

# MODELAGEM DE UM PROCESSO PRODUTIVO UTILIZANDO REDES NEURAIS ARTIFICIAIS

**Simone Massulini Acosta (UTFPR/PUCPR)**

smacosta@utfpr.edu.br

**Anderson Levati Amoroso (UTFPR)**

andersonl@utfpr.edu.br

**Ângelo Márcio Oliveira Santanna (UFBA)**

angelo.santanna@ufba.br



*Para melhorar a qualidade de seus produtos ou serviços as empresas buscam reduzir a variabilidade nos seus processos e produtos para que estes atendam as especificações preestabelecidas. Em um processo industrial muitas vezes não se consegue controlar todas as causas ou fatores de variação que produzem determinado efeito sobre uma ou mais características da qualidade dos produtos. Como resultado pode-se ter a produção de produtos não conformes às especificações, que podem ser mensuradas através da fração de produtos não conformes. A modelagem da fração de produtos não conformes pode ser realizada utilizando-se diferentes técnicas e modelos de regressão. As redes neurais artificiais são muito utilizadas para a modelagem de dados de processos. O objetivo deste artigo é modelar a fração de produtos não conformes às especificações de uma indústria curtidora de couro utilizando redes neurais artificiais e comparar os resultados obtidos com outros modelos de regressão desenvolvidos para o mesmo processo produtivo. O modelo de regressão utilizando redes neurais*

*artificiais apresentou melhor desempenho no ajuste do modelo aos dados do processo quando comparado ao modelo de regressão de Quase-verossimilhança e ao modelo de regressão Beta e desempenho próximo ao modelo utilizando regressão por vetores suporte.*

*Palavras-chave: Inteligência computacional, redes neurais artificiais, modelagem de dados.*

## 1. Introdução

Os processos produtivos estão sujeitos a variabilidade, que está relacionada com as diferenças existentes entre as unidades produzidas. Segundo Costa, Epprecht e Carpinetti (2010), se a variabilidade do processo for grande, as diferenças entre as unidades produzidas serão fáceis de observar mas, se a variabilidade for pequena, tais diferenças serão difíceis de observar.

O conceito de qualidade tem sido definido de diferentes maneiras na literatura. Para Montgomery (2004) a qualidade é inversamente proporcional à variabilidade e a melhoria da qualidade é a redução da variabilidade nos processos e produtos. Taguchi define que um produto de qualidade é aquele que atende perfeitamente às especificações atingindo o valor alvo com a menor variabilidade possível em torno dele. Os elementos que, em conjunto, descrevem a adequação de um produto ao uso podem ser denominados características da qualidade ou indicadores de desempenho (RIBEIRO; CATEN, 2012).

Em um processo industrial muitas vezes não se consegue controlar todas as causas ou fatores de variação que produzem determinado efeito sobre as características da qualidade dos produtos, pois certas causas são inerentes ao processo. As causas de variação que podem ser controladas podem resultar na produção de produtos com características da qualidade não conformes às especificações preestabelecidas, que podem ser mensuradas através da fração de produtos não conformes (MONTGOMERY, 2004).

A fração não conforme é definida como a razão entre o número de unidades não conformes da amostra e o tamanho da amostra e podem compreender a razão entre dois números discretos, denominada de percentual, ou a razão entre dois números contínuos, denominada de proporção. As características da qualidade do tipo fração possuem as observações no intervalo  $[0,1]$  e geralmente não possuem distribuição Normal (SANT'ANNA; CATEN, 2010).

A análise de regressão consiste de uma técnica estatística para modelar e investigar a relação entre as características da qualidade com os fatores controláveis que podem afetá-las (CHARNET et al., 2008). A modelagem da fração de produtos não conformes é uma

relevante estratégia para investigar a relação entre as causas de variação que interferem no processo e as características da qualidade dos produtos.

Segundo Sant'Anna e Caten (2010) os modelos estatísticos de regressão linear não são adequados quando se monitoram características da qualidade do tipo fração de produtos não conformes, pois esses pressupõem que a característica da qualidade possui distribuição Normal e possibilita a previsão de valores fora do limite do intervalo  $[0,1]$ . Os modelos estatísticos pertencentes à classe dos Modelos Lineares Generalizados e o modelo de regressão Beta são alternativas para a modelagem de variáveis com distribuição não Normal.

Um método muito utilizado para a modelagem de processos é baseado em Rede Neural Artificial (RNA). As RNAs são modelos de processamento de dados que emulam redes de neurônios biológicos, capazes de recuperar rapidamente uma grande quantidade de dados e reconhecer atributos baseados na experiência (HAYKIN, 2009).

As RNAs utilizam o princípio de minimização do risco empírico (*Empirical Risk Minimization*, ERM) baseado no princípio da indução para inferir uma função que minimize o erro sobre os dados de treinamento conhecidos e, desta forma, espera-se que esse procedimento leve também a um menor erro sobre os dados de teste (VAPNIK, 1998).

