

**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**FACULDADE DE ENGENHARIA QUÍMICA**

**ÁREA DE CONCENTRAÇÃO**

**SISTEMAS DE PROCESSOS QUÍMICOS E INFORMÁTICA**

**DESENVOLVIMENTO DE UM MODELO DE PREDIÇÃO DE  
CONDIÇÕES OPERACIONAIS DE ALTOS-FORNOS SIDERÚRGICOS  
COM BASE EM REDES NEURAIS ARTIFICIAIS**

Autor: Fernando Tadeu Pereira de Medeiros

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup> Ana Maria Frattini Fileti

Co-orientador: Dr André Pitasse da Cunha

Tese de Doutorado apresentada à Faculdade de Engenharia Química como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Química.

Campinas – São Paulo

Março / 2006

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

M467d	<p>Medeiros, Fernando Tadeu Pereira de Desenvolvimento de um modelo de predição de condições operacionais de altos-fornos siderúrgicos com base em redes neurais artificiais / Fernando Tadeu Pereira de Medeiros.--Campinas, SP: [s.n.], 2006.</p> <p>Orientadores: Ana Maria Frattini Fileti, André Pitasse da Cunha Tese (Doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química.</p> <p>1. Siderurgia. 2. Altos-fornos. 3. Modelos matemáticos. 4. Redes neurais (Computação). 5. Inteligência artificial. I. Fileti, Ana Maria Frattini. II. Cunha, André Pitasse da. III. Universidade Estadual de Campinas. Faculdade de Engenharia Química. IV. Título.</p>
-------	--

Titulo em Inglês: Development of a prediction model for ironmaking blast furnace operational conditions based on artificial neural networks

Palavras-chave em Inglês: Ironmaking, Blast furnace, Hybrid model, Artificial neural network

Área de concentração: Sistemas de Processos Químicos e Informática

Titulação: Doutor em Engenharia Química

Banca examinadora: José Adilson de Castro, Rodney Alexandre Ferreira Rodrigues, Antonio Carlos Luz Lisboa e Roger Josef Zemp

Data da defesa: 17/03/2006

Este exemplar corresponde à versão final de tese de doutorado em engenharia química

---

Profa. Dra. Ana Maria Frattini Fileti

Tese de doutorado defendida por Fernando Tadeu Pereira de Medeiros e aprovada em 17 de março de 2006 pela banca constituída pelos doutores:

---

Profa. Dra. Ana Maria Frattini Fileti  
FEQ/UNICAMP – Orientadora

---

Prof. Dr. José Adilson de Castro  
Universidade Federal Fluminense

---

Dr. Rodney Alexandre Ferreira Rodrigues  
CPQBA/UNICAMP

---

Prof. Dr. Antonio Carlos Luz Lisboa  
FEQ/UNICAMP

---

Prof. Dr. Roger Josep Zemp  
FEQ/UNICAMP

## **AGRADECIMENTOS**

À Professora Doutora Ana Maria Frattini Fileti, pela confiança, pela orientação, conhecimento transmitido e generosidade.

Ao Doutor André Pitasse da Cunha pelo incentivo, orientação e companheirismo.

Aos executivos da Companhia Siderúrgica Nacional, Enéas Garcia Diniz, Diretor Executivo de Produção, Walter Luiz da Costa Reis, Gerente Geral de Processos Siderúrgicos, e Fabiam Franklin, Gerente Geral de Altos-Fornos, pela oportunidade e apoio.

Ao Professor Doutor José Adilson de Castro pelas discussões a cerca do modelamento de processos e apoio na UFF.

Aos colegas Sebastião Jorge Xavier Noblat, José Borges de Araújo Neto e José Maria Reinaldo pelo apoio e incentivo para a implantação industrial do modelo.

À minha esposa, Celina, e filhas, Maitê e Tatiana, pelo incentivo e motivação.

## RESUMO

Este trabalho descreve o desenvolvimento de um modelo capaz de prever as condições operacionais do processo de redução de minério de ferro em alto-forno a partir de características das matérias-primas e da composição da carga e do sopro, fornecendo aos operadores e engenheiros de processo, uma ferramenta de planejamento e de análise de desempenho da operação.

A motivação para este trabalho deveu-se à constatação de que existe uma deficiência de modelos capazes de identificar cotidianamente os fatores críticos na operação dos altos-fornos siderúrgicos. Os modelos encontrados são excessivamente simples para preverem condições futuras de processo ou são demasiadamente complexos para o uso operacional diário ou no planejamento estratégico da produção.

Os altos-fornos e os conversores de aço são as unidades centrais de uma usina siderúrgica integrada. O primeiro obtém o ferro primário a partir dos óxidos contidos nos minérios e o segundo refina e ajusta a composição química do metal produzindo o aço. O modelo desenvolvido e aplicado tem natureza híbrida, combinando algoritmos simuladores baseados em balanços de massa e energia com variável grau de desvio estequiométrico e térmico. O valor de cada indicador de desvio é previsto por uma rede neural cujas variáveis de entrada quantificam características das matérias-primas e condições de sopro e carga do alto-forno.

A aplicação industrial do modelo comprovou sua capacidade de prever as condições do processo e sua aplicação resultou em aumento da produtividade média do processo e menor consumo específico de redutores, decorrentes da melhor efetividade das ações operacionais. Além disso, o modelo, associado a um módulo de balanço do setor primário da usina, vem sendo aplicado na simulação de alternativas de padrões operacionais, atividade fundamental para o planejamento estratégico do negócio.

## **ABSTRACT**

This work describes the development of a model capable of evaluating and predicting iron ore reduction process in blast-furnaces based on raw materials characteristics as well as burden and blast composition. It provides a planning and analysing tool to operators and process engineers.

The motivation for this development resides on the lack of this kind of model in the ironmaking industry. The many models found are either too simple and not capable of predicting raw materials parameters effects on the furnace performance or too complex. The latter are useful in technology potential identifications but not in the daily work or ordinary operation planning.

Blast-furnaces and oxygen converters are the core units in an integrated steel works. The former one produces primary iron from oxides bared by iron ores and the latter one refines molten iron into steel, adjusting its chemical composition. The developed and applied model is hybrid in nature, combining simulating algorithms based on mass and energy balances with variable lack of fitness. The value of each lack-of-fitness index is estimated by a neural model in which the input quantify burden materials characteristics as well as blow and charging conditions.

The model has shown its predicting capacity during its industrial application, which lead to higher average productivity and lower specific fuel consumption are expected following better operation action and process planning efficiency. In addition, the model is now being associated to a plant mass balance module to asses alternative operation pattern simulation for the hole company which is a fundamental activity in strategic business planning.

