

USO DE TÉCNICAS DE RECONHECIMENTO DE PADRÃO EM UM PROCESSO DE PRODUÇÃO DE ALUMINA PARA AUXÍLIO NO CONTROLE AMBIENTAL

Ruy Gomes da Silva

Programa de Pós-graduação em Engenharia de Produção e Sistemas - PUCPR

E-mail: ruyrugs@gmail.com

Maria Teresinha Arns Steiner

Programa de Pós-graduação em Engenharia de Produção e Sistemas - PUCPR

E-mail: maria.steiner@pucpr.br

Anderson Roges Teixeira Góes

Programa de Pós Graduação em Métodos Numéricos - UFPR

E-mail: artgoes@ufpr.br

Resumo: O principal processo de extração de Alumina que está contida na bauxita foi desenvolvido em 1858 e aperfeiçoado posteriormente em 1888 por *Karl Bayer*, sendo conhecido atualmente como processo *Bayer*. Cada empresa deste ramo, mesmo que utilize o mesmo processo de extração, apresenta diferentes valores de concentração de soda cáustica no rejeito do processo. Estima-se um valor de 15g/l de soda cáustica para cada tonelada de rejeito gerado sendo que, atualmente, este material é acondicionado em bacias próprias expostas ao meio ambiente. Assim sendo, este trabalho tem por objetivo utilizar, inicialmente, técnicas de Reconhecimento de Padrões (RP), de forma comparativa, visando reconhecer no rejeito de um processo de produção de alumina, denominado lama vermelha, os diferentes teores de soda cáustica gerados pelas diferentes entradas. Realizado este RP, pretende-se construir, então, um modelo matemático capaz de minimizar o teor cáustico contido no rejeito, através da melhor combinação entre variáveis de entrada do processo aumentando, desta forma, a eficiência da lama vermelha. Com o seu reaproveitamento na indústria, como na fabricação de cerâmica, na indústria civil, como componente do concreto ter-se-á, como consequência, a minimização dos impactos ambientais gerados pela mesma. As técnicas de RP aqui abordadas (Regressão Múltipla e Redes Neurais Artificiais) consideram os registros históricos referentes às variáveis de entrada e de saída da etapa do processo que realiza a lavagem da lama vermelha. Tais registros dizem respeito à leitura das medições dos diversos instrumentos envolvidos; a análise dos resultados laboratoriais e do comportamento do filtro rotativo que realiza a lavagem da lama vermelha. Pretende-se utilizar a técnica de RP com melhor desempenho para a construção do referido modelo matemático que “investigará” como “trabalhar” com os dados de entrada de forma a obter como resultado o mínimo teor de soda cáustica. Alguns testes preliminares foram realizados com as técnicas de RP e a técnica de Redes Neurais Artificiais apresenta melhor aderência aos dados da problemática, mostrando um elevado percentual de acertos nos dados testados, sendo assim possível a formulação de um modelo que possa minimizar o teor de soda cáustica na lama vermelha. Viabilizando assim seu reaproveitamento na indústria civil, auxiliando na redução dos impactos ambientais.

Palavras-chave: Redes Neurais Artificiais. Regressão Múltipla. Previsão. Alumina.

1 INTRODUÇÃO

O processo mais utilizado industrialmente para a produção de alumina é denominado Processo Bayer, que consiste em misturar a bauxita moída a uma solução de soda cáustica, com a qual a mesma reage sob pressão e temperatura (LUZ e LINS, 2005). O Processo Bayer, desenvolvido pelo químico austríaco Karl Joseph Bayer, em 1889, é, atualmente, o único

processo economicamente viável para a produção de alumina (HIND *et al.*, 1999; SILVA FILHO *et al.*, 2007).

Dentro do complexo processo de extração da alumina destaca-se a “lavagem e filtração da lama vermelha”, que tem como principal finalidade, a retirada de soda cáustica contida no rejeito (lama vermelha) antes de seu descarte.

Os filtros rotativos que compõem o processo de lavagem da lama vermelha operam com a utilização de uma grande quantidade de variáveis, que combinadas promovem a retirada de soda cáustica desta lama que, em seguida, é transferida para uma bacia de contenção exposta no meio ambiente preparada para tal fim.