A aplicação de RNAs para modelar diferentes processos químicos e industriais tem sido apresentado na literatura nos últimos anos. Bakar et al. (2013) realizaram um estudo para identificar os parâmetros principais de um trocador de calor em uma planta de produção de bio metanol. Osuolale e Zhang (2014) avaliaram diferentes modelos de RNAs para a previsão da carga em um mercado de eletricidade. Yong et al. (2015) utilizaram RNAs para minimizar o projeto de um permutador de calor, permitindo menor tempo e custo de fabricação. Acosta, Sant'Anna e Canciglieri (2016) analisaram a concentração de fósforo nas ligas de aço produzidas por uma empresa siderúrgica, pois altos níveis de fósforo causam danos as características físicas das ligas de aço.

O objetivo deste artigo é modelar a fração de produtos não conformes às especificações de uma indústria curtidora de couro utilizando Redes Neurais Artificiais e comparar os

resultados obtidos com outros modelos de regressão desenvolvidos para o mesmo processo produtivo.

## 2. Redes Neurais Artificiais

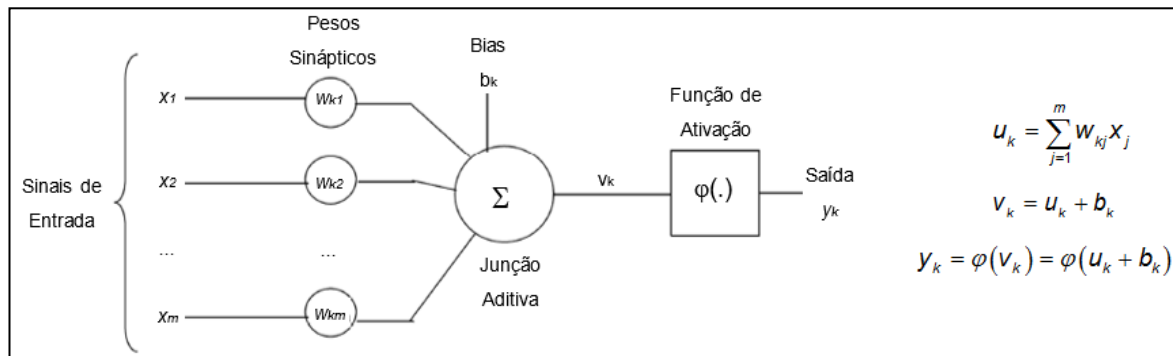
Segundo Braga, Carvalho e Ludermir (2012), as RNAs são sistemas paralelos distribuídos compostos por unidades de processamento simples (neurônios artificiais), dispostas em uma ou mais camadas e interligadas por um grande número de conexões, que calculam determinadas funções matemáticas (normalmente não lineares). Na maioria dos modelos as conexões estão associadas a pesos, que armazenam o conhecimento adquirido pelo modelo e servem para ponderar a entrada recebida por cada neurônio da rede.

O procedimento usual na solução de problemas utilizando RNAs inicia por uma fase de aprendizagem, em que um conjunto de exemplos é apresentado para a rede, que extrai as características necessárias para representar a informação fornecida. Essas características são utilizadas para gerar respostas para o problema (BRAGA; CARVALHO; LUDERMIR, 2012).

De acordo com De Castro (2006), as RNAs podem ser caracterizadas por três aspectos principais: (1) um grupo de neurônios artificiais, denominados nós, unidades ou neurônios; (2) um padrão de conectividade entre os neurônios, denominado de arquitetura ou estrutura; e (3) um método para determinar os valores dos pesos, denominado de algoritmo de treinamento ou de aprendizado.

Cada neurônio em uma rede neural recebe uma série de entradas. Uma função de ativação é aplicada a esses valores de entrada, o que resulta no nível de ativação do neurônio, que é o valor de saída do neurônio. A Figura 1 apresenta um modelo não linear de um neurônio artificial e sua representação matemática (HAYKIN, 2009).

Figura 1 - Modelo não linear de um neurônio artificial e sua representação matemática



Fonte: Adaptado de Haykin (2009)

Existem várias funções de ativação que podem ser utilizadas para o projeto de RNAs. As funções de base radial (por exemplo, gaussiana) e sigmoide (por exemplo, logística e tangente hiperbólica) são totalmente diferenciáveis e facilitam o processo de aprendizado da rede. A função sigmoide é a mais utilizada em RNAs (HAYKIN, 2009).