# ÍNDICE

<b>CAPITULO 1</b>	<b>INTRODUÇÃO</b>	1
<b>CAPITULO 2</b>	<b>DESCRIÇÃO DO PROCESSO E REVISÃO BIBLIOGRÁFICA</b>	11
	2.1 ESTADO ATUAL DA TECNOLOGIA DE ALTOS-FORNOS	11
	2.2 EVOLUÇÃO DA TECNOLOGIA	16
	2.3 EVOLUÇÃO DOS MODELOS DE PROCESSO	19
	2.4 REDES NEURAIS ARTIFICIAIS	26
	2.5 APLICAÇÕES NA SIDERURGIA	42
	2.6 JUSTIFICATIV DA PROPOSTA DE TESE	47
<b>CAPITULO 3</b>	<b>FORMULAÇÃO DO MODELO HÍBRIDO</b>	51
	3.1 ÍNDICE DE DESEMPENHO FLUIDO-DINÂMICO	54
	3.2 ÍNDICE DE DESEMPENHO FÍSICO-QUÍMICO	60
	3.3 ARQUITETURA DA REDE NEURAL	67
	3.4 DEFINIÇÃO DE VARIÁVEIS E DO BANCO DE DADOS	68
	3.5 PAPEL DO MODELO NA PRÁTICA INDUSTRIAL	81
<b>CAPÍTULO 4</b>	<b>RESULTADOS E DISCUSSÃO</b>	83
	4.1 TREINAMENTO E GENERALIZAÇÃO DA REDE	83
	4.2 SIMULAÇÃO “OFF-LINE”	92
	4.2 TESTE INDUSTRIAL	95
<b>CAPITULO 5</b>	<b>CONCLUSÕES E SUGESTÕES</b>	99
	<b>NOMENCLATURA</b>	103
	<b>BIBLIOGRAFIA</b>	107
<b>ANEXO I</b>	<b>LISTAGENS DE PROGRAMAS</b>	113
<b>ANEXO II</b>	<b>PARÂMETROS DA REDE ADOTADA</b>	117
<b>ANEXO III</b>	<b>TELA DO PROTÓTIPO DO APLICATIVO DESENVOLVIDO</b>	119
<b>APÊNDICE I</b>	<b>CONTRIBUIÇÃO TÉCNICA AO CONGRESSO <i>ADCHEM 2006</i></b>	
<b>APÊNDICE II</b>	<b>ARTIGO SUBMETIDO AO CORPO EDITORIAL DA REVISTA <i>IRONMAKING &amp; STEELMAKING</i></b>	

## **CAPITULO 1 - INTRODUÇÃO**

O processo siderúrgico se dá em etapas: redução, refino, solidificação, tratamento termomecânico e tratamento superficial.

Na etapa de redução, os óxidos de ferro presentes nos minérios são reduzidos a ferro metálico impuro, ou ferro primário, matéria prima para a produção do aço. As principais impurezas são o silício, o enxofre e o fósforo.

O ferro primário quando obtido em estado líquido é chamado de ferro-gusa, ou simplesmente gusa, enquanto aquele produzido sem fusão é chamado de ferro-esponja devido à sua natureza porosa. Uma alternativa ao consumo de minérios de ferro é a fusão de sucata de ferro e de aço.

No refino, o teor de impurezas presentes no ferro primário é ajustado enquanto outros elementos de liga são adicionados para compor o aço com as propriedades desejadas. O tratamento termo-mecânico posterior confere a forma (chapa, folha, bobina, perfil etc), e contribui também para as propriedades mecânicas necessárias a cada aplicação. O tratamento superficial visa à proteção do material contra a corrosão através de um revestimento metálico ou orgânico.

As rotas possíveis não são necessariamente equivalentes. Há uma certa relação entre rotas e classes de produtos, que foi se cristalizando ao longo do tempo. As grandes usinas integradas a alto-forno tendem a concentrar sua produção nos aços planos. As usinas médias e pequenas tradicionais se deslocaram para a fabricação de aços longos (barras e perfis). Usinas originalmente com amplo espectro de produtos se adaptaram nas últimas décadas à nova ordem optando por uma das duas vertentes. Uma nova geração de mini-usinas, com capacidade de 0,5 a 1,0 milhão de toneladas de aço por ano vem tentando quebrar esta regra com dificuldade pois seus custos são mais elevados que os das grandes usinas de planos. Estas têm capacidade maior que 3,0 milhões de toneladas por ano.

Os conversores a oxigênio necessitam de ferro primário líquido, ferro-gusa, e por isso são geralmente associados a altos-fornos. A aciaria elétrica, a base de fornos a arco elétrico, permite a fusão da carga metálica antes do refino e, portanto, pode receber 100% de sucata ou uma mistura de sucata e ferro-esponja. O ferro-esponja é produzido nos processos de redução direta. O ferro-gusa também pode ser produzido em processos de fusão redutora. Esses ainda não estão ainda tecnologicamente maduros e não têm produção significativa no cenário mundial.

Estima-se que dentre 1 bilhão de toneladas de aço produzidas no mundo durante o ano 2000, cerca de 64% seguiram a rota alto-forno + aciaria a oxigênio. A Figura 1.1 ilustra a evolução dessa distribuição. No Brasil, esta relação gira em torno de 78% (IBS, Estatísticas da Siderurgia, 2005)