Para a lavagem da lama vermelha, uma série de variáveis é combinada visando o melhor desempenho possível da atividade.

A lama vermelha é descartada em bacias planejadas de tal forma que não tenham contato direto com o solo. Porém gera um grande desmatamento poderia ser amenizado, caso esta lama vermelha fosse reaproveitada em outro processo de fabricação, uma vez que tal rejeito não pode ser reaproveitado devido ao elevado volume de soda cáustica contida na lama.

O objetivo deste trabalho é utilizar, inicialmente, uma Rede Neural Artificial (RNA) para reconhecer no rejeito de um processo de produção de alumina, a Lama Vermelha, os diferentes teores de soda cáustica gerados pelas diferentes entradas.

Após o Reconhecimento de Padrão (RP), pretende-se construir, então, um modelo matemático capaz de minimizar o teor cáustico contido no rejeito, através da melhor combinação entre variáveis de entrada do processo minimizando, desta forma, o teor de soda cáustica contida na lama vermelha.

2 REFERENCIAL TEÓRICO

2.1 Redes neurais artificiais

Realizar uma otimização não é uma tarefa simples, uma vez que se pode deparar com variáveis qualitativas e quantitativas, correlacionadas, ou não. No entanto, existem inúmeras técnicas que podem nos auxiliar neste processo.

Para os processos com utilizam sistemas de controle complexos é possível aplicar técnicas de Reconhecimento de Padrão (RP) com a finalidade de identificar o comportamento, o padrão, das variáveis deste processo. Dentre as técnicas de RP que podem ser utilizadas têm-se as Redes Neurais Artificiais (RNA), que além de possuírem uma capacidade de modelar sistemas complexos lineares e não-lineares, possuem também mecanismos de aprendizagem, sendo estes os motivos que levam a utilização de RNA em controles de processos adaptativos que estão sujeitas a incertezas nos controles das variáveis e dos processos (CAMPOS & SAITO, 2004).

Uma RNA funciona baseada na estrutura do cérebro humano que é responsável pela execução das funções sensoras, motoras e autônomas do corpo, além das emoções, pensamento e percepção. O cérebro com suas redes de neurônios tem ainda a capacidade de reconhecer padrões, armazenar conhecimento através de experiências e interpretar observações (BRAGA *et al.*, 2000).

Haykin (2005) afirma que uma RNA pode ser vista como uma máquina adaptativa com capacidade de aprendizagem, graças à interligação maciça de células computacionais denominadas neurônios. Ela pode ser definida ainda como “um processador paralelamente distribuído constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso”.

Uma RNA é composta por várias unidades de processamento, cujo funcionamento é bastante simples. Essas unidades, geralmente são conectadas por canais de comunicação que estão associados a determinado peso. As unidades fazem operações apenas sobre seus dados locais, que são entradas recebidas pelas suas conexões. O comportamento inteligente de uma RNA vem das iterações entre as unidades de processamento da rede.

São muitos os algoritmos de aprendizagem de RNA: backpropagation; Levenberg-Marquardt (LM); dentre outros (HAGAN & MENHAJ, 1994; FAUSSET, 1994; HAYKIN, 2005). Arquiteturas neurais são tipicamente organizadas em camadas, com unidades que podem estar conectadas às unidades da camada posterior, como descrito na Figura 1 a seguir.

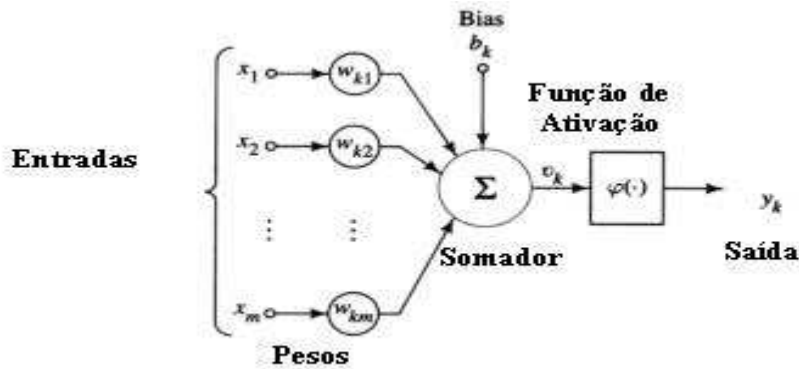


Figura 1 – Arquitetura típica de um neurônio artificial.
 Fonte: Haykin (2005).

