# PREVISÃO DO TEOR DE SILÍCIO NO FERRO-GUSA UTILIZANDO REDES NEURAIS ARTIFICIAIS E ALGORTIMO DE PODA Ana P. M. Diniz\*, José L. F. Salles\*, Klaus F. Coco\*, Flávio S. V. Gomes\*\*

\*Departamento de Engenharia Elétrica, Universidade Federal do Espírito Santo - UFES Av. Fernando Ferrari, 514 - Goiabeiras, Vitória – ES, 29075-910

*Emails*: ana.diniz@aluno.ufes.br, jleandro@ele.ufes.br, klaus.coco@ufes.br

\*\*Departamento de Engenharia Elétrica, Universidade Federal Rural de Pernambuco – UFRPE Unidade Acadêmica do Cabo de Santo Agostinho, 300 – Garapu, Santo Agostinho - PE E-mail: EMAIL:flavio.svgomes@ufrpe.br

**Abstract**— The silicon content forecasting models in molten-iron can assist the blast furnace operators in decision making during the production process, contributing not only to the quality of the final product but also to reducing the costs associated with their production. This paper proposes the use of an Artificial Neural Network (ANN), associated to a pruning algorithmfor for the selection of input variables of the model, in order to forecast the silicon content time series. After the application of the pruning algorithm, in addition to a decrease in the size of the network, it was observed a significant improvement of the silicon content forecasting performed up to 6 hours in advance.

Keywords-Blast Furnace, Silicon Content, Forecasting, Artificial Neural Networks, Pruning Algorithm.

**Resumo**— Os modelos de previsão do teor de silício no ferro-gusa podem auxiliar os operadores de alto-forno na tomada de decisão durante o processo produtivo, contribuindo não só na qualidade do produto final como também para a redução dos custos associados à sua produção. Este artigo propõe o uso de uma Rede Neural Artificial (RNA), associada a um algoritmo de poda para a seleção das variáveis de entrada do modelo, objetivando-se prever a série temporal do teor de silício. Após a aplicação do algoritmo de poda, além de uma diminuição do tamanho da rede, observou-se uma melhoria significativa das previsões do silício realizadas com antecedência de até 6 horas.

Palavras-chave- Alto-forno, Teor de Silício, Previsão, Redes Neurais Artificiais, Algoritmo de Poda.

# 1 Introdução

As usinas siderúrgicas produzem o ferro-gusa no alto-forno e posteriormente o encaminham para a refino na aciaria, resultando no aço. A qualidade do ferro-gusa produzido é importante para determinar o quão dispendioso será produzir o aço, bem como restringir em que tipos finais de aço o ferro-gusa poderá ser empregado (Mourão et al., 2011).

No entanto, devido às condições hostis no interior do alto-forno e à constante ocorrência de distúrbios durante o processo, a avaliação do estado interno do processo por medição direta não é uma tarefa trivial. Em vez disso, o estado interno pode ser avaliado baseando-se no comportamento de outras variáveis do processo que podem ser medidas diretamente e que potencialmente afetam o estado do processo, i.e., utilizando dados históricos que descrevem a evolução temporal de uma ou mais variáveis do processo, é possível, então, monitorar e até mesmo prever o estado de funcionamento do alto-forno (Saxén et al., 2013).

No caso em questão, o teor de silício no ferrogusa é utilizado por ser um importante indicador do estado térmico de um alto-forno e, consequentemente, ser capaz de refletir qual a qualidade do aço e o quão dispendioso será produzi-lo (Geerdes et al., 2009). Os valores elevados do teor de silício indicam um excedente de coque dentro do forno (uma grande reserva de energia), enquanto que valores mais baixos indicam a diminuição do mesmo. Assim, uma vez que o custo do coque é dominante na fabricação do ferro-gusa, há benefícios econômicos bastante evidentes no que concerne ao controle mais rigoroso do teor de silício (Saxén and Pettersson, 2007; Saxén et al., 2013).

Na aciaria, o conteúdo de silício no ferro-gusa é fundamental para o fornecimento da energia ao processo, devido à sua oxidação. No entanto, teores muito elevados de silício podem comprometer a produtividade em virtude da constante necessidade de ajustes do processo (Mourão et al., 2011).

Somando-se o longo tempo transcorrido desde o carregamento da matéria-prima no topo do alto-forno até sua retirada na forma de ferro e o tempo necessário para que os resultados da análise laboratorial da composição química do ferro-gusa, bem como do percentual de silício presente no mesmo, fique à disposição dos operadores, faz com que atitudes necessárias, visando a garantia de um produto final de boa qualidade e com baixo custo, sejam tomadas tardiamente (Bag, 2007).