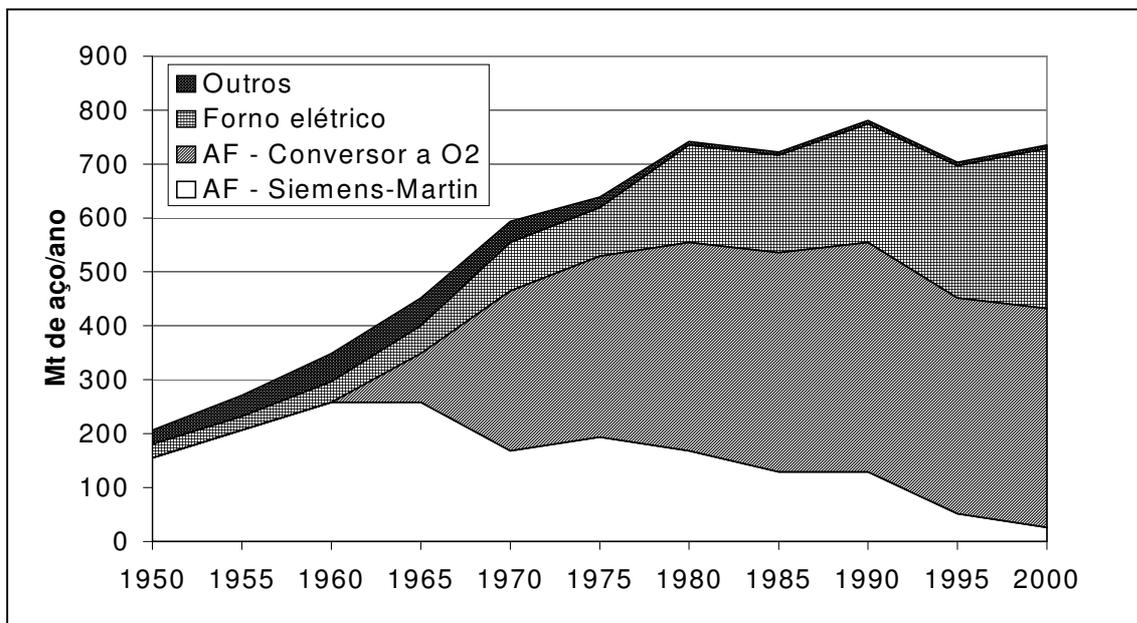


Figura 1.1 - Evolução da distribuição da produção mundial de aço por processo

O crescimento da fatia relativa à aciaria elétrica se explica pela crescente disponibilidade de sucata e de ferro esponja. Os fornos Siemens-Martin se tornaram obsoletos, tendo sido, em muitas usinas, substituídos por conversores a oxigênio. As rotas que envolvem o alto-forno são a base da indústria de aços

planos que abastece o setor automobilístico, o de eletrodomésticos, o de embalagens, o naval, o da construção civil e ferroviária. O aumento de produtividade industrial nestas áreas é uma das razões para a estagnação da demanda mundial de aços planos observada na Figura 1.1. Em todos os setores se produz cada dia com menos aço, seja pela redução do peso dos componentes, seja pelo menor rendimento dos processos de fabricação. O uso de materiais sucedâneos, principalmente plásticos, também é fator relevante.

No ano 2004 o Brasil foi o oitavo maior produtor de aço e o maior exportador de minério de ferro do mundo segundo o Instituto Brasileiro de Siderurgia (IBS, 2005).

Países de maior tradição industrial, como os Estados Unidos da América e o Japão, têm reduzido a sua produção siderúrgica devido aos seus elevados custos operacionais. Os recentes episódios em que o governo dos EUA criou novas barreiras alfandegárias aos produtos siderúrgicos estrangeiros atestam essa dificuldade. Dados mercadológicos indicam que já existe uma produção excedente nos mercados europeu e principalmente no americano. Hoje a produção ultrapassa demanda sendo necessário recorrer às exportações com preços abaixo dos aplicados no mercado interno para garantir uma rotação mais dinâmica dos estoques segundo o Iron and Steel International Institute (IISI, 2005).

A Ásia tem uma participação bastante intensa na produção mundial de aço bruto, seguida pela Europa e pela América. A China desponta como a maior potência siderúrgica da atualidade apresentando crescimentos anuais superiores a 10% neste setor. Em 2004 a China produziu 272 milhões de toneladas de aço, 29% da produção mundial segundo dados do Instituto Brasileiro de Siderurgia..

Mas o que se observa hoje, em relação há 20 anos, é o aumento da participação dos países em desenvolvimento na produção siderúrgica mundial que se elevou de 30% para 55%.

As mudanças que vêm ocorrendo nos novos estados-membros da União Europeia têm importância significativa na siderurgia mundial. Para aderir à União Europeia a indústria siderúrgica da Europa Central e Oriental passou e passa por profundas alterações, à medida que os países preparam o setor e as suas economias para se adaptarem à nova ordem. Tais mudanças implicaram no encerramento das instalações de produção ineficientes e aumento da produtividade, como fizeram os atuais estados-membros na década de 80, objetivando concluir a grande vaga de reestruturações até ao final de 2006.

A siderurgia é um setor industrial importante na Europa Oriental, especialmente na Polónia, República Checa, Eslováquia, Hungria, Eslovênia, Bulgária, Romênia e também na Turquia. Neste grupo, a produção anual de cerca de 45 milhões de toneladas representa 5% do total da produção mundial e proporciona cerca de 220 mil postos de trabalho. Esta reformulação do setor visa a estabelecer uma base económica sólida para garantir que as empresas destes países prosperem e concorram em condições de igualdade com as outras empresas dos outros países membros da UE. Embora os auxílios estatais para a reestruturação da indústria sejam proibidos pela legislação comunitária durante a etapa transitória, acordos foram estabelecidos visando minimizar os problemas sociais e as perturbações do mercado.

No Brasil, estima-se que o setor siderúrgico nacional tenha investido cerca de US\$ 15 bilhões na década de 90. Esses investimentos foram voltados à modernização e ampliação da capacidade das usinas, logo após a privatização do setor. A partir do final dos anos 90, os investimentos começaram a diminuir e passaram a privilegiar não tanto ao aumento da capacidade instalada mas a produtividade dos processos e a melhoria de qualidade do aço. O Instituto Brasileiro de Siderurgia estima que o investimento no setor será de US\$ 12,5 bilhões de 2005 a 2010. Entre 1994 e 2004 o investimento foi contabilizado em quase US\$ 14,0 bilhões.

Uma das explicações para a retração em termos de investimentos no aumento da capacidade instalada é a demanda por aço. O consumo aparente de aço bruto cresce a uma taxa média de 1% a 2% ao ano no Brasil. Verifica-se na Figura 1.2 que o Brasil não tem conseguido, devido a problemas estruturais, um período de crescimento regular com as taxas necessárias a uma boa consolidação do consumo aparente de aço bruto no país. Tem-se um consumo per capita pequeno em relação a outros mercados: 100 kg por habitante, contra os 400 a 650 kg registrados nos Estados Unidos, no Japão, na União Européia e na Coréia do Sul.

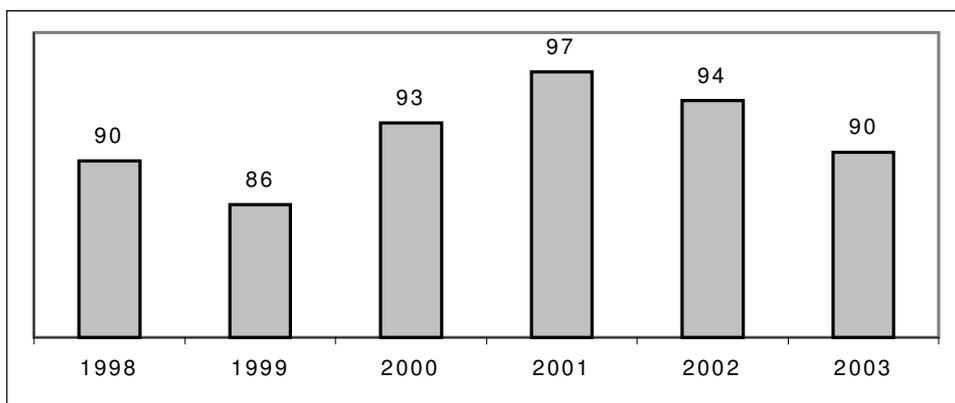


Figura 1.2 Consumo per-capita do Brasil (kg aço bruto/hab.ano) - Fonte: IBS

No mundo, a produção também segue estagnada, com se vê na Figura 1.3. O contraste é a China, hoje a maior produtora e consumidora de aço do mundo. As importações chinesas elevaram as cotações do aço bruto apesar da baixa demanda mundial.