Conforme a figura 1 anterior, o vetor X , que representa um conjunto de m entradas (X_1, X_2, \dots, X_m) , é multiplicados por um vetor peso, $W (w_{k1}, w_{k2}, \dots, w_{km})$, resultando em $P = X \cdot W$. Na sequência é calculada a soma de todos os elementos de P , sendo cada um deles multiplicado por b_k . Esta soma é processada por uma função de ativação, resultando no “sinal” de saída do neurônio. Este procedimento poderá ser repetido n vezes, denominado de iterações, que será definido pelo usuário durante a fase de treinamento da rede.

2.2 Regressão Múltipla

A técnica estatística de Regressão Múltipla é usada para estudar a relação entre uma variável dependente e várias variáveis independentes. Normalmente estas variáveis têm correlações entre si e o grau de importância das mesmas está relacionado com o impacto de cada uma no processo produtivo. Segundo Larson e Farber (2010): “uma correlação é uma relação entre duas variáveis. Os dados podem ser representados por pares ordenados (x, y) , onde x é a variável independente (ou explanatória) e y é a variável dependente (ou resposta)”.

O modelo genérico é dado pela equação 1 a seguir, segundo LEVINE *et al.* (2005), quando aplicado a uma amostra de tamanho n .

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

Onde

\hat{y} = Variável dependente ou explicada, também chamada de resposta;

β_0 = Interceptor ou termo independente de variável;

β_1 = inclinação de Y em relação à variável X_1 , mantendo constantes as variáveis X_2, X_3, \dots, X_K .

β_2 = inclinação de Y em relação à variável X_2 , mantendo constantes as variáveis X_1, X_3, \dots, X_K

β_k = inclinação de Y em relação à variável X_K , mantendo constantes as variáveis X_1, X_2, \dots, X_K

ε = erro aleatório em Y .

2.2.1 Poder de explicação do modelo

Uma característica importante para verificação do ajuste do modelo, que está intimamente ligado à estatística F , é o coeficiente de determinação múltipla que, informa o poder de explicação do modelo em função das p variáveis independentes consideradas. Este coeficiente é obtido dividindo-se a variação explicada pela variação total do modelo, ou seja, a razão entre a parcela SQ_{Regr} e SQ_T , conforme a equação 2 a seguir.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad 0 < R^2 < 1 \quad (2)$$

O R^2 é conhecido como coeficiente de correlação múltipla ao quadrado ou coeficiente de determinação. Quando o ajuste é bom o modelo explica boa parte da variação total e

consequentemente o valor de R^2 é próximo de 1. O coeficiente de determinação é uma medida da qualidade do ajuste.

2.2.3 Relação entre as variáveis

O grau de relação entre as variáveis, que expressa como as variáveis estão relacionadas entre si, é definido numericamente pelo Coeficiente de Correlação, representado por (r). Com base em n observações do par (X, Y) este parâmetro é estimado pela estatística apresentado na equação 3 a seguir.

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{(X_i - \bar{X})^2 (Y_i - \bar{Y})^2}} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y} \quad (3)$$

Onde:

\bar{X} é a média da variável independente X;

\bar{Y} é a média da variável dependente Y;

$\hat{\sigma}_{xy}$ é a covariância amostral entre X e Y;

$\hat{\sigma}_x$ é o desvio padrão amostral de X;

$\hat{\sigma}_y$ é o desvio padrão amostral de Y.

O coeficiente de correlação varia entre os limites -1 e 1 podendo, portanto, ser positivo ou negativo ($-1 \leq r \leq 1$). Quando o coeficiente de correlação é nulo ($r = 0$), significa que não existe nenhum relacionamento entre as variáveis. E quando o coeficiente de correlação é igual a unidade, -1 ou +1, tem-se um relacionamento perfeito entre elas. O sinal (+) ou (-) das variáveis indica a relação direta ou indireta existente entre as variáveis. O grau de relacionamento entre as variáveis, definido numericamente pelo valor r, pode ser assim interpretado:

Coeficiente	Correlação
$r = 0$	relação nula;
$0 < r \leq 0,30$	relação fraca;
$0,30 < r \leq 0,70$	relação média;
$0,70 < r \leq 0,90$	relação forte;
$0,90 < r \leq 0,99$	relação fortíssima;
$r = 1$	relação perfeita.