Ao longo dos anos, foram publicados diversos trabalhos abordando o uso de técnicas lineares e nãolineares para a identificação dos mais variados modelos, em geral utilizando entradas exógenas, com o propósito de prever, e até mesmo controlar, o conteúdo de silício do ferro-gusa e demais parâmetros relativos à qualidade do processo de alto-forno. No entanto, em virtude do grande número de variáveis envolvidas no processo e que, potencialmente influenciam a variável de saída, a maioria dos trabalhos destaca a dificuldade na escolha das entradas exógenas, bem como dos seus respectivos atrasos, durante a concepção do modelo (Bag, 2007; Chen, Wang and Han, 2010; Nurkkala, Pettersson and Saxén, 2011; Saxén and Pettersson, 2007; Saxen et al., 2013).

Neste contexto, em virtude do grande interesse demonstrado por uma siderúrgica local em aprimorar a qualidade de seu produto final, este trabalho visa a elaboração de uma ferramenta capaz de prever, com até 6 horas de antecedência, o conteúdo de silício no ferro-gusa, baseando-se não só em seus dados históricos como também em demais variáveis do processo que influenciam de forma significativa no valor presente do silício. Para tanto serão utilizadas redes neurais com atraso de tempo (Time Delay Neural Networks - TDNN), auxiliadas pela expertise dos operadores e especialistas de alto-forno e por ferramentas de otimização, como o Algoritmo de Poda, para a seleção das variáveis de entrada bem como dos seus respectivos atrasos. Portanto, os resultados advindos desta pesquisa possibilitarão não apenas o aumento na qualidade do produto final, como também a redução dos custos associados à sua produção.

# 2 O Processo de Alto-Forno

O alto-forno é um grande reator termo-químico empregado na produção de ferro-gusa, principal matéria-prima para a produção do aço. Pela parte superior do alto-forno, são carregadas as matériasprimas ferrosas, combustível e fundentes. Na parte inferior, efetua-se o sopro de ar quente pelas ventaneiras, podendo haver também injeções auxiliares de combustível, com o intuito de gerar calor ao processo através da combustão do carbono (Mourão et al., 2011).

O objetivo deste processo é produzir uma liga, denominada de ferro-gusa, a uma temperatura de aproximadamente 1500°C, no estado líquido. Além disto, são também gerados, como subprodutos, a escória e o gás de alto-forno. No cadinho do altoforno, o ferro-gusa e a escória produzidos se separam por diferença de densidade: a escória possui um peso específico menor do que o gusa e, portanto, flutua sobre ele (Geerdes et al., 2009; Mourão et al., 2011).

Geerdes et al. (2009) explicam que, em geral, o alto-forno é vazado de 8 a 14 vezes por dia através do furo de gusa, correspondendo à uma duração média de vazamento entre 90 e 180 minutos. Além disto, o tempo de residência da carga no interior do forno é de aproximadamente 6 horas. Assim, um vazamento de 2 horas, representa um terço do conteúdo do alto forno transformado em ferro e escória. Grande parte dos fornos de alta produtividade, o vazamento ocorre de maneira alternada e simétrica, assim, no momento em que um furo de gusa é tamponado, o outro simetricamente oposto é aberto. Cada vazamento corresponde à uma corrida de gusa e o intervalo de tempo entre as corridas pode ser reduzido para zero (vazamento contínuo) ou, em virtude de desvios operacionais, ocorrer aberturas paralelas.

De uma maneira geral, dentre os elementos de liga que compõem o ferro-gusa, o silício é um dos mais relevantes. A sua importância é devido a sua capacidade em expressar não somente o estado térmico do forno, como também a qualidade do produto final e os custos associados à produção (Mourão et al., 2011). Um baixo teor de silício indica um arrefecimento do forno, podendo comprometer sua operação, enquanto que altos valores indicam um desperdício de coque, mostrado através da excessiva geração de calor (Saxén and Pettersson, 2007; Saxén et al., 2013).

Ao sair do alto-forno, o gusa liquido é levado à aciaria, onde é refinado e transformado em aço, através dos convertedores. O aporte térmico do gusa é o principal responsável pelo fornecimento de calor ao processo e, por isso, este material deve apresentar características físico-quimicas adequadas e com a menor variação possível, caso contrário, a produtividade é comprometida em virtude da constante necessidade de ajustes ao processo. Logo, um alto teor de silício pode prejudicar o processo de conversão do ferro em aço e, consequentemente, gerar gastos ainda maiores (Campos, 1983; Geerdes et al., 2009).

Em contrapartida, a concentração do silício no gusa líquido representa boa parte da energia necessária para o sistema de convertedores, uma vez que a oxidação do silício promove uma reação exotérmica. Por esta razão, baixos teores de silício alteram as condições de oxidação, prejudicando a formação de escória no início do processo, aumentando a eminência de projeções de metal para fora do forno (Campos, 1983).

#### 2.1 Principais variáveis envolvidas no processo

A Tabela 1 apresenta as 28 potenciais variáveis de entrada, selecionadas a partir de uma criteriosa análise da literatura existente, bem como pela avaliação dos engenheiros especialistas e operadores do alto-forno estudado.