Certamente as exportações brasileiras de minério para a China continuarão crescendo, mas o comercio futuro de aço entre os dois países ainda é uma incógnita. O balanço comercial no setor siderúrgico dependerá da competitividade da industria brasileira. Se a tecnologia neste setor no Brasil não acompanhar a evolução que se observa hoje na China corremos o risco de, mesmo sendo exportadores líquidos de aço, re-importarmos unidades de ferro na forma de alguns graus de aço.

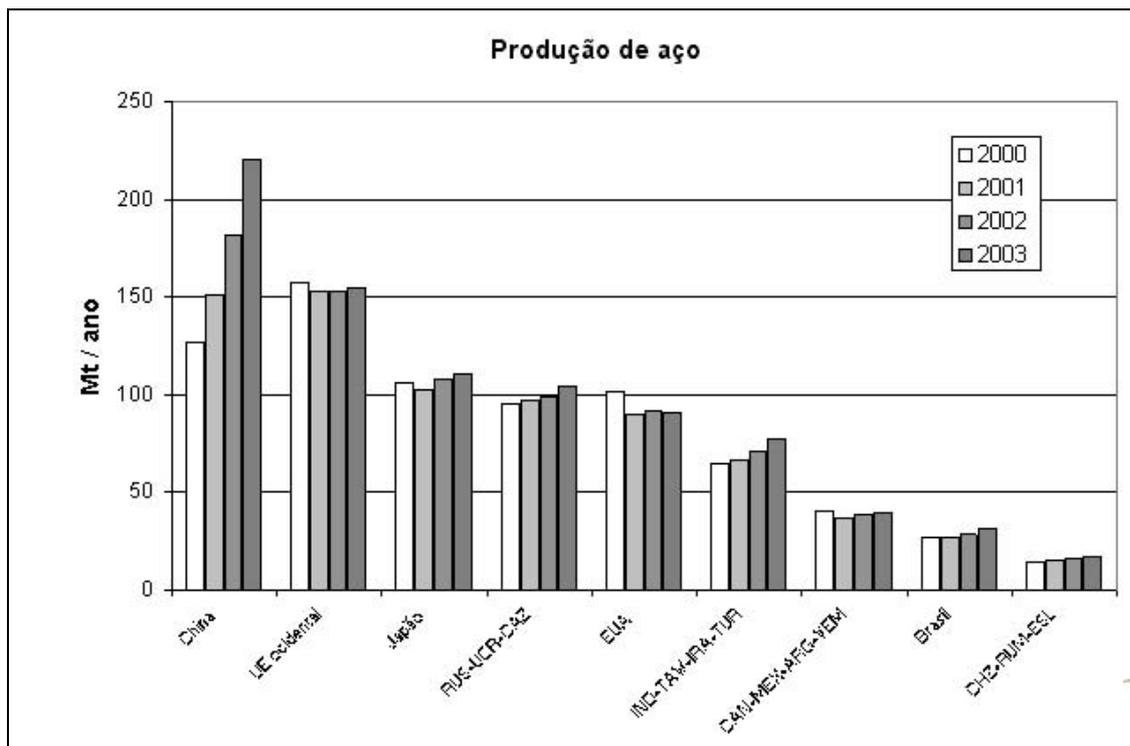


Figura 1.3 – Produção de aço em alguns países regiões. Fonte IISI

O setor siderúrgico brasileiro sente a redução do lucro líquido devido às despesas financeiras contraídas em dólar. Por questões de estabilidade cambial, não ocorrendo apreciação do real diante do dólar, há a necessidade de despendar mais reais para pagar o mesmo montante de dívida em dólares. Embora parcela considerável das receitas do setor seja proveniente das exportações, as siderúrgicas têm que arcar com seus insumos cotados em dólares.

As empresas têm que importar coque e carvão mineral, que, além de constituírem custos em dólares, tiveram os preços elevados depois do aumento da demanda mundial provocado pela expansão da China.

O mercado chinês importou, em 2003, 28,5 milhões de toneladas de aço. Quantidade bem próxima da produção total brasileira, que ficou em 31,2 milhões de toneladas. A produção do gigante oriental, 272 milhões de toneladas em 2004 e crescendo a 27,9% este ano, enquanto a brasileira cresceu apenas 4% sobre o ano passado e chegou a 32,9 milhões de toneladas. A preocupação é redobrada,

levando-se em conta a possibilidade do caminho da auto-suficiência da China, uma vez que a Ásia se constituiu em importante mercado consumidor das exportações brasileiras. Em 2004 absorveu cerca de 42% do total de 18,4 milhões de toneladas vendidas ao exterior.

Espera-se que ocorra uma desaceleração da economia chinesa, fato este que ameaça o crescimento global (causando o impacto que seus investidores haviam previsto há muito tempo). O crescimento abrupto do país respondeu por mais de um sexto do crescimento da economia mundial em 2003. O esforço do governo chinês para reduzir a taxa de crescimento do país, fazendo-a passar dos 9,1% registrados em 2003, para 7% este ano, pode cortar meio ponto percentual da taxa de crescimento mundial em 2005, além de fazer a expansão da Ásia diminuir um ponto percentual, disse Stephen Roach, economista do Morgan Stanley de Nova York (Gazeta Mercantil, 2005). Tendo em vista este panorama, pode-se esperar que as siderúrgicas daquele país também sofram algum tipo de efeito relacionado à desaceleração econômica.

Por outro lado, a produção brasileira de ferro-gusa não pára de crescer desde 1996. Há quatro anos, as empresas do setor retornaram aos níveis atingidos no final dos anos 80. Em 2004, registraram expansão de 43% em relação a 2002, um salto de 6,8 milhões para 9,7 milhões de toneladas. A falta de investimento em reflorestamento, porém, poderá interromper a atual fase de expansão do setor.(IBS,2005)

Neste setor, a predominância no Brasil é do alto-forno de pequeno porte. Neste caso, os novos investimentos no país são feitos mantendo-se a tecnologia existente, apesar de algumas consultas sobre o sistema de redução direta, ainda sem definições. O grupo Gerdau herdou do estado uma usina com a tecnologia de redução direta na Bahia, destacada como a pioneira do gênero em território nacional, a Usiba. É possível que com o aumento nos preços das matérias-primas e da energia elétrica no mercado brasileiro e disponibilidade de gás natural, a

instalação de equipamentos de redução direta possa se ampliar no setor de não-planos ou exportador de ferro primário. A redução direta, porém, não parece estar nos planos das grandes usinas nacionais de aços planos por apresentar menor escala típica e custos atrelados ao mercado derivados de petróleo, mais instável que o de carvão.