3 OBTENÇÃO DOS RESULTADOS

Uma série de registros (amostras) foi coletada pelo especialista de processo da área de produção da lama vermelha, que os identificou como sendo os atributos de entrada e a variável de saída. A tabela 1, a seguir, mostra alguns destes registros, onde a variável x1 representa a densidade, x2 a pressão de vácuo, x3 o nível da bacia, x4 a rotação do filtro, x5 o condensado de lavagem, x6 o teor total de soda, e como variável de saída a concentração de soda na lama (Y).

Amostras	Densidade (g/cm ³)	Vácuo (kgf)	Nível de Bacia (%)	Rotação (rpm)	Spray (m ³ /h)	TTS (g/l)	Concentração (g/l)
n	X1	X2	X3	X4	X5	X6	Y
1	1,49	0,51	25,55	1,60	24,81	66,57	5,85
2	1,52	0,51	25	1,60	25,23	70,43	6
3	1,51	0,52	19,64	1,60	22,25	68,4	14,65
.
.
.
298	1,43	0,45	18,66	2,20	23,71	64,1	8,49
299	1,43	0,44	19,96	2,10	24,02	58,83	8,16
300	1,52	0,42	18,16	2,10	23,76	59,99	8,24

Tabela 1 – Dados referente as amostragens
Fonte – Autores (2013)

Tais registros serviram de base para o treinamento e teste da RNA e foram coletadas diariamente (duas por dia), sempre em horários pré-estabelecido, perfazendo um total de 300 amostras.

Foi desenvolvida uma RNA de multicamadas, onde se têm os seis neurônios na camada de entrada; um número variável na camada oculta/intermediária e um neurônio na camada de saída. A tabela 2 a seguir mostra alguns testes realizados com os dados, de acordo com a variação na quantidade de neurônios na camada intermediária. Para a construção/treinamento da rede utilizou-se 70% do total de amostras para o treinamento e os demais 30% para os testes/validação. Foram realizadas 3000 épocas (condição de parada).

RNA			
Camada de Entrada	Camada Oculta	Camada de Saída	MSE
6	10	1	4
6	15	1	4,15
6	20	1	4,19
6	25	1	4,56
6	30	1	5,34

Tabela 2 – Configurações da RNA.

Fonte: Autor.

Para cada uma ds topologias (6-variável-1), o treinamento da rede ocorreu com 70% do total de registros, e os outros 30% dos registros foram utilizados para os testes; tal procedimento foi realizado 10 vezes, com os pesos iniciais aleatórios, variando no intervalo (-1, 1). As taxas de aprendizagem e de momento, ambas variando no intervalo de (0, 1), ficaram definidas em 0,8 no início do treinamento e foram sendo ajustadas de acordo com o aprendizado da rede.

A análise gráfica obtida pela resposta da rede treinada permite inferir que os valores previstos pela RNA em sua grande maioria conseguem retratar os valores reais do processo, conforme podemos visualizado na Figura 2, a seguir.