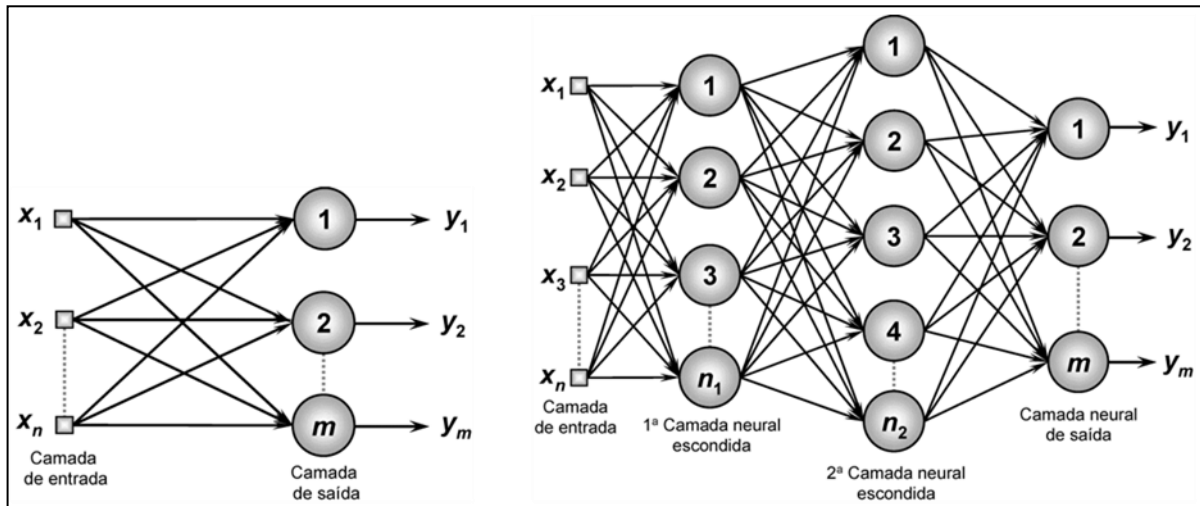
A arquitetura de uma RNA define a forma como seus neurônios estão arranados, uns em relação aos outros, através do direcionamento das conexões sinápticas destes neurônios. A topologia de uma RNA pode ser definida como sendo as diferentes formas de composições estruturais que esta poderá assumir (SILVA; SPATTI, FLAUZINO, 2010).

De acordo com De Castro (2006), em geral é possível distinguir três tipos principais de estruturas de RNAs: redes *feedforward* de camada única, redes *feedforward* de camadas múltiplas e redes recorrentes. Nas redes *feedforward* não existe conexão entre os neurônios da mesma camada e não existe realimentação entre as camadas. A Figura 2 apresenta um exemplo de uma rede *feedforward* de camada única e de uma rede *feedforward* de camadas múltiplas (SILVA; SPATTI; FLAUZINO, 2010).

Figura 2 - Exemplos de rede *feedforward*: (a) camada única, (b) camadas múltiplas

(a)

(b)



Fonte: Silva, Spatti e Flauzino (2010)

Para resolver problemas não linearmente separáveis utilizando RNAs, a alternativa mais utilizada são as redes *feedforward* de camadas múltiplas. As camadas escondidas utilizam funções de ativação não lineares. Entre os principais tipos de redes com esta arquitetura estão o Perceptron Multicamadas (*Multilayer Perceptron*, MLP) e as Redes de Base Radial (*Radial Basis Function*, RBF) (FACELI et al., 2011).

Segundo Du e Swamy (2013), o teorema de Kolmogorov garante que uma rede *feedforward* com uma camada escondida, com um número suficiente de neurônios, pode aproximar qualquer função contínua com qualquer grau de precisão. Baseado no teorema de Kolmogorov, o teorema de Hecht-Nielsen apresenta que qualquer função contínua pode ser aproximada por uma rede *feedforward* com  $n$  nós de entrada,  $2n+1$  neurônios escondidos e  $m$  neurônios de saída. A utilização de duas camadas escondidas permite a aproximação de qualquer função (SCARCELLI; TSOI, 1998; FACELI et al., 2011).

Segundo Haykin (2009), a propriedade de importância primordial para uma RNA é a sua habilidade de aprender a partir de seu ambiente e de melhorar o seu desempenho através da aprendizagem. Uma RNA aprende através de um processo iterativo de ajustes dos pesos sinápticos e níveis de bias, tornando-se mais instruída sobre seu ambiente após cada iteração do processo de aprendizagem. Um conjunto preestabelecido de regras para a solução de um

problema de aprendizagem é denominado de algoritmo de aprendizagem, tendo-se como objetivo final a generalização de soluções a serem produzidas pelas saídas.

Os algoritmos de aprendizagem diferem entre si na maneira de obter o ajuste do peso sináptico de um neurônio, podendo ser agrupados em aprendizado supervisionado, não supervisionado e por reforço (BRAGA; CARVALHO; LUDERMIR, 2012).