Os dados e argumentos apresentados indicam que o Brasil tem grande potencial para ter papel mais significativo na siderurgia mundial. Tanto a produção de aço como a de tecnologia no setor têm sido inferiores ao que se poderia esperar do maior exportador de minério de ferro do mundo. Do lado da produção a reação, embora tardia, começa a se esboçar. É preciso que o meio acadêmico e os setores de desenvolvimento das empresas apoiem o setor produtivo neste esforço. A presente tese se enquadra nesta proposta.

Portanto é proposta da presente tese a formulação e parametrização de um modelo híbrido, com base em uma rede neural para prever parâmetros que alimentam módulos fenomenológicos. Trata-se do desenvolvimento de uma ferramenta de planejamento operacional capaz de avaliar padrões de processo alternativos indicando os mais adequados aos objetivos empresariais. A ferramenta torna a operação dos altos-fornos, hoje uma prática reativa, em uma atividade pró-ativa e, portanto, mais eficiente.

No Capítulo 2 serão apresentados aspectos tecnológicos e sua evolução histórica além dos modelos de processos e ferramentas matemáticas aplicadas.

O Capítulo 3 trata das dificuldades encontradas na operação dos altos-fornos, destacando a necessidade do desenvolvimento de ferramentas matemáticas de suporte mais eficientes. Apresenta-se uma contribuição para a maior eficácia do planejamento operacional a curto, médio e longo prazo, na forma de um modelo híbrido com base em redes neurais artificiais e em um balanço termo-químico que adota parâmetros previstos pela rede neural.

A coleta, a análise e o tratamento dos dados, resultados e discussões são descritos no Capítulo 4 evidenciando a necessidade de estruturação de um banco de dados dinâmico e representativo para a aplicação do modelo desenvolvido.

Conclusões e sugestões são apresentadas no Capítulo 5.

## **CAPÍTULO 2 - DESCRIÇÃO DO PROCESSO E REVISÃO BIBLIOGRÁFICA**

A comunidade siderúrgica mundial é relativamente pequena quando comparada a de outros setores industriais. Isso leva a uma grande uniformização de conceitos entre os especialistas. A relativa calma nos investimentos fora da China tem motivado a fusão das empresas de engenharia dedicadas a este ramo, de forma que os equipamentos, que pouco diferem em tecnologia, sejam cada vez mais parecidos. Essa característica se reflete na bibliografia especializada.

Neste capítulo, o processo será descrito procurando realçar os aspectos físico-químicos relevantes para o modelo desenvolvido neste trabalho. A evolução da tecnologia é apresentada visando a evidenciar como a tecnologia neste setor, na maior parte do tempo, andou na frente da ciência.

### **2.1 ESTADO ATUAL DA TECNOLOGIA DE ALTOS-FORNOS**

O alto-forno siderúrgico propriamente dito consiste de um reator formado por segmentos de cilindros e troncos de cone. A maior parte do volume interno opera em regime de contra corrente. Minério de ferro, em sua maior parte aglomerado, e coque de carvão mineral são carregados pelo topo enquanto ar modificado é soprado por ventaneiras situadas no seguimento inferior. Neste segmento está também o cadinho que acumula ferro-gusa e escória líquidos que são vazados por furos de corrida, ou furos de gusa. A Figura 2.1 mostra uma vista do Alto-Forno número 3 da CSN, realçando o contorno reator dentro da estrutura.

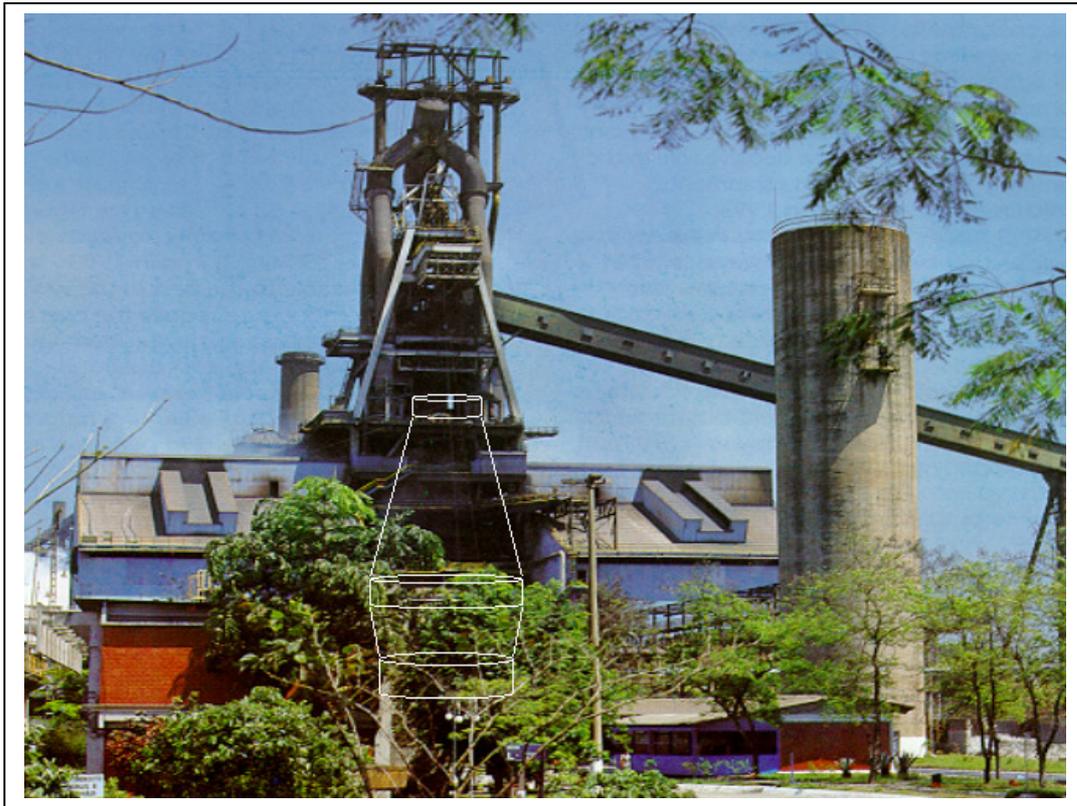


Figura 2.1 – Vista do Alto-Forno número 3 da CSN destacando o contorno do reator, propriamente dito.

