

Modelagem empírica de evaporador de fábrica de celulose por regressão linear múltipla sem entradas atrasadas

Silveira, B. S.^{1*}; Pinheiro, O. S.²; Costa, A. O. S.³; Costa Jr, E. F.³

1 Curso de Graduação em Engenharia Química, Universidade Federal do Espírito Santo, Alegre, ES, Brasil.

2 Celulose Nipo-Brasileira S.A., Belo Oriente, MG, Brasil.

3 Pós-Graduação em Engenharia Química e Graduação em Engenharia Química, Departamento de Engenharia Rural, Universidade Federal do Espírito Santo, ES, Brasil.

* e-mail: bryan_es_95@yahoo.com.br

Resumo

O processo Kraft de obtenção de celulose possui como principal vantagem o sistema de recuperação de produtos químicos associados a ele. Nesse processo, durante o cozimento da madeira no digestor ocorre a formação de um subproduto denominado licor preto. Esse licor preto contém grande quantidade de água e não pode aplicado diretamente na caldeira de recuperação. Por esta razão, este licor é concentrado em um sistema de evaporadores de múltiplo efeito. Diante da importância do sistema de evaporadores de múltiplo efeito para o processo, é interessante que sejam desenvolvidos modelos desses equipamentos. Neste artigo são apresentados modelos de regressão linear múltipla usando o software STATISTICA 9.1. Isso foi feito utilizando um conjunto de dados provenientes do banco de dados de uma empresa de celulose. Apesar de o processo ser dinâmico, foram testados somente modelos estáticos. A estratégia de modelagem se deu em acrescentar novas variáveis na regressão a partir das variáveis operacionais. Os resultados obtidos foram melhores do que os reportados na literatura para redes neurais sem entradas atrasadas.

Keywords: empirical modeling, parameters estimation, multiple effect evaporators

Abstract

The Kraft process for obtaining cellulose has as its main advantage the chemical recovery system. In this process occurs the formation of a byproduct called black liquor during the wood cooking in the digester. This black liquor contains a large amount of water that cannot be directly applied in the recovery boiler. For this reason, this liquor is concentrated in a multiple effect evaporator system. Given the importance of the multiple effect evaporator system for the process, the development of such equipment models becomes of interest. In this article, multiple linear regression models are presented using the STATISTICA 9.1 software. This was done using a data set from the database of a cellulose company. Although the process is dynamic, only static models were tested. The modeling strategy was done by adding new variables in the (initial) regression from operating variables. The results obtained were better than those reported in the literature for neural networks without delayed inputs.

Keywords: modelagem empírica, estimação de parâmetros, evaporadores de múltiplo efeito.

1. Introdução

O processo Kraft de produção de celulose possui como função dissolver a lignina, composto que confere rigidez a parede celular, com o objetivo de separar a lignina das cadeias de carboidratos (celulose e hemicelulose) com o mínimo de degradação das fibras. O sistema de recuperação de produtos químicos

associados a esse processo é sua principal vantagem [1]. Durante o cozimento da madeira, é produzido o licor preto, subproduto que é direcionado ao ciclo de recuperação. Esse fluido deve ser direcionado a caldeira de recuperação, mas o licor que sai do digestor possui concentração de sólidos em torno de 14 a 20%. A aplicação direta desse licor na caldeira resultaria em um desempenho ruim do processo, com

isso, é necessária a utilização de um sistema de evaporadores de múltiplo efeito para que a concentração de sólidos seja aumentada a no mínimo 65% [2].

Os modelos empíricos são aqueles que não são baseados em quaisquer pressupostos teóricos, ou seja, em sua formulação são empregados apenas dados experimentais. Cabe lembrar que grande parte das correlações utilizadas em engenharia são modelos empíricos [3]. A regressão múltipla a ser utilizada neste trabalho é uma técnica estatística que pode ser usada para analisar a relação entre uma única variável dependente e várias variáveis independentes [4]. O modelo de regressão linear múltipla é dado pela equação (1).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

em que: n = tamanho amostral; k = número de variáveis independentes; x_1, x_2, \dots, x_k são as variáveis independentes; β_0 = valor de y quando as variáveis preditoras são 0; $\beta_1, \beta_2, \dots, \beta_k$ são os coeficientes das variáveis preditoras; ε = erro aleatório dado pela diferença entre o valor observado de y e o valor obtido pelo modelo de regressão.