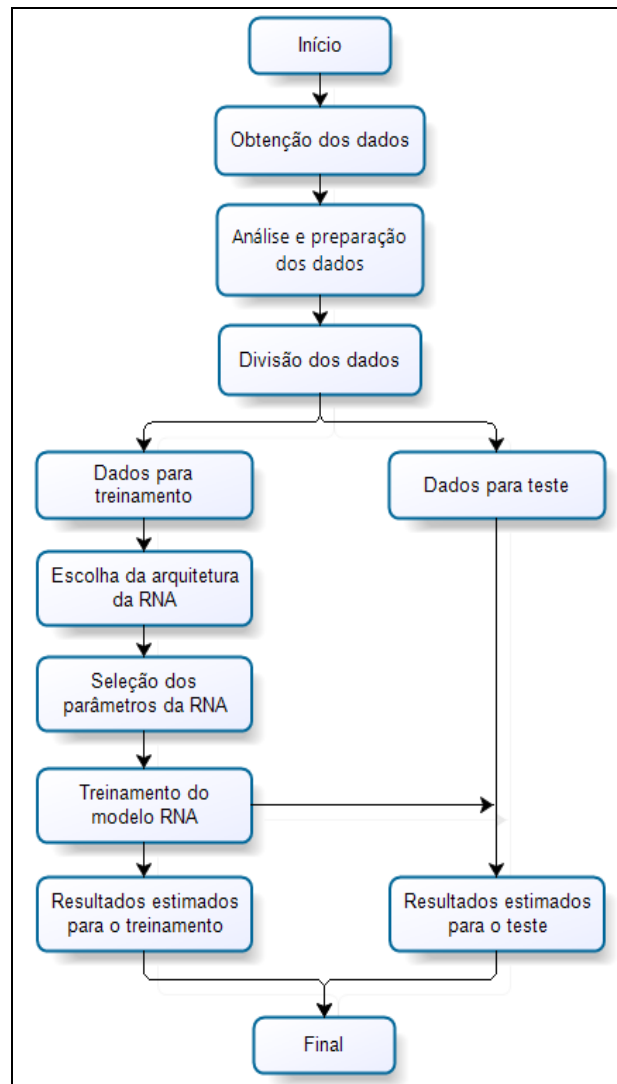
No treinamento de redes *feedforward* de camada única por meio do aprendizado supervisionado e correção de erros, o erro é obtido diretamente através da diferença entre a saída desejada e a saída atual da rede. No entanto, nas redes *feedforward* de camadas múltiplas esse procedimento pode ser aplicado somente para a camada de saída, pois não existem saídas desejadas definidas para as camadas escondidas (DE CASTRO, 2006).

Conforme Braga, Carvalho e Ludermir (2012) para as redes MLP, que se caracterizam por utilizarem funções de ativação sigmoidais nas camadas escondidas e sigmoidais ou lineares na saída, é utilizado o algoritmo de retropropagação de erros ou *error backpropagation*. Este algoritmo utiliza o gradiente descendente e estima o erro das camadas escondidas por meio de uma estimativa do efeito que estas causam no erro da camada de saída. Existem várias modificações do algoritmo *backpropagation* que visam acelerar seu tempo de treinamento e melhorar seu desempenho, sendo uma destas modificações a Levenberg-Marquardt.

Normalmente, o procedimento para o desenvolvimento do modelo RNA segue os seguintes passos: obtenção dos dados, análise e preparação dos dados, divisão dos dados, escolha da arquitetura da rede, seleção dos melhores parâmetros para a rede, simulações de treinamento e de teste e análise do desempenho. A Figura 3 apresenta um fluxograma geral do procedimento para a obtenção do modelo RNA.

Figura 3 - Fluxograma para o desenvolvimento do modelo RNA





Fonte: Autoria própria

Para análise do desempenho e avaliação da capacidade de generalização do modelo desenvolvido com RNAs podem ser utilizadas estratégias de minimização do erro de predição ou erro residual. O erro residual ( $e_i$ ) é a diferença entre o valor real ( $y_i$ ) e o valor estimado pelo modelo ( $\hat{y}_i$ ), conforme apresentado na equação (1). A equação (2) apresenta o erro médio absoluto (*Mean Absolute Error*, MAE), a equação (3) apresenta o erro médio quadrático (*Mean Squared Error*, MSE) e a equação (4) apresenta a raiz do erro médio quadrático (*Root Mean Squared Error*, RMSE), onde  $n$  representa o número de amostras.

$$e_t = y_i - y_t \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_t| = \frac{1}{n} \sum_{i=1}^n |e_t| \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_t)^2 = \frac{1}{n} \sum_{i=1}^n (e_t)^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_t)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_t)^2} \quad (4)$$

Outro critério que pode ser utilizado para verificar a adequação do modelo é o Coeficiente de Informação de Akaike corrigido (*corrected Akaike Information Criterion*, AICc), equação (5) (SANT'ANNA; CATEN, 2010).

$$AICc = \ln \left( \frac{\sum_{t=1}^n e_t^2}{n} \right) + \frac{2n(W+1)}{n-W-2} \quad (5)$$

onde  $n$  representa o número de amostras e  $W$  representa o número de parâmetros do modelo. Para as RNAs o número de parâmetros do modelo corresponde ao número de pesos e de *bias* a serem ajustados, sendo  $W = k(j+1) + i(j+1)$ , onde  $k$  é o número de entradas,  $j$  é o número de neurônios na camada oculta e  $i$  é o número de neurônios na camada de saída (YU; WANG; LAI, 2006).