Hoje é consensualmente aceito que, no processo de redução de minérios de ferro em altos-fornos, os fenômenos se distribuem nas zonas internas segundo a forma descrita a seguir:

a) Zona granular superior:

A carga é seca e pre-aquecida pelos gases quentes em ascensão que deixam a zona com temperatura ente 500K a 700K. As reações ocorrem em intensidade desprezível em função da velocidade dos gases e da baixa temperatura, embora os gases ainda tenham potencial redutor suficiente para reduzir a hematita ( $\text{Fe}_2\text{O}_3$ ) a magnetita ( $\text{Fe}_3\text{O}_4$ ).

b) Zona granular intermediária:

Nesta zona ocorre a reação de redução indireta, ou seja, sem regeneração do CO pelo carbono presente no coque, ou ainda, antes do início da reação de Boudouard, ou seja:



A temperatura em que a reação de Boudouard ocorre com maior intensidade depende da reatividade do coque, e varia na prática entre 1200K e 1350 K. Nos altos-fornos a carvão vegetal a temperatura de Boudouard é da ordem de 1100 K, o que leva a diferenças bastante importantes entre os dois processos. Os carbonatos presentes na carga são parcialmente decompostos nesta zona, completando-se a calcinação na zona granular inferior (von Bogdandy et alli, 1971).

c) Zona granular inferior:

Nesta região o coque toma parte efetivamente da redução dos óxidos de ferro. Ocorre a chamada redução direta. Idealmente todos os óxidos de ferro teriam se reduzido a wustita ( $\text{Fe}_y\text{O}$ ) antes de entrar nesta região. Como será visto mais tarde a quantificação deste desvio fará parte da avaliação do desempenho do processo.

d) Zona de coesão:

Enquanto as três zonas superiores se classificam através das diferentes reações químicas, esta região se caracteriza pela mudança do estado físico da carga portadora de óxidos de ferro. Nela a fusão parcial da carga torna as camadas ferríferas pastosas e depois líquidas. A fusão total da carga marca o limite inferior da zona. Nesta zona forma-se a escória primária, ou seja, aquela formada apenas pelos óxidos da carga ferrífera, antes da incorporação da cinza dos redutores.

e) Núcleo de coque:

O núcleo de coque compreende um volume limitado pela zona de coesão e é constituído pelo coque que se desloca continuamente em direção à zona de

combustão e pelo chamado "homem-morto", ou coque estacionário. Este é lentamente consumido pela dissolução do coque no gusa líquido acumulado no cadinho e em menor escala pela reação de Boudouard. A denominação "homem-morto", tradução equivocada de tronco ou bloco inerte (Lighter,1994), é bastante antiga e o conhecimento de suas características físicas não são recentes. O conhecimento dos fenômenos físico-químicos que ali ocorrem, por outro lado, é bem recente e precário em alguns aspectos.

f) Cadinho:

No cadinho, gusa e escória líquidos se acumulam em camadas imiscíveis ocupando o espaço entre as pedras de coque. O gusa, a mais densa das duas fases líquidas, escoam através da escória transferindo a ela a maior parte do enxofre que recebeu ao percolar o núcleo de coque.

Essa estrutura interna pode ser vista na Figura 2.2.

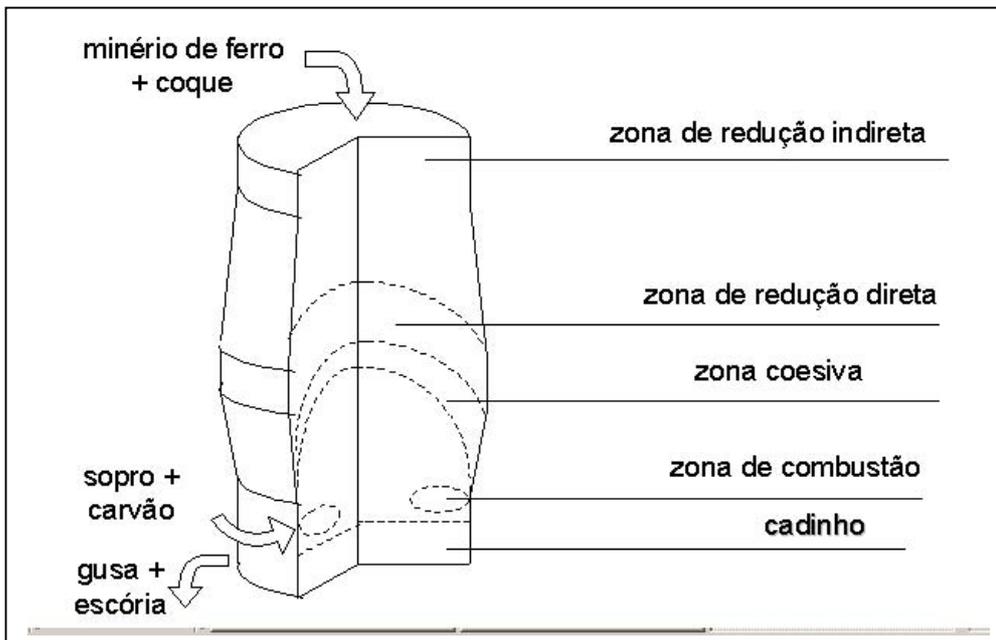


Figura 2.2 - Estrutura interna dos altos-fornos

## A Operação dos Altos-Fornos

O sistema operacional dos altos-fornos é bastante semelhante nas diversas usinas siderúrgicas do mundo. Há um operador líder que controla as condições de sopro e coordena as operações de carregamento e de vazamento. O carregamento é realizado automaticamente mas acompanhado por um outro operador na própria sala de controle. O vazamento é realizado por equipes que permanecem nas casas de corrida. Pode haver mais um operador para acompanhar o sistema de aquecimento do ar soprado, em muitos casos também responsável por outros equipamentos periféricos.

A Figura 2.3 mostra a sala de controle do Alto-Forno 3 da CSN, permitindo observar as três seções de comando, cada uma em um setor do painel. No centro fica o operador líder, responsável pelo sopro. Ao fundo o operador do carregamento e adiante o setor de aquecimento do ar e de injeção de carvão pulverizado.



Figura 2.3 – Aspecto da sala de controle do Alto-Forno 3 da CSN

O planejamento da operação é feito por um grupo de engenheiros e técnicos que consideram as metas de produção e qualidade, as condições das matérias-primas e definem os parâmetros de sopro e carga para os próximos dias. Correções mais imediatas são feitas pelo operador líder, o supervisor de turno ou membros da equipe técnica. Essa equipe faz os planos mensais e anuais de operação em conjunto com os fornecedores internos de matérias-primas e o cliente, a aciaria.