É importante destacar que, apesar das variáveis pré-selecionadas apresentarem correlações com o percentual de silício no ferro-gusa, esta correlação não deve ser analisada separadamente, uma vez que, devido à complexidade da dinâmica do alto-forno, a ação de uma variável em conjunto com as demais, pode influenciar positivamente ou negativamente a ação de outra. As redundâncias diretas ou indiretas entre as variáveis são necessárias, uma vez que também em virtude da complexa dinâmica do processo, não se sabe *a priori* qual ou quais parâmetros serão mais representativos para o modelo.

De acordo com os especialistas e operadores do alto-forno estudado, em geral, o tempo de residência da carga no interior do forno é de cerca de 6 horas. Com relação às variáveis descritas na Tabela 1, com exceção das variáveis *Coke Rate* e Relação Minério/Coque, que possuem tempo médio de resposta no processo de 8 horas, todas as demais variáveis apresentam uma resposta de 3 horas.

Com relação às informações pertinentes as corridas de ferro-gusa, que acompanham a série de silício, nota-se que a cada corrida ocorrem de dois a três carregamentos de carro torpedo e que, normalmente, em cada carregamento é coletada uma amostra do ferro-gusa é coletada no canal de gusa para análise laboratorial. Posteriormente, o percentual de silício encontrado nesta análise, referente ao determinado carregamento, é atualizado no sistema. Todos estes fatores colaboram para a irregularidade do tempo de amostragem do percentual de silício, que varia em torno de 50 minutos.

Tabela 1. Potenciais variáveis de entrada do modelo e suas principais características.

N°	Variável de Entrada	Unidade de Medida	
1	Temperatura de Chama	°C	
2	Relação CO/CO <sub>2</sub>	-	
3	Eficiência do gás H <sub>2</sub>	%	
4	Velocidade de Produção de Gusa	ton gusa/min	
5	Coke Rate	kg/ton gusa	
6	PCI Rate	kg/ton gusa	
7	Redução Direta	%	
8	Relação Minério/Coque	-	
9	Índice térmico H0	10 <sup>3</sup> kcal/ton gusa	
10	IFGCN: Índice de Fluxo Gasoso Central Norte	-	
11	IFGPN: Índice de Fluxo Gasoso Periférico Norte	-	
12	IFGCS: Índice de Fluxo Gasoso Central Sul	-	
13	IFGPS: Índice de Fluxo Gasoso Periférico Sul	-	
14	Volume de Sopro	Nm <sup>3</sup> /min	
15	Umidade de Sopro	g/Nm <sup>3</sup>	
16	Taxa de Enriquecimento de O2	%	
17	Temperatura de Sopro	°C	
18	Pressão de Sopro	Kg/cm <sup>2</sup>	
19	Pressão de topo	Kg/cm <sup>2</sup>	
20	Temperatura de topo	°C	
21	Composição de H <sub>2</sub> no topo	%	
22	Composição de N2 no topo	%	
23	Composição de CO no topo	%	
24	Composição de CO2 no topo	%	
25	Eficiência do gás CO	%	
26	Perdas Térmicas pelo Stave	10 <sup>4</sup> Kcal/h	
27	Temperatura do ferro-gusa	°C	
28	Força-Eletromotriz	%	

Outras variáveis como a temperatura do ferrogusa e a força-eletromotriz (*ElectroMotive Force* – EMF) também são mensuradas a cada carregamento de carro torpedo e, por isso, também apresentam mesmo período de amostragem irregular que o teor de silício. As demais variáveis descritas na Tabela 1 possuem um período de amostragem de 30 minutos.

# **3 Redes Neurais Artificiais**

As Redes Neurais Artificiais (RNA) são bastante favoráveis na modelagem de processos não-lineares que possuem características complexas e/ou desconhecidas, em contrapartida com os métodos determinísticos tradicionais. Elas funcionam como processadores constituídos de unidades de processamento simples, denominados neurônios, que possuem capacidade de armazenar o conhecimento obtido através de experiências e adaptar-se perante às alterações do ambiente, através de um processo de aprendizagem (Haykin, 2001).

Uma RNA capaz de modelar a dinâmica de um processo, através de uma série temporal, é conhecida como rede neural dinâmica. Esta topologia de rede é conhecida como *Time Delay Neural Networks* (TDNN): tratam-se de uma extensão das redes *MultiLayer Perceptron* (MLP), com a introdução de operadores de atraso unitário em cada uma das variáveis de entrada, objetivando-se, assim, representar a dinâmica (relações de causa e efeito) do processo não-linear que gerou a série temporal analisada (Hu and Hwang, 2002).

Uma das maneiras de se representar a TDNN é através de um modelo não-linear autorregressivo com entradas exógenas (*Nonlinear AutoRegressive model with eXogenous inputs* – NARX), a qual inclui operadores de atraso tanto nas entradas quanto na própria saída, realizando uma espécie de filtragem dinâmica, nos quais os valores passados das entradas são usados para prever valores futuros (Haykin, 2001).