### 3. Estudo aplicado

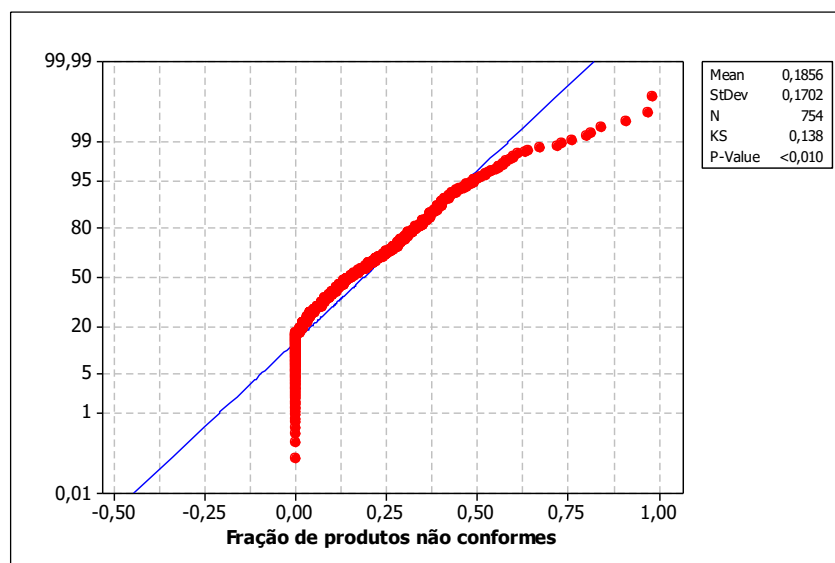
Sant'Anna e Caten (2010) apresentam a modelagem dos dados de uma empresa curtidora de couro, produtora de couro acabado e fornecedora para as indústrias de calçados e artefatos em couro, utilizando o modelo de regressão de Quase-verossimilhança, pertencente à classe dos Modelos Lineares Generalizados, e o modelo de regressão Beta.

A etapa *wet blue* do processo produtivo avaliado consiste em: o classificador recebe um lote de diferentes tamanhos contendo as matérias-primas e verifica se as características de qualidade satisfazem às especificações, por métodos cognitivos. As matérias-primas que não

satisfazem às especificações são classificadas como produtos não conformes e a fração de produtos não conformes às especificações, por lote, é considerada a variável dependente. Os dados coletados contemplaram uma amostra de 754 lotes.

A fração de produtos não conformes do processo não segue a distribuição Normal. A Figura 4 apresenta o teste de normalidade dos dados do processo baseado no método Kolmogorov-Smirnov com 95% de confiança ( $\alpha = 0,05$ ), sendo o valor obtido para *p-value* menor que 0,01. A Figura 5 apresenta o histograma da fração de produtos não conformes, onde se observa que o comportamento dos dados não segue a distribuição Normal e segue a distribuição Beta.

Figura 4 - Teste de normalidade da fração de produtos não conformes

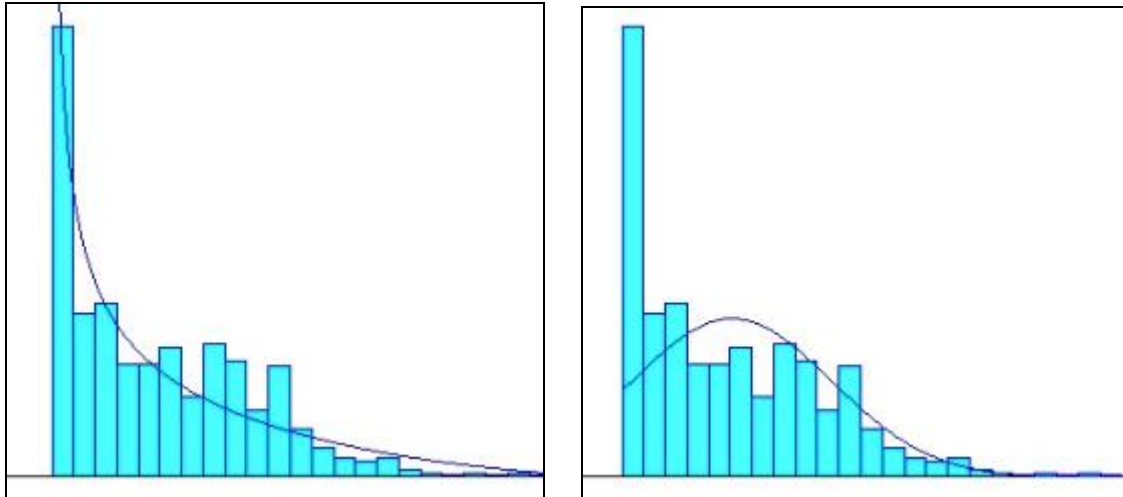


Fonte: Autoria própria

Figura 5 - Distribuição da fração de produtos não conformes com distribuição sobreposta:

(a) Beta

(b) Normal



Fonte: Autoria própria

Os fatores controláveis definidos como variáveis independentes para a modelagem da fração de produtos não conformes foram: a seleção da matéria prima conforme qualidade e preço ( $x_1$ ), com cinco níveis diferentes; a procedência da matéria-prima adquirida pela empresa ( $x_2$ ), com cinco níveis; o classificador que inspeciona as matérias-primas ( $x_3$ ), com três níveis, e o estado de rebaixamento da matéria-prima ( $x_4$ ), com dois níveis.

