1) 3-14, 1999

BHUSHAN B. & ROMAGNOLI J.A., Self-organizing self-clustering network: A strategy for unsupervised pattern classification with its application to fault diagnosis. *Ind. & Engng. Chem. Research*, 47 (12), 4209-4219, 2008

DEBELJAK, Z., STRAPAC, M., MEDIC-SARIC, M., Application of selforganizing maps for the classification of chromatographic systems and prediction of values of chromatographic quantities, *J. Chromatogr. A* 925, 31-40, 2001.

KOHONEN, T., *Self-Organizing Maps*, third ed., Springer, Berlin, Germany, 2001.

KOHONEN, T., The self-organizing map, *Neurocomputing* 21, 1-6, 1998.

KOLEHMAINEN, M., RONKKO, P., RAATIKAINEN, O., Monitoring of yeast fermentation by ion mobility spectrometry measurement and data visualization with self-organizing maps, *Anal. Chim. Acta*, 484, 93-100, 2003.

LOW, L.D., KALELKAR S., DOW E.R., Self-organizing maps for the analysis of NMR spectra, *BioSilico* 2, 157-163, 2004.

RANZAN, C., TRIERWEILER, L., HITZMANN, B. e TRIERWEILER, J. O. Modelagem quimiométrica para caracterização de bioprocessos através de espectroscopia fluorescente-2D. In: SIMPOSIO NACIONAL DE BIOPROCESSOS 2011, 18. 2011, Caxias do Sul, Anais. (a)

RANZAN, C., TRIERWEILER, L., HITZMANN, B. e TRIERWEILER, J. O. Validação com Dados de Espectroscopia Fluorescente 2D de Modelo Dinâmico para Fermentações Batelada de *Saccharomyces cerevisiae*. In: SIMPOSIO NACIONAL DE BIOPROCESSOS 2011, 18. 2011, Caxias do Sul, Anais. (b)

RHEE, J.I., LEE, K.I., KIM, C. K., YIM, Y.S., CHUNG, S.W., WEI J. Classification of two-dimensional fluorescence spectra using self-organizing maps. *Biochem. Eng. J.*, 22:135-44, 2005.

SOLLE, D., GEISLER, D., STARK, E., SCHEPER, T., HITZMANN, B. Chemometric modeling based on 2D-fluorescence spectra without a calibration measurement. *Bioinformatics*. v. 19, p. 173-177, 2003.

TRYBA, V., GOSER, K. Self-organizing feature maps for process control in chemistry, *Artificial Neural Networks*, North- Holland, Amsterdam, pp. 847-852, 1991.

VESANTO J., SOM-based data visualization methods, *Intell. Data Anal.* 3, 111-126, 1999.

YEW N. & SRINIVASAN R., Multivariate Temporal Data Analysis Using Self-Organizing Maps. 1. Training Methodology for Effective Visualization of Multistate Operations, *Ind. Eng. Chem. Res.*, 47 (20), 7744-7757, 2008

ZUPAN, J., GASTEIGER J., *Neural Networks in chemistry and Drug Design*, 2º ed., Wiley-VCH, Weinheim, USA, 1999.