# 3.1 Algoritmo de Poda

Para a solução de problemas com redes neurais, frequentemente são necessárias o uso de redes altamente estruturadas e de tamanhos bastante elevados. Uma regra geral para que seja obtida uma boa generalização é a minimização do tamanho da rede, sem que o desempenho seja comprometido (Norgaard et al., 1996).

Em sua pesquisa, Sietsma and Dow (1988) demostraram os benefícios da redução do número de neurônios em redes neurais previamente treinadas. Além da redução do tamanho operacional, estas redes aumentaram a sua capacidade de generalização e reduziram o efeito do ruído.

Uma estratégia para a poda de redes neurais dinâmicas baseadas em modelos entrada-saída, é a aplicação de algoritmos fundamentados no cálculo da matriz hessiana, isto é, utilizando a informação sobre as derivadas de segunda ordem da superfície do erro. O objetivo desta técnica é estabelecer uma relação entre a complexidade da rede e o desempenho do erro de treinamento, a partir do cálculo da sensibilidade (Norgaard et al., 1996),

Através de uma rede completamente conectada e com um determinado número de parâmetros, os suficientes para mapear as relações entrada-saída, é construído um modelo local da superfície do erro para prever analiticamente o efeito da retirada de cada peso sináptico sobre o erro da saída da rede. A partir de então, os pesos que apresentarem as menores influências sobre a rede são podados. O resultado é uma rede com menos pesos e parcialmente conectada (LeCun et al., 1990).

No método intitulado Dano Cerebral Ótimo (*Optimal Brain Damage* – OBD), LeCun et al. (1990) simplifica o cálculo das sensibilidades fazendo a suposição de que a Hessiana seja uma matriz diagonal, reduzindo substancialmente o esforço computacional. Entretanto, a computação completa da matriz Hessiana torna o método mais poderoso. Este procedimento é intitulado Cirurgião Cerebral Ótimo (*Optimal Brain Surgeon* – OBS).

### 4 Metodologia

Para o desenvolvimento deste estudo, serão utilizados dados oriundos de um Alto-forno de uma siderúrgica localizada na região da Grande Vitória, Espírito Santo.

Inicialmente, serão coletadas as 28 variáveis de processo que mais influenciam no percentual de silício do ferro-gusa. Estas variáveis são responsáveis por descrever a dinâmica do processo e, por isso, são utilizadas pelos próprios operadores como parâmetros de avaliação da qualidade da produção. Também serão obtidos os dados relativos ao percentual de silício, resultante da análise laboratorial.

Devido ao interesse da siderúrgica em precisar os dados a cada 60 minutos, todo o conjunto de dados será reamostrado de forma a possuir um período de amostragem regular de 60 minutos. Além disto, todos os dados coletados serão tratados, para a correção de *outliers* e ausência de medições, bem como para a exclusão de grandes períodos de "nãonormalidade" ou paradas de alto-forno. Segundo os especialistas, entende-se por operação "normal" ou períodos de "normalidade" os momentos de regularidade do processo, isto é, instantes em que não ocorreram contingências e anomalias.

Em virtude da diversidade de magnitude existente entre as variáveis, após as etapas de tratamento, o conjunto de dados será escalonado para dentro da mesma faixa de valores, através da normalização pela amplitude para o intervalo [0,1; 0,9], fazendo uso da Equação (1).

$$x_{norm} = 0.1 + 0.8. \left(\frac{x_i - x_{min}}{x_{max} - x_{min}}\right)$$
 (1)

em que  $x_i$  é o valor da variável a ser normalizada,  $x_{min}$  e  $x_{max}$  são, respectivamente, os valores mínimo e máximo de assumidos pela variável e  $x_{norm}$  é o valor normalizado desta variável (Chen, Wang and Han, 2010).

Com o intuito de reduzir o nível de ruído, será aplicado um filtro não-linear, utilizando uma *wavelet* de base do tipo Symlet de ordem 4, por assemelharse mais com a forma característica dos sinais coletados. É interessante destacar que, após a decomposição *Wavelet*, a maior parte das informações referentes aos ruídos são encontradas nas subséries de altafrequência e, em geral, estas subséries apresentam pouca informação a respeito do sinal original. Assim, é possível ajustar o sinal resultante de maneira a conter maior nível de informação de baixa frequência e pouca informação de alta frequência (Misiti et al., 2017).

Serão avaliados e comparados modelos de previsão utilizando a topologia de rede do tipo NARX, com uma camada de entrada, uma camada oculta e uma camada de saída. A princípio, o atraso máximo de cada uma das entradas será igual ao seu respectivo maior tempo médio de resposta no processo. Neste sentido, com exceção das variáveis Relação Minério/Coque e Coke Rate, que possuem tempo de resposta mais lento (8 horas), em todas as demais será aplicado um atraso de 3 amostras, i. e., 3 horas.