O objetivo deste trabalho é a modelagem empírica de um evaporador de múltiplo efeito de uma empresa de celulose por regressão linear múltipla sem entradas atrasadas.

2. Materiais e Métodos

Os dados utilizados foram coletados de um banco de dados de uma empresa de celulose por um período de 10 dias com intervalo de 1 minuto. Com isso, foi definido o conjunto de variáveis que seriam utilizados no modelo. Foram introduzidas novas variáveis, calculadas a partir das variáveis operacionais, com o objetivo de otimizar os resultados. As respostas obtidas eram avaliadas através do coeficiente de determinação e análise de resíduos.

O programa utilizado para o desenvolvimento dos modelos foi o software STATISTICA 9.1.

3. Resultados e Discussão

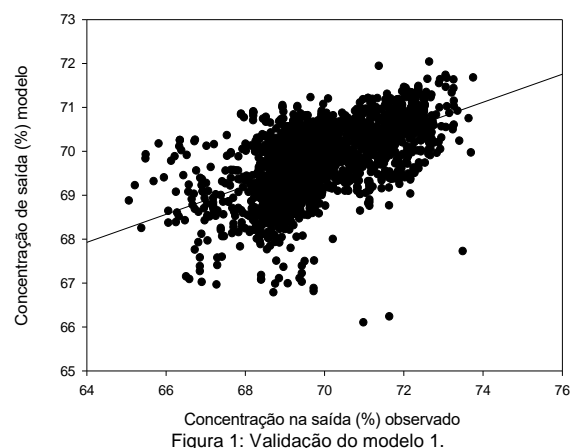
Inicialmente o conjunto de dados utilizados possuía 37441 pontos operacionais que foram reduzidos a

26904 pontos operacionais, sendo que 90% deles foram utilizados para o treinamento e 10% para a validação. Essa redução ocorreu principalmente devido a dados que correspondem às paradas do equipamento. Além disso, foi estipulado para cada variável um intervalo de valores que seriam utilizados nas regressões, para excluir valores das variáveis que estivessem muito fora da faixa operacional normal do equipamento. A definição destes pontos de corte foi realizada com base nos histogramas de cada variável. A Tabela 1 descreve quais são as variáveis operacionais utilizadas e suas respectivas faixas operacionais (pontos de corte dos dados empregados no desenvolvimento dos modelos).

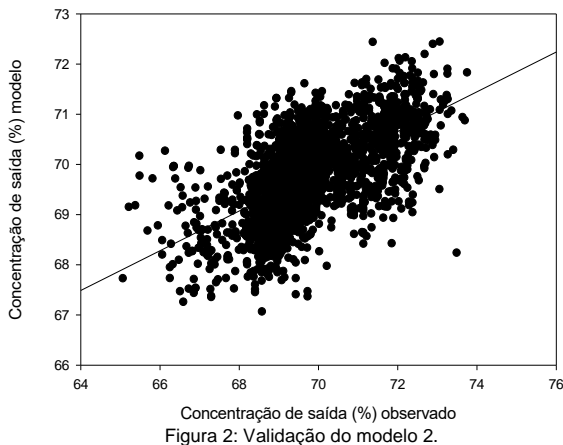
Tabela 1: Variáveis e suas respectivas faixas operacionais.

Variáveis	Valor mínimo	Valor máximo
Concentração na saída (%)	65,0	74,0
Vazão de vapor vivo (ton/h)	9,0	21,0
Vazão de licor na saída (m³/h)	40,0	160,0
Nível do evaporador (%)	35,0	60,0
Concentração na entrada (%)	58,0	71,0
Temperatura do licor na saída (°C)	104,0	116,0
Pressão do vapor vivo (Kgf/cm²)	-	0,0
Temperatura do vapor vivo (°C)	-	165,0
Pressão no vaso flash (Kgf/cm²)	3,0	-

O primeiro modelo testado utilizou como variáveis preditoras apenas as variáveis operacionais, sendo que o valor do coeficiente de determinação (R^2) na validação foi de 0,2980.