O automatismo da operação, por mais sofisticado que seja o sistema de aquisição de dados, limita-se às ações repetitivas. Qualquer modificação de padrão operacional é feita pelos operadores. Os sistemas especialistas adotados em alguns altos-fornos, como os da Voest-Alpine na Áustria, limitam-se a manipular uma série de regras formuladas por operadores experientes e a sugerir uma ou mais ações. Naquele sistema o operador pode tomar uma ação diferente das sugeridas, mas é obrigado a registrá-la em um relatório no próprio sistema para análise posterior pela equipe de engenharia de processo (Mauhart et alli, 2003).

## **2.2 EVOLUÇÃO DA TECNOLOGIA**

Considera-se que os altos-fornos surgiram na fase mais tardia da Idade Média, na forma dos "stückerfen" alemães, embora esses operassem em bateladas. Eram fornos de cuba, carregados pelo topo, com ventaneiras laterais próximas do fundo. Foram eles os primeiros a produzir industrialmente ferro no estado líquido, ou ferro-gusa, diferente do produto anteriormente obtido nas forjas de redução, as "forjas catalãs", que era pastoso e mais puro. Como resultado do aumento do tamanho desses fornos e do uso de foles movidos a água o processo passou a ser contínuo. A necessidade de purificar o ferro e ajustar o seu teor de carbono resultou no aparecimento dos processos de refino estabelecendo o padrão de processo em duas etapas, redução e refino, adotado até hoje na siderurgia (Bouman, 1978).

Durante o século XVII um alto-forno produzia menos de duas toneladas por dia a partir de minério bitolado e carvão vegetal. A preocupação com as florestas

resultou na conversão dos fornos para carvão e depois para coque. O uso de máquinas a vapor no final do século XVIII permitiu o aumento do tamanho dos fornos, mas o topo permaneceu aberto para a atmosfera. Passou-se a adicionar calcário para controle da viscosidade da escória. As taxas de consumo de coque eram da ordem de duas toneladas por uma de ferro gusa. Desde as origens do processo foi adotada a prática de carregar o redutor e o minério em camadas alternadas.

Durante a Revolução Industrial ocorreu significativa evolução dos componentes mecânicos do sistema mas os princípios do processo não mudaram. No começo do século XIX a escala de produção dos maiores altos-fornos era da ordem de 30 toneladas ao dia. A idéia de pré-aquecer o ar soprado foi outro passo significativo no desenvolvimento do processo, permitido cerca de 30% a mais de produção nos mesmos equipamentos. Ao mesmo tempo a campanha do revestimento refratário sofria considerável prolongamento através do uso de novos materiais que até 1890 não ultrapassavam dois anos, como descreve Holowaty (1969).

Em 1884 Bell descobriu que havia relação entre a proporção de CO e de CO<sub>2</sub> no gás e a produtividade do forno. Aparentemente ele foi o primeiro a tentar estabelecer as condições de equilíbrio no sistema Fe-O. Antes, em 1838, von Bunsen havia coletado amostras de gases através de um tubo introduzido a partir do topo do forno. O objetivo, no entanto, era o de tentar aproveitá-lo como combustível em caldeiras. Os primeiros dados, criteriosamente catalogados, aplicáveis ao processo apareceram com Darken e Gurry em 1946.

Só no século XX é que os materiais, minério e coque, passaram a ser peneirados antes do carregamento, marcando as primeiras práticas de preparação e controle da qualidade da carga (Bell, 1984).

Nas décadas de 50 e 60, experimentos e o tratamento de dados de operação dos altos-fornos se tornaram freqüentes e sistemáticos. Foram desenvolvidos

sistemas mecânicos para minimizar a heterogeneidade na permeabilidade da carga em decorrência da segregação dos materiais durante a operação de carga no topo dos fornos.

Os fenômenos associados à zona de amolecimento e fusão da carga ferrífera, hoje reconhecida como o coração do alto-forno, só foram iniciados na segunda metade da década de 70, após a dissecação de fornos japoneses (Sasaki e all, 1976; Kanbara et all, 1977;). Esses trabalhos foram de grande importância para o desenvolvimento da tecnologia e garantiram a liderança do Japão nesta área por muitos anos. A Figura 2.4 sintetiza as observações obtidas nesses estudos.

Engenheiros soviéticos tinham realizado experimentos semelhantes uma década antes sem, no entanto, tratar os dados obtidos de forma tão detalhada (Muraviev et all, 1970).

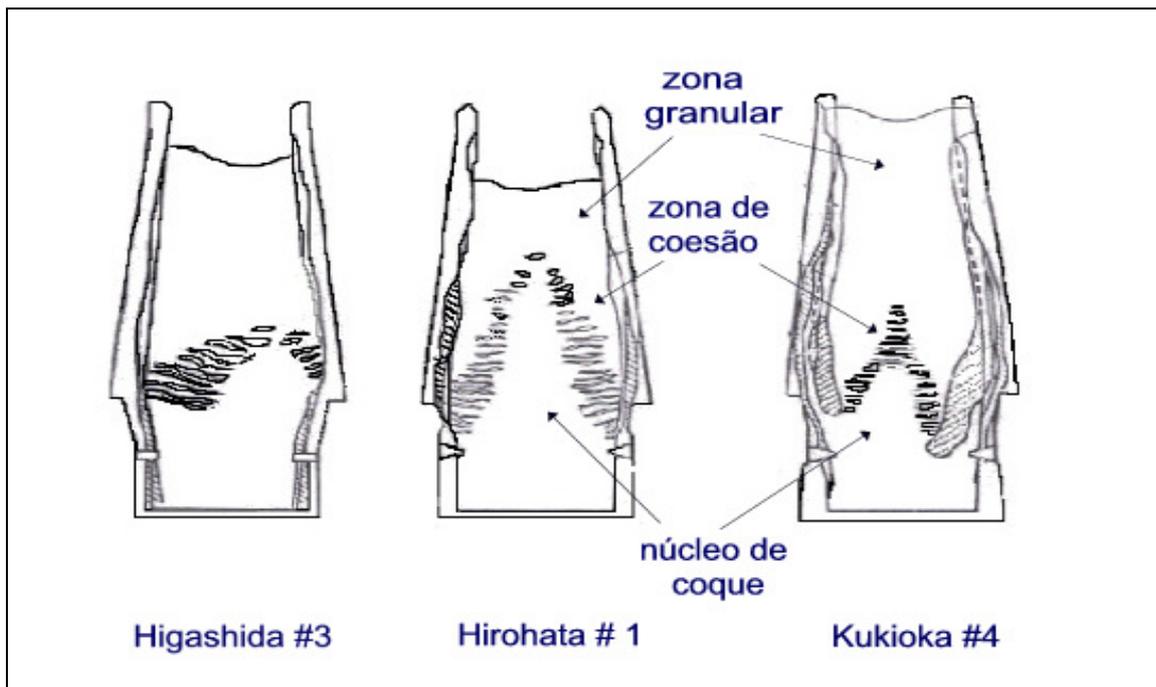


Figura 2.4 - Resultado do trabalho de dissecação de altos-fornos japoneses

A década de setenta marcou também o início da injeção de carvão pulverizado nos altos-fornos. Taxas de injeção de 200 kg por tonelada de gusa foram

praticadas duas décadas depois, mudando significativamente o desempenho e características do processo. No mesmo período os fornos passaram a trabalhar com pressão cada vez mais alta no topo. Na União Soviética, altas taxas de injeção de gás natural também foram praticadas. Esta prática chegou ao Brasil por volta em 1985, através da Companhia Siderúrgica Nacional (de Medeiros et alli, 1987).