Considerando que a topologia NARX utiliza amostras autorregressivas da própria série a ser prevista, torna-se necessário analisar a dinâmica do processo de produção para que seja possível estipular o número de atrasos de tempo os quais serão aplicados nesta série, na entrada da rede. Seguindo este raciocínio, para implementação desta rede, será estipulado um atraso de tempo de 6 amostras, ou seja, para prever o teor de silício das próximas amostras, serão necessários os 6 valores passados de suas entradas.

O tamanho da camada oculta será determinado testando-se redes com estruturas variando de 5 a 100 neurônios ocultos, incrementado 5 neurônios a cada teste. Para cada rede, será calculada a média do Erro Quadrático Médio (*Mean Squared Error* – MSE), obtido após cinco treinamentos consecutivos, utilizando o algoritmo de aprendizagem de Levenberg-Marquardt. Após todos os testes, será escolhida o tamanho da camada oculta através da rede que apresentar o menor MSE médio.

Considerando o grande número de variáveis envolvidas no processo, também serão aplicados Algoritmos de Poda (do tipo OBS) nas mesmas, visando remover as entradas e/ou atrasos pouco significativos para o modelo, facilitando assim, a escolha de variáveis relevantes e possibilitando a redução do número de neurônios da camada oculta (Nurkkala, Pettersson and Saxén, 2011; Saxén and Pettersson, 2007).

Para todos os testes serão selecionadas as primeiras 2.880 amostras do conjunto de dados, divididas em dois grandes blocos. O primeiro bloco consiste das 2.160 primeiras amostras que serão utilizadas para o treinamento supervisionado da rede neural e o segundo bloco, composto pelas 720 amostras restantes, será aplicado na validação dos testes.

Os critérios escolhidos para mensurar o desempenho dos modelos foram o MSE, o Erro Percentual Absoluto Médio (*Mean Absolute Percentage Error* – MAPE) e o Percentual de Erro Absoluto (PEA) dados, respectivamente, por:

$$MSE = \frac{1}{N} \sum_{t=1}^{N} (y_t - \hat{y}_t)^2$$
(2)

$$MAPE (\%) = \frac{1}{N} \sum_{t=1}^{N} \frac{|y_t - \hat{y}_t|}{y_t} .100$$
(3)

$$PEA (\%) = \frac{1}{N} \sum_{t=1}^{N} N_E . 100$$
(4)

onde,

$$N_E = \begin{cases} 1, & se |y_t - \hat{y}_t| < 0.05 \\ 0, & caso \ contrário \end{cases}$$

em que N é o número total de amostras,  $y_t$  representa o valor real da amostra e  $\hat{y}_t$  refere-se ao valor estimado.

O MSE e o MAPE são medidas já consolidadas na validação de modelos de previsão. No entanto, Chen, Wang and Han (2010) afirmam que, na prática, um importante critério a ser avaliado quando se trata de modelos de previsão do conteúdo de silício é o PEA menor que uma tolerância pré-definida. Em geral, a variação do PEA é definida de acordo com as necessidades dos especialistas do alto-forno e, por esta razão, para este artigo será definida uma variação de 0,05.

### **5** Resultados

Durante o treinamento, o desempenho das redes foi mensurado através do MSE. A partir de então, determinou-se o número de neurônios da camada oculta que resulte em uma rede que melhor se ajuste à estimação dos dados um passo à frente.

A Tabela 2 apresenta a média dos MSE obtidos para cada tamanho de rede testado. Observa-se a partir desta que a estrutura que apresentou a menor média do MSE durante o treinamento, foi a que possuía 90 neurônios ocultos.

Tabela 2. Desempenho resultante durante o treinamento de uma
rede NARX com diferentes tamanhos de camada oculta.

Número de neurônios da camada oculta	MSE	
5	2,18 x 10 <sup>-3</sup>	
10	1,91 x 10 <sup>-3</sup>	
15	1,67 x 10 <sup>-3</sup>	
20	1,50 x 10 <sup>-3</sup>	
25	1,41 x 10 <sup>-3</sup>	
30	1,30 x 10 <sup>-3</sup>	
35	1,14 x 10 <sup>-3</sup>	
40	8,52 x 10 <sup>-4</sup>	
45	7,91 x 10 <sup>-4</sup>	
50	6,23 x 10 <sup>-4</sup>	
55	5,72 x 10 <sup>-4</sup>	
60	4,16 x 10 <sup>-4</sup>	
65	4,29 x 10 <sup>-4</sup>	
70	1,89 x 10 <sup>-4</sup>	
75	1,53 x 10 <sup>-4</sup>	
80	4,02 x 10 <sup>-5</sup>	
85	9,81 x 10 <sup>-6</sup>	
90	1,06 x 10 <sup>-10</sup>	
95	6,81 x 10 <sup>-10</sup>	
100	9,98 x 10 <sup>-10</sup>	

Com o intuito de eliminar a redundância dos dados e, consequentemente, aumentar o desempenho da previsão, empregou-se o Algoritmo de Poda do tipo OBS. É importante comentar que em virtude da complexidade do OBS e do consequente longo tempo de processamento, as redes foram inicializadas com 10 neurônios ocultos.