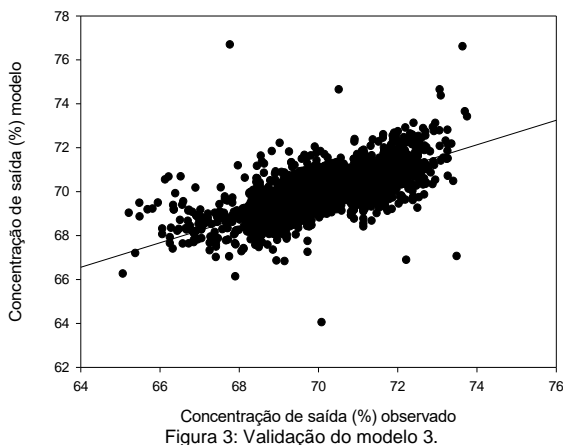


Verifica-se através da Figura 1 e do valor de R^2 que o modelo 1 não foi um resultado satisfatório. Com isso, foi desenvolvido um novo modelo acrescido de novas variáveis de entrada. Essas novas variáveis eram o quadrado e inverso das variáveis operacionais. O valor do R^2 para a validação desse segundo modelo foi de 0,3782.

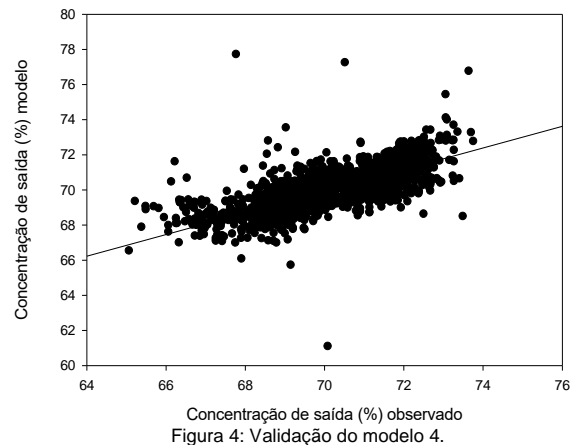


Novamente verifica-se através da Figura 2 e do valor de R^2 que o resultado obtido não foi satisfatório, apesar disso, esse segundo modelo descreve melhor o sistema que o primeiro modelo.

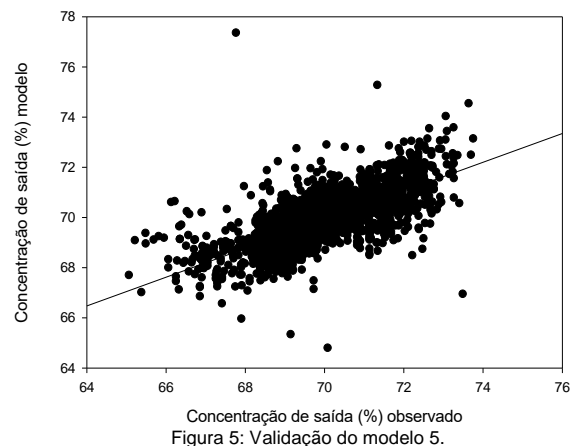
O terceiro e quarto modelos testados introduziram novas variáveis além das já adicionadas no segundo modelo. No modelo 3 foram adicionadas os produtos dois a dois entre as variáveis operacionais e seus respectivos quadrados e inversos. O valor do coeficiente de determinação na validação nesse modelo foi de 0,4881.



No modelo 4, foram adicionadas novamente os produtos dois a dois entre as variáveis operacionais, os quadrados, os inversos e o seno e cosseno das variáveis operacionais. O valor de R^2 obtido na validação foi de 0,5106.