Para a construção dos modelos de regressão as variáveis independentes qualitativas seleção, procedência, classificador e rebaixamento foram substituídas pelas variáveis *dummy*. As novas variáveis independentes foram definidas como: seleção tipo 2 ( $x_1$ ), seleção tipo 3 ( $x_2$ ), seleção tipo 4 ( $x_3$ ), seleção tipo 5 ( $x_4$ ), procedência 2 ( $x_5$ ), procedência 3 ( $x_6$ ), procedência 4 ( $x_7$ ), procedência 5 ( $x_8$ ), classificador 2 ( $x_9$ ), classificador 3 ( $x_{10}$ ) e rebaixamento ( $x_{11}$ ).

Para a definição das variáveis independentes relevantes para a construção do modelo de regressão foi realizado o teste de correlação de Pearson. Após o teste foram selecionadas as variáveis seleção (tipo 2, tipo 3, tipo 4 e tipo 5), classificador (2 e 3) e rebaixamento como estatisticamente significativas para explicar a variável dependente fração de produtos não conformes.

Para o desenvolvimento dos modelos baseados em RNAs foi utilizada a biblioteca para redes neurais do programa Matlab®. Após a análise e preparação dos dados o conjunto total de

dados foi dividido randomicamente em dois subconjuntos: 70% para o grupo de treinamento e 30% para o grupo de teste.

A arquitetura selecionada para a RNA foi a MLP com sete variáveis de entrada, uma camada oculta e uma variável de saída. Após diversas simulações os parâmetros selecionados para as RNAs foram: algoritmo de treinamento Levenberg-Marquardt, função de ativação logística para a camada oculta e linear para a camada de saída e taxa de aprendizado de 0,01.

Para a obtenção do número de neurônios da camada oculta foi utilizado o teorema de Hecht-Nielsen e foram analisados MSE, RMSE, MAE e AICc para redes com diferentes números de neurônios na camada oculta. A Tabela 1 apresenta os resultados de algumas simulações com os dados de treinamento para a obtenção do número de neurônios da camada oculta. Em função dos resultados obtidos para os erros e para AICc foi selecionada a RNA com 19 neurônios na camada oculta.

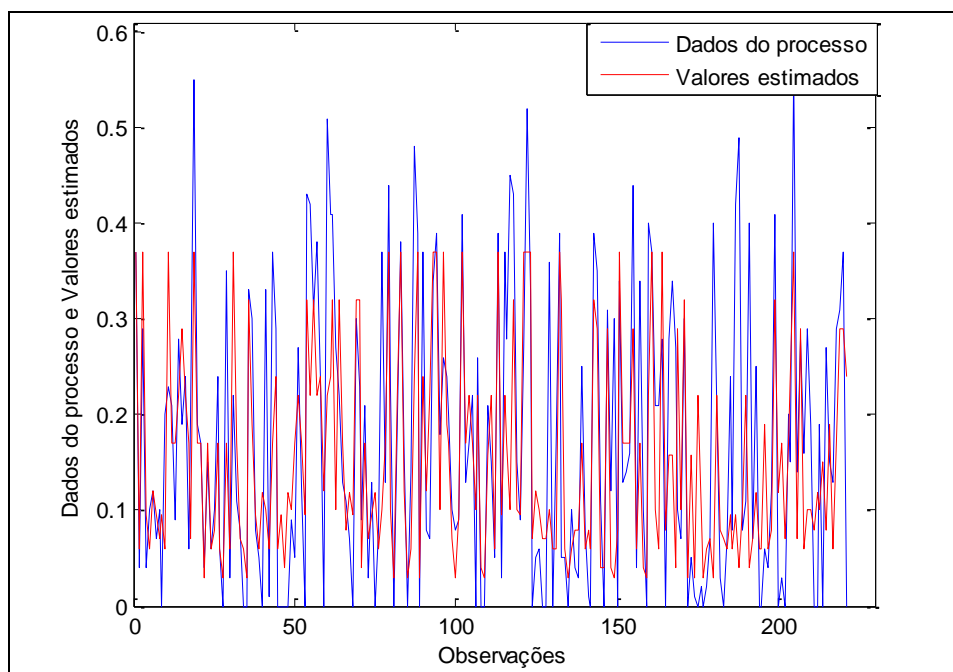
Tabela 1 - Resultados das simulações para obtenção do número de neurônios da camada oculta

Neurônios	14	15	16	17	18	19	20	21
MSE	0,01407	0,01285	0,01288	0,01281	0,01852	0,01262	0,01269	0,01271
RMSE	0,1186	0,1134	0,11349	0,1132	0,1361	0,11235	0,11265	0,11275
MAE	0,09502	0,08939	0,08934	0,08763	0,1138	0,08850	0,08727	0,08878
AICc	339,10	372,46	407,66	444,705	484,28	524,935	568,49	614,61

Fonte: Autoria própria

Após a seleção da arquitetura e dos parâmetros, a RNA foi treinada e, após a fase de treinamento, o modelo obtido foi utilizado para estimar as saídas utilizando os dados do grupo de teste. A Figura 6 apresenta que os valores estimados pelo modelo RNA se aproximam dos valores dos dados experimentais do processo do grupo de teste.