No entanto, para a avaliação das entradas mais significativas resultantes do OBS, o Algortimo de Poda foi implementado 10 vezes. No final de cada implementação (poda), as entradas remanescentes, bem como os seus respectivos atrasos, foram armazenadas. Através da Tabelas 3 e 4 é possível verificar, respectivamente, as entradas/atrasos remanescentes no final de cada implementação do OBS, bem como os critérios de desempenho obtidos para uma previsão de 1 passo à frente.

Tabela 3. Entradas e atrasos obtidos após cada iteração do Algoritmo de Poda aplicado à Rede NARX.

Poda	Entrada (Atraso)		
1	4 (2); 8(3); 9(2); 11(2);16(1); 17(2); 28(2);		
	Saída Autorregressiva (1, 2, 3, 4, 5)		
2	1(1, 2, 3); 2(1, 2, 3); 3(1, 2, 3); 4(1, 2, 3); 5(1, 2, 3, 4, 5, 6, 7); 6(1, 2, 3); 7(1, 2, 3); 8(1, 2, 3, 4, 5, 6, 7); 9(1, 2, 3); 10(1, 2, 3); 11(1, 2, 3); 12(1, 2, 3);		
	(2, 3); 13(1, 2, 3); 14(1, 2, 3); 15(1, 2, 3); 15(1, 2, 3); 16(1, 2,		
	10(1, 2, 3); 17(1, 2, 3); 19(1, 2, 3); 20(1, 2, 3); 21(1, 2, 3); 22(1, 2, 3); 24(1, 2, 3); 25(1, 2, 3);		
	21(1, 2, 3); 22(1, 2, 3); 24(1, 2, 3); 25(1, 2, 3); 26(1, 2, 2); 27(1, 2, 2); 28(1, 2, 2);		
	20(1, 2, 3); 27(1, 2, 3); 28(1, 2, 3); Solda Automographica (1, 2, 2, 4, 5)		
	Salda Autorregressiva $(1, 2, 3, 4, 5)$		
	1(1, 2, 3); 2(1, 2, 3); 3(1, 2, 3); 4(1, 2, 3);		
	5(1, 2, 3, 4, 5, 6, 7); 6(1, 2, 3); 7(1, 2, 3);		
	8(1, 2, 3, 4, 5, 6, 7); 9(1, 2, 3); 10(1, 2, 3); 11(1, 2, 3); 12(1, 3, 3); 12(1,		
3	(2, 3); 13(1, 2, 3); 14(1, 2, 3); 15(1, 2, 3); 15(1, 2, 3); 16(1, 2, 3); 17(1, 2, 2), 17(1, 2)		
	16(1, 2, 3); 1/(1, 2, 3); 19(1, 2, 3); 20(1, 2, 3);		
	21(1, 2, 3); 22(1, 2, 3); 24(1, 2, 3); 25(1, 2, 3);		
	26(1, 2, 3); 27(1, 2, 3); 28(1, 2, 3);		
	Salda Autorregressiva $(1, 2, 5, 4, 5)$		
	1(1, 3); 3(1, 2); 4(3); 5(1, 3, 5); 6(3); 8(1, 3); 9(1, 2, 3); 10(2, 2), 11(2, 2), 10(1), 12(1, 2), 14(2, 2), 15(1, 2, 2);		
4	10(2, 3); 11(2, 3); 12(1); 13(1, 3); 14(2, 3); 15(1, 2, 3); 16(2, 2); 17(1, 2); 10(1, 2); 20(1); 21(2); 24(1); 25(1);		
4	16(2, 3); 17(1, 3); 19(1, 3); 20(1); 21(2); 24(1); 25(1);		
	20(1, 2, 3); 27(2); 28(2, 3);		
	Saida Autorregressiva $(1, 2, 3, 4, 5)$		
5	20(1); Saida Autorregressiva $(1, 2, 3, 4, 5)$		
6	Saída Autorregressiva (1, 2, 3, 4, 5)		
7	12(1, 3); 14(1); 17(3); 20(1); 27(2);		
7	Saída Autorregressiva (1, 2, 3, 4, 5)		
8	3(2); $5(7)$ ; $6(3)$ ; $9(3)$ ; $11(3)$ , $12(1, 3)$ ; $13(2)$ ; $15(1)$ ; $17(2)$ ;		
	23(2); $25(1)$ ; $26(2)$ ; $26(2)$ ; $26(2)$ ; $12(1, 2$		
	2(2), 2(2), 4(2), 5(1, 2, 2), 7(2), 9(1), 11(1, 2), 12(2		
9	2(2); 3(3); 4(3); 3(1, 2, 3); 7(3); 8(1); 11(1, 2); 12(2); 14(1); 15(1); 16(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2); 20(2); 24(2); 25(2); 27(1, 2); 29(1); 15(2); 20(2);		
	14(1); 15(1); 10(3); 20(2); 24(2); 25(2); 27(1, 2); 28(1);		
10	Salua Autoriegressiva $(1, 2, 3, 4, 5)$		
10	5(3); Saida Autorregressiva $(1, 2, 3, 4, 5)$		