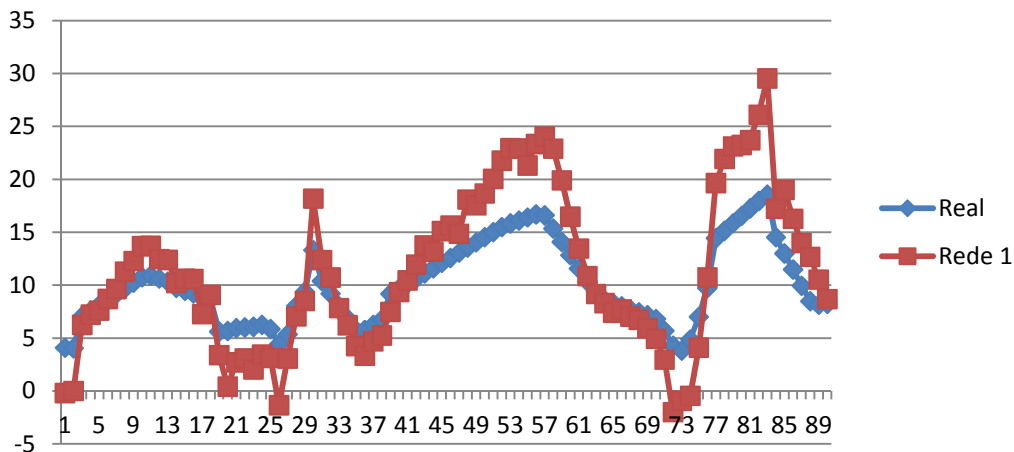


Figura 2 – Gráfico resultados da rede neural

Fonte: Autores (2013)

Após a simulação, a RNA usada conseguiu gerar valores aproximados aos dados reais com um bom desempenho. Comparando as saídas dos dados reais com os dados simulados obteve-se um erro médio quadrático calculado de 0,015192 dentro do intervalo de especificação do processo para os limites de segurança verificados previamente.

Com o mesmo banco de dado, foi aplicada a técnica de Regressão Múltipla, porém os resultados estatísticos não foram muito animadores, tanto o coeficiente de correlação entre as variáveis (R múltiplo), quanto o coeficiente de regressão (R-Quadadro), apresentaram valores baixos conforme apresentado na tabela 3 a seguir.

<i>Estatística de regressão</i>	
R múltiplo	0,381314572
R-Quadrado	0,145400802

Tabela 3 – Estatística da Regressão.
Fonte: Autores (2013)

Com o valor apresentado do R múltiplo em 0,381 pode-se inferir que as variáveis possuem uma correlação média, tal afirmação pode ser observar na tabela 4, onde se mostra correlação que cada variável tem uma com as outras e com a variável de saída.

	X1	X2	X3	X4	X5	X6	Y
X1	1						
X2	-0,11668	1					
X3	-0,09803	-0,04976	1				
X4	-0,14671	-0,07285	0,050942	1			
X5	0,143056	0,128523	0,070061	-0,28208	1		
X6	-0,05685	0,067694	0,029424	0,044162	-0,08152	1	
Y	0,062714	0,003894	0,039672	0,126995	-0,10102	0,111559	1

Tabela 4 – Correlação entre as variáveis
Fonte: Autores (2013)

Observa-se também que o valor do R-Quadrado sugere que o modelo gerado tem um baixo poder de explicação, ou seja, o resultado previsto pelo modelo terá baixa confiabilidade. O modelo de previsão gerado por este método apresenta-se a seguir na equação 4, e seus dados simulados logo em seguida na Figura 3.

$$\hat{y} = -1,21 + 13,26x_1 - 4,41x_2 - 0,09x_3 - 0,77x_4 - 0,33x_5 + 0,07x_6 \quad (4)$$

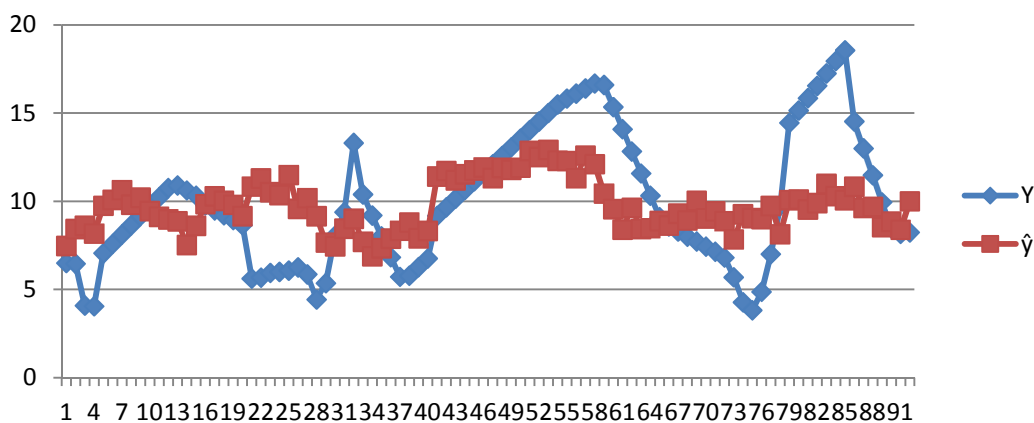


Figura 3 – Dados reais x dados simulados.
Fonte: Autores (2013)

Nota-se que os resultados de \hat{y} (valor previsto), não acompanha com precisão as oscilações de Y (valor real). Isso é explicado pelo baixo valor do R-múltiplo e R-Quadrado, mostrados na tabela 3 anterior.