Figura 6 - Valores dos dados do processo do grupo de teste e valores estimados pelo modelo RNA



Fonte: Autoria própria

A Tabela 2 apresenta os valores dos erros calculados para os dados do grupo de treinamento e do grupo de teste. Os valores dos erros obtidos pelo modelo RNA para os grupos de treinamento e de teste não diferem significativamente, indicando não haver sobreajuste do modelo. Analisando a Figura 6 e a Tabela 2 pode-se considerar que o modelo RNA representa adequadamente a fração de produtos não conforme do processo.

Tabela 2 - Valores dos erros calculados para os grupos de treinamento e de teste

Erro	Grupo de treinamento	Grupo de teste
MSE	0,0126	0,0138
RMSE	0,1123	0,1176
MAE	0,0885	0,0937

Fonte: Autoria própria

Na sequência foi realizada a comparação do modelo RNA com o modelo de regressão de Quase-verossimilhança e modelo de regressão Beta apresentados em Sant'Anna e Caten

(2010) e com o modelo utilizando regressão por vetores suporte apresentado em Acosta, Amoroso e Sant'Anna (2016).

A forma de regressão para o modelo de Quase-verossimilhança é descrita pela equação (6) e para o modelo Beta é descrita pela equação (7) (SANT'ANNA; CATEN, 2010).

$$g(\hat{\mu}) = -3,7692 + 2,0475 * x_1 + 2,2857 * x_2 + 2,5226 * x_3 + 2,7583 * x_4 + 0,2672 * x_9 + 0,4483 * x_{10} - 0,8357 * x_{11} \quad (6)$$

$$g(\hat{\mu}) = -2,4225 + 1,5822 * x_1 + 1,1202 * x_2 + 1,2295 * x_3 + 1,4469 * x_4 + 0,2883 * x_9 + 0,3740 * x_{10} - 0,6807 * x_{11} \quad (7)$$

A aplicação de máquinas de vetores suporte (*Support Vector Machine*, SVM) foi originalmente desenvolvida para problemas de classificação de padrões, sendo posteriormente estendida para o tratamento de problemas de regressão e denominada regressão por vetores suporte (*Support Vector Regression*, SVR). O modelo utilizando SVR possui o parâmetro C igual a 10, o parâmetro  $\epsilon$  igual a 0,03 e a função *kernel* RBF com parâmetro  $\gamma$  igual a 0,2 (ACOSTA; AMOROSO; SANT'ANNA, 2016).

Os dados do grupo de teste utilizados no modelo RNA foram utilizados com o modelo Quase-verossimilhança, modelo Beta e modelo SVR. A Tabela 3 apresenta os valores dos erros calculados para os quatro modelos. Analisando os valores dos erros verifica-se que o modelo RNA possui um desempenho melhor para a modelagem da fração de produtos não conformes às especificações do que o modelo Quase-verossimilhança e modelo Beta. O modelo RNA possui um desempenho um pouco menor, porém próximo, ao do modelo SVR.

Tabela 3 - Valores dos erros calculados para o grupo de teste nos modelos

Modelo	MSE	RMSE	MAE
Quase-verossimilhança	0,0161	0,1268	0,0998
Beta	0,0151	0,1231	0,0961
SVR	0,0131	0,1144	0,0853
RNA	0,0138	0,1176	0,0937

Fonte: Autoria própria

As máquinas de vetores suporte utilizam o princípio de minimização do risco estrutural (*Structural Risk Minimization*, SRM) que envolve a minimização de um limite superior sobre o erro de generalização. Vapnik, Golowich e Smola (1996) mostraram que o princípio SRM é superior ao princípio ERM, utilizado em redes neurais artificiais. Desta forma, modelos de aprendizado de máquina baseados no princípio SRM tendem a apresentar uma maior habilidade para generalizar bem frente a dados de teste não observados (VAPNIK, 1998).

#### 4. Conclusões

Os efeitos das variáveis independentes (fatores controláveis) sobre a variável dependente (característica da qualidade) em processos industriais podem ser analisados através da modelagem dos dados destes processos usando modelos de regressão. Para esta modelagem podem ser utilizadas técnicas estatísticas, como os modelos lineares generalizados e o modelo de regressão Beta, e técnicas de inteligência computacional, como as redes neurais artificiais e a regressão por vetores suporte.