Conforme pode ser observado nas Tabelas 3 e 4, deve-se destacar a forte ocorrência dos valores autorregressivos da série de silício frente às demais entradas para o modelo NARX. Neste sentido, é possível verificar a forte influência dos valores autorregressivos durante a modelagem. Outro detalhe importante foi observado na poda de número seis, na qual foi obtida uma rede com apenas valores autorregressivos da série de silício, resultando assim, em um modelo autorregressivo puro (*Nonlinear Autoregressive –*  NAR) com 4 neurônios ocultos, MSE de 0,001, MAPE de 7,65% e PEA < 0,05 de 89,32%.

No entanto, a estrutura que apresentou o melhor desempenho foi o modelo de número 9, com um MSE de 0,0011, MAPE de 6,41%, PEA < 0,05 de 93,82% e 6 neurônios ocultos e, portanto, a mesma será escolhida para aplicação das previsões.

Poda	Neurônios Ocultos	MSE	MAPE (%)	PEA < 0,05 (%)
1	3	0,0089	7,21	92,97
2	10	0,0011	7,32	91,292
3	5	0,0008	7,35	92,696
4	8	0,0010	7,79	90,59
5	3	0,0008	6,32	93,11
6	4	0,0010	7,65	89,32
7	6	0,0011	8,00	89,18
8	7	0,0007	6,89	92,13
9	6	0,0011	6,41	93,82
10	3	0,0008	7,36	92,696

Tabela 4. Critérios de desempenho obtidos após cada iteração do Algoritmo de Poda aplicado à Rede NARX.

#### 6 Discussão

Para uma melhor observação dos resultados obtidos, a Fig. 1 apresenta uma comparação entre as previsões com e sem a implementação do Algoritmo de Poda para previsões de 1, 3 e 6 passos à frente, das primeiras 300 amostras do conjunto de testes. Observa-se que, em geral, embora as previsões sem a aplicação da poda consigam acompanhar, mesmo com um pequeno atraso, o comportamento da série real, há pequenas discrepâncias nas amplitudes dos sinais. Com relação às previsões resultantes após a aplicação da poda, é possível notar que o comportamento destas apresentam maiores semelhanças com série real, embora ainda ocorram pequenas diferenças entre os sinais.

Na Tabela 5 são apresentados os desempenhos obtidos pela Rede NARX, para as previsões de 1, 3 e 6 passos à frente antes e depois da implementação do Algoritmo de Poda.

Tabela 5. Comparativo dos critérios de desempenho das previsões antes e depois da aplicação do Algoritmo de Poda na Rede NARX.

Horizonte de Previsão		1 hora	3 horas	6 horas
Antes do	MSE	0,0041	0,0134	0,0118
Algoritmo	MAPE (%)	17,58	32,15	29,04
de Poda	PEA < 0,05 (%)	55,61	33,80	36,63
Depois do	MSE	0,0010	0,0053	0,0064
Algoritmo	MAPE (%)	6,41	17,98	20,13
de Poda	PEA < 0,05 (%)	93,82	56,33	49,50

Observa-se uma melhoria significativa no desempenho das redes, além de uma considerável diminuição no tamanho da camada oculta, por exemplo, a rede NARX que antes possuía 90 neurônios ocultos, com MSE de 0,0134, MAPE de 32,15% e PEA < 0,05 de 33,80% (horizonte de previsão de 3 horas), agora com 6 neurônios ocultos, foi obtido um aumento do PEA < 0,05 em aproximadamente 60% e uma redução do MAPE em cerca de 44,17%.

No entanto, nota-se que, conforme aumenta-se o horizonte de previsão, maior é o atraso entre a saída prevista e o sinal original. Alguns autores (Waller and Saxén, 2000) sugerem que estes atrasos podem ser devido aos termos autorregressivos, que tendem a tornar as previsões mais inertes.



Figura 1. Comparativo entre as previsões da Rede NARX antes e depois da aplicação do Algoritmo de Poda para horizontes de previsão de a) 1 hora; b) 3 horas; c) 6 horas.

# 7 Conclusão

Este artigo utilizou uma rede neural do tipo NARX para prever amostras do percentual de silício no ferro-gusa com horizontes de previsão variando de 1 a 6 horas. Inicialmente, para o desenvolvimento do modelo NARX foram selecionadas 28 variáveis de processo que mais influenciam no percentual de silício do ferro-gusa, através da avaliação dos engenheiros especialistas e operadores de alto-forno, bem como em trabalhos anteriormente desenvolvidos. Os atrasos de cada variável foram escolhidos baseandose no seu tempo de resposta no processo de altoforno. Este modelo é então comparado ao modelo resultante após a aplicação do Algoritmo de Poda, utilizando o método OBS, responsável por apurar quais destas variáveis apresentam maior grau de significância ao modelo. Além disto, devido à natureza deste algoritmo, os atrasos de cada entrada também serão definidos através do mesmo.