No modelo 5, as variáveis operacionais foram normalizadas entre 0 e 1 e similar ao quarto modelo, foram introduzidas como variáveis preditoras o produto dois a dois das variáveis operacionais normalizadas, os quadrados, os inversos e o seno e cosseno das variáveis operacionais normalizadas. Espera-se que com as variáveis normalizadas os problemas provenientes da dispersão dos dados sejam minimizados. Esse modelo obteve o valor de R^2 na validação de 0,4939.



Para avaliar qual é o melhor modelo, além de fazer uma análise do R^2 e dos gráficos de validação, também deve ser feito a análise de resíduos.

Tabela 2: Média dos erros absolutos dos modelos testados.

Modelo	Média dos erros absolutos no treinamento	Média dos erros absolutos na validação
1	0,8167	0,8072
2	0,7715	0,7606
3	0,7616	0,6538
4	0,6085	0,6191
5	0,6383	0,6481

Tabela 3: Média dos erros ao quadrado dos modelos testados.

Modelo	Média dos erros ao quadrado no treinamento	Média dos erros ao quadrado na validação
1	1,1578	1,1481
2	1,0316	1,0169
3	1,1836	0,8372
4	0,7058	0,8004
5	0,7704	0,8277

Verifica-se a partir das Tabelas 2 e 3 que o modelo 4 apresenta valores de erros médios menores do que os dos outros modelos e além disso, é o modelo que apresenta maior valor de R^2 . Portanto, pode-se afirmar que o modelo 4 descreve melhor o comportamento do sistema.

4. Conclusões

A partir dos resultados apresentados neste trabalho, pode-se concluir que, no processo de modelagem por regressão linear múltipla dos evaporadores de múltiplo efeito, a inserção de novas variáveis de regressão, obtidas a partir das variáveis originais do problema, resulta em modelos de melhor desempenho.

O valor coeficiente de determinação do melhor modelo obtido neste trabalho indica que um pouco mais de 51% da variabilidade dos dados de validação foi explicada. Este resultado é surpreendente porque no

desenvolvimento do modelo não foram empregadas entradas atrasadas e o evaporador é um equipamento que não opera em regime estacionário. Mais além, o melhor modelo empírico da literatura (baseado redes neurais) para este mesmo sistema sem entradas explicou 40% da variabilidade dos dados de validação [5].

5. Agradecimentos

Os autores agradecem a FAPES pela bolsa pesquisador capixaba e ao CNPq pela bolsa de produtividade DTI.

6. Referências

- [1] CAMPOS, S. G. S. Modelagem matemática e Análise do coeficiente global de transferência de calor para o processo de concentração do licor negro de eucalipto em sistema de evaporadores de múltiplo efeito. Dissertação (Mestrado em Engenharia Industrial), Programa de Pós-Graduação em Engenharia, Centro Universitário do Leste de Minas Gerais, Coronel Fabriciano, 2009.
- [2] ALMEIDA, A. M. Modelagem, Simulação e Controle de um Sistema de Evaporação Múltiplo Efeito de Licor Negro. Monografia (Trabalho de Conclusão de Curso) Faculdade de Telêmaco Borba, Paraná, 2009.
- [3] MELO, A. L. V., Figueiredo, L. S., COSTA, A. O. S., COSTA Jr., E. F., 2011, Modelagem por Regressão Linear Múltipla do Sistema de Caustificação de uma Fábrica de Celulose. 2011. In: Anais do IX Congresso Brasileiro de Engenharia Química – Iniciação Científica – COBEQIC 2011, v. 1, p. 1-6.
- [4] HAIR Jr., J. F., ANDERSON, R. E., TATHAM, R. L., BLACK, W. C., 2005, Análise Multivariada de Dados; trad. Adonai Schlup Sant'Anna e Anselmo Chaves Neto. – 5ed. – Porto Alegre: Bookman.
- [5] ARAUJO, M. N.; PINHEIRO, O. S.; COSTA, A. O. S.; COSTA Jr, E. F.; "Descrição de um Evaporador de Licor Preto de Fábrica de Celulose por Redes Neurais sem Entradas Atrasadas", p. 57-58 . In: Anais do V Encontro Científico de Física Aplicada [= Blucher Physics Proceedings, n.1, v.1]. São Paulo: Blucher, 2014.