Neste artigo foram utilizadas as redes neurais artificiais para a modelagem da fração de produtos não conformes às especificações de uma indústria curtidora de couro em que a característica da qualidade é mensurada no intervalo [0,1]. O modelo obtido foi comparado com outros modelos de regressão desenvolvidos para o mesmo processo produtivo.

Analisando os valores dos erros obtidos para os modelos verifica-se que o modelo RNA apresentou melhor desempenho no ajuste do modelo aos dados do que o modelo Quase-verossimilhança e o modelo Beta e um desempenho próximo ao do modelo SVR.



Através dos resultados verifica-se que os modelos obtidos utilizando as duas técnicas de inteligência computacional obtiveram um desempenho melhor do que os modelos utilizando as duas técnicas estatísticas. Estes resultados estão de acordo com trabalhos disponíveis na literatura onde os autores concluem que as redes neurais artificiais e a regressão por vetores suporte apresentam desempenho similar ou superior quando comparadas as técnicas estatísticas convencionais. Apesar destas técnicas de inteligência computacional apresentarem bons resultados no modelamento de processos, estas ainda são mais complexas que as técnicas estatísticas convencionais.

Para trabalhos futuros sugere-se a modelagem da fração de produtos não conformes às especificações de diferentes processos industriais utilizando redes neurais artificiais e regressão por vetores suporte e a comparação com modelos obtidos através de técnicas estatísticas convencionais.

## REFERÊNCIAS

ACOSTA, Simone M.; SANT'ANNA, Ângelo M. O.; CANGIOLIERI Osiris J. Forecasting modeling for energetic efficiency in an industrial process. **Chemical Engineering Transactions**, v. 52, p. 1081-1086, 2016.

ACOSTA, Simone M.; AMOROSO, Anderson L.; SANT'ANNA, Ângelo M. O. Modelagem de um processo produtivo utilizando regressão por vetores suporte. In: **Encontro Nacional de Engenharia de Produção**, 2016, João Pessoa. Encontro Nacional de Engenharia de Produção. Rio de Janeiro: ABEPRO, 2016.

BAKAR, Suraya H. A. et al. A. Flexible and operable heat exchanger networks. **Chemical Engineering Transactions**, v. 32, p. 1297-1302, 2013.

BRAGA, Antônio de P.; CARVALHO, André P. de L. F.; LUDERMIR, Teresa B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2012.

CHARNET, Reinaldo et al. **Análise de modelos de regressão linear com aplicações**. Campinas: Editora da Unicamp, 2008.

COSTA, Antônio F. B.; EPPRECHT, Eugênio K.; CARPINETTI, Luiz C. R. **Controle estatístico de qualidade**. 2. ed. São Paulo: Atlas, 2010.

DE CASTRO, Leandro N. **Fundamentals of natural computing**: basic concepts, algorithms, and applications. London: Chapman & Hall/CRC, 2006.

DU, Ke-Lin; SWAMY, M. N. S. **Neural networks and statistical learning**. Heidelberg: Springer, 2013.

FACELI, Katti et al. **Inteligência artificial**: uma abordagem de aprendizado de máquina. Rio de Janeiro: LTC, 2011.

HAYKIN, Simon. **Neural networks and learning machines**. 3rd ed. New York: Prentice Hall, 2009.

MONTGOMERY, Douglas C. **Introdução ao controle estatístico da qualidade**. Rio de Janeiro: LTC, 2004.

OSUOLALE, F. N., ZHANG, J. Energy efficient control and optimization of distillation column using artificial neural network. **Chemical Engineering Transactions**, n. 39, p. 37-42, 2014.

RIBEIRO, José L. D.; CATEN, Carla S. Controle estatístico do processo. **Série Monográfica Qualidade**. Porto Alegre: FEENG/UFRGS, 2012.

SANT'ANNA, Ângelo M. O.; CATEN, Carla S. Modelagem da fração de não-conformes em processos industriais. **Pesquisa Operacional**, v. 30, n. 1, p. 53-72, janeiro/abril, 2010.

SCARCELLI, Franco, TSOI, Ah C. Universal approximation using feed-forward neural networks: a survey of some existing methods and some new results. **Neural Networks**, v. 11, p. 15-37, 1998.

SILVA, Ivan N. da; SPATTI, Danilo H.; FLAUZINO, Rogério A. **Redes neurais artificiais**: para engenharia e ciências aplicadas. São Paulo: Artliber: 2010.

VAPNIK, Vladimir N.; GOLOWICH, Steven; SMOLA, Alex J. Support vector method for function approximation, regression estimation and signal processing. **Advances in Neural Information Processing Systems**, v. 9, p. 281-287, 1996.

VAPNIK, Vladimir N. **Statistical learning theory**. New York: John Wiley & Sons, 1998.

YONG, J. Y. et al. Heat exchanger network modification for waste heat utilization under varying feed conditions. **Chemical Engineering Transactions**, n. 43, p. 1279-1284, 2015.

YU, L., WANG, S., LAI, K.K. An integrated data preparation scheme for neural network data analysis. **IEE Transactions on Knowledge and Data Engineering**, v. 18, p. 217-230, 2006.