4 CONSIDERAÇÕES FINAIS

Estes resultados, apesar de preliminares, demonstram que é possível a utilização das técnicas de RP que possa servir de base para criação de um modelo de previsão para os

resultados de concentração cáustica na lama vermelha. Neste pequeno experimento pode-se observar que a modelagem por RNA apresentou uma melhor aderência aos dados da referida problemática. A topologia de RNA utilizada foi uma rede de multicamadas, com aprendizado supervisionado. O algoritmo utilizado para o treinamento da rede foi o Backpropagation.

Para avaliação do treinamento da RNA foi utilizado o MSE, conforme recomenda a literatura. Foi realizada, também, uma avaliação através da média do erro absoluto entre os valores reais e valores estimados. O erro encontrado foi considerado razoável dentro de um limite de tolerância aceitável sob o ponto de vista operacional. Como foram desenvolvidos poucos modelos de rede, por enquanto, a avaliação através deste método ainda é preliminar.

Os resultados apresentados são preliminares. Deverão ser realizados outros testes com outro banco de dados mais completo, da mesma problemática, que está em fase de coleta.

Com relação à aplicação da técnica de RNA, novos testes serão realizados com variação da quantidade de neurônios da camada escondida e com variação das funções de ativação em todas as camadas, a fim de buscar uma melhor combinação no treinamento da rede, para um menor erro com relação aos dados reais.

Após a definição da “melhor” técnica será, então, construído o modelo matemático que terá como objetivo minimizar o teor cáustico contido na lama vermelha, cuja função objetivo será a minimização dos custos operacionais e as restrições envolverão, além dos limites de controle das variáveis de entrada do processo, também a função discriminante da “melhor” técnica. Com isso o rejeito da produção de alumina terá a possibilidade de ser utilizada em outros processos produtivos como, por exemplo, da construção civil.

REFERÊNCIAS

- BRAGA, A. P.; LUDEMIR, T. B.; CARVALHO, A. C. P. L. *Redes Neurais Artificiais: Teoria e Aplicações*. LTC – Livros Técnicos e Científicos / Editora S.A. – Rio de Janeiro, 2000.
- CAMPOS, M. M.; SAITO, K. *Sistemas inteligentes em controle e automação de processos*. Rio de Janeiro: Editora Ciência Moderna Ltda, 2004.
- FAUSSET, L. V. *Fundamentals of Neural Networks: architecture, algorithms, and applications*. New Jersey: Prentice Hall International, 1994.
- HAYKIN, S. *Redes Neurais: Princípios e Práticas; trad. Paulo Martins Engel. – 2. ed. – Porto Alegre: Bookman, 2005.*
- HAGAN, M. T. e MENHAJ, M. B. Training feedforward networks with Marquardt algorithm. *IEEE Transactions on Neural Networks*, v.5, n.6, p.989-993, 1994.
- HIND, R. A.; BHARGAVA, S. K. e GROCOTT, S. C. The surface chemistry of Bayer process solids: a review. *Colloids and surfaces A. Physicochemical and engineering aspects*, n.146, p.359-374, 1999.
- LARSON, R.; FARBER, B. *Estatística Aplicada. 4ª. ed. São Paulo: Pearson Prentice Hall, 2010.*
- LEVINE, D. M.; BERENSON, M. L.; STEPHAN, D. *Estatística: Teoria e Aplicações*. Rio de Janeiro: LTC, 2005.
- LUZ, A. B.; LINS, F. A. F. *Rochas & Minerais Industriais: usos e especificações. 2. ed. Rio de Janeiro: CETEM/MCT, 2005.*
- SILVA FILHO, E. B.; ALVES, M. C. M. e MOTA, M. Lama vermelha da indústria de beneficiamento de alumina: produção, características, disposição e aplicações alternativas. *Revista Matéria*, v.12, n.2, p.322-338, 2007.