Após a aplicação do Algoritmo de Poda, além de uma considerável diminuição do tamanho da rede (redução de 75% dos neurônios de entrada e 40% da camada oculta), observou-se uma melhoria significativa no desempenho da rede neural. Esta melhoria é justificada pelo fato de que em virtude do grande número de variáveis exógenas empregadas na camada de entrada, o excesso de informações e consequente redundância dos dados podem confundir a rede durante o processo de aprendizagem.

Tendo em vista que, uma variação de 5% na previsão do teor de silício não ocasiona graves impactos no processo produtivo em questão, a rede NARX resultante da poda, com apenas 6 neurônios na camada oculta foi capaz de prever dentro da faixa de erro tolerável. Foi obtido um MAPE de 6,41% e 93,82% de acerto na faixa de tolerância, para um horizonte de previsão de 1 hora, frente à um MAPE de 20,13% e 49,5% de acerto na faixa de tolerância, para um horizonte de previsão de 6 horas.

Portanto, este artigo demonstra a viabilidade da modelagem e previsão utilizando RNA e ferramentas de otimização, embora que em trabalhos futuros será discutida a utilização de outras ferramentas, com objetivo de melhorar ainda mais a previsão para horizontes maiores. Para trabalhos futuros, pretendese implementar o sistema de previsão na planta real.

## Agradecimentos

Os autores agradecem ao apoio do PPGEE, ao suporte financeiro propiciado pelo CAPES e ao incentivo a realização desta pesquisa por parte da Arcelor Mittal Tubarão.

#### Referências

- Bag, S. K. (2007). ANN based prediction of blast furnace parameters. *Journal – The Institution of Engineers*, vol. 68, no. 1, pp. 37-42.
- Campos, V. F. (1983). *Tecnologia de fabricação do aço líquido*. 2. ed. UFMG, Belo Horizonte.
- Chen W.; Wang, B. X. and Han, H. L. (2010). Prediction and control for silicon content in pig iron of blast furnace by integrating artificial neural network with genetic algorithm. *Ironmaking & Steelmaking*, vol. 37, no. 6, pp. 458–463.
- Geerdes, M.; Toxopeus, H. and Van Der Vliet, C. (2009). *Modern Blast Furnace Ironmaking: an introduction*. 2nd ed. IOS Press., Amsterdam.
- Haykin, S. (1999). *Neural networks and learning machines*. 3rd ed. Pearson, New Jersey.
- Hu, Y. H. and Hwang, J. (2002) *Handbook of neural* network signal processing. CRC Press, Boca Raton.
- Lecun, Y., Denker, J. S., Solla, S. A., Howard, R. E.; Jackel, L. D. (1990). Optimal brain damage. In: Touretzky, D., ed., *Advances in Neural Information Processing Systems*. Denver: Morgan Kaufmann, pp. 598-605.
- Misiti, M.; Misiti, Y.; Oppenheim, G.; Poggi, J-M. (2017). *Wavelet Toolbox User's Guide*. The MathWorks Inc., Natick.

- Mourão, M. B.; Yokoji, A.; Malynowskyj, A.; Leandro, C. A. S.; Takano, C; Quites, E. E. C.; Gentile, E. F.; Lenz E Silva, G. F. B; Bolota, J. R.; Gonçalves, M. and Faco, R. J. (2011). *Introdução à Siderurgia.* Associação Brasileira de Metalurgia, Materiais e Mineração, São Paulo.
- Norgaard, M.; Ravn, O.; Hansen, L.; Poulsen, N. (1996). The nnsysid toolbox a matlab toolbox for system identification with neural networks. In: *IEEE Symposium on Computer-Aided Control System Design*. Dearborn: IEEE Press., pp. 374-379.
- Nurkkala, A.; Pettersson, F. and Saxén, H. (2011). Non linear modeling method applied to prediction of hot metal silicon in the ironmaking blast furnace. *Ind. Eng. Chem. Res.*, vol. 50, no. 15, pp. 9236–9248.
- Saxén, H.; Gao, C. and Gao, Z. (2013). Data-Driven Time Discrete Models for Dynamic Prediction of the Hot Metal Silicon Content in the Blast Furnace — A Review. *IEEE Transactions On Industrial Informatics*, vol. 9, no. 4, pp. 2213-2225.
- Saxén, H. and Pettersson, F. (2007). Nonlinear prediction of the hot metal silicon content in the blast furnace, *ISIJ Int.*, vol. 47, no. 12, pp. 1732– 1737.
- Sietsma, J. and Dow, R. J. F. (1988). Neural net pruning-why and how. In: *IEEE International Conference on Neural Networks*, San Diego: IEEE Press., pp. 325-333.