O desenvolvimento de sistemas de distribuição de carga nos altos-fornos foi decorrência dos estudos de dissecação. Os estudos levaram a uma melhor definição da estrutura interna do reator, caracterizando as interações-chave em cada uma das zonas, sob os aspectos térmicos e químicos, lançando as bases para técnicas de ajuste da distribuição do fluxo gasoso através do leito. Este tema deu margem a uma série de modelos matemáticos que visavam prever a geometria das camadas de materiais no interior do forno (Heynert et alli, 1978; Hattori et alli, 1993).

Os fatos descritos nesta seção caracterizam a evolução de uma arte, que avançou por muito tempo através de uma sucessão de tentativas com erros e acertos. Nos últimos 50 anos é que a engenharia tomou o controle do desenvolvimento do processo transformando arte em tecnologia. Os operadores sabiam razoavelmente o que entrava e o que saia dos fornos porém muito pouco do que dentro deles acontecia. Até hoje muitos fenômenos que ocorrem nos altos-fornos ainda esperam ser explorados embora existam altos-fornos capazes de produzir quase 12 mil toneladas de gusa por dia.

## **2.3 EVOLUÇÃO DOS MODELOS DE PROCESSO**

### **Modelos de avaliação termoquímica**

Reichardt, em 1927, propôs um método de análise do desempenho de altos-fornos bastante avançado para a época e que é usado até hoje, isoladamente, ou contido

em modelos mais amplos. Consiste na divisão do forno em 5 zonas e no cálculo de um balanço térmico para cada uma delas. As zonas são aproximadamente aquelas apresentadas no item anterior, com exceção da zona de coesão, conceito não existente na época. Em um diagrama cartesiano, Reichardt designou o eixo horizontal para a entalpia e calor sensível acumulados na carga e disponível nos gases. O eixo vertical referencia a temperatura do gás e da carga. Na Figura 2.5 um diagrama simplificado é apresentado para um alto-forno hipotético.

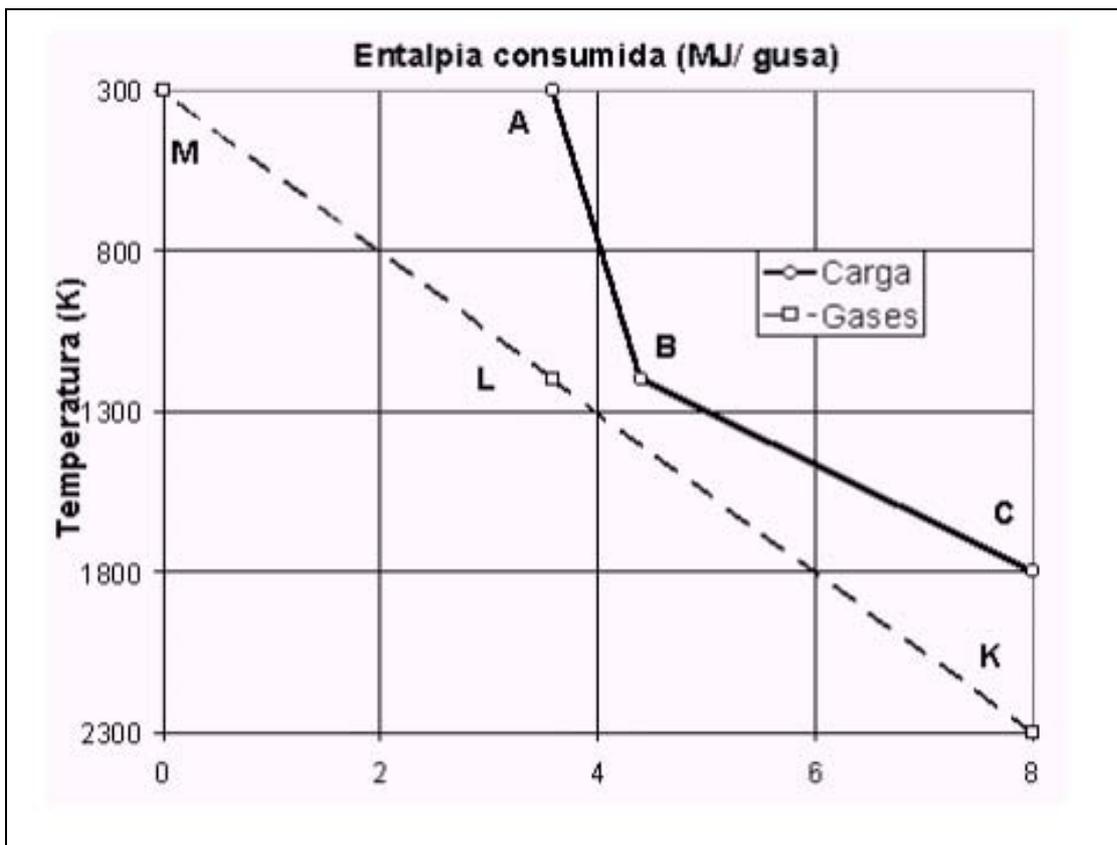


Figura 2.5 - Diagrama de Reichardt

No espaço em  $R^2$  aparecem duas linhas, uma para a carga e outra para os gases. O ponto A representa a carga e o ponto M o gás no topo do forno. Os pontos B e L fazem o mesmo, respectivamente, para o limite entre as zonas de redução direta e indireta. O ponto C representa o gusa e a escória no cadinho e o ponto K os gases

na saída da zona de combustão. A distância entre elas indica o potencial de trocas térmicas entre as duas fases.

O modelo constitui uma forma simples e eficiente para a análise das condições de trocas térmicas no processo. Outros pesquisadores, entre eles Ridgion (1962) e Witmann (1962), procuraram avançar nos conceitos originais incluindo maior número de reações químicas e maior número de zonas. Foi Kitaev, em 1967, quem melhor definiu os parâmetros para a análise dos fenômenos.

Rist e Meysson (1967) desenvolveram o chamado diagrama operacional que quantifica e representa graficamente o regime de troca de átomos de oxigênio entre gases e sólidos no alto-forno. Esse modelo é bastante útil e tem sido a base de muitos modelos mais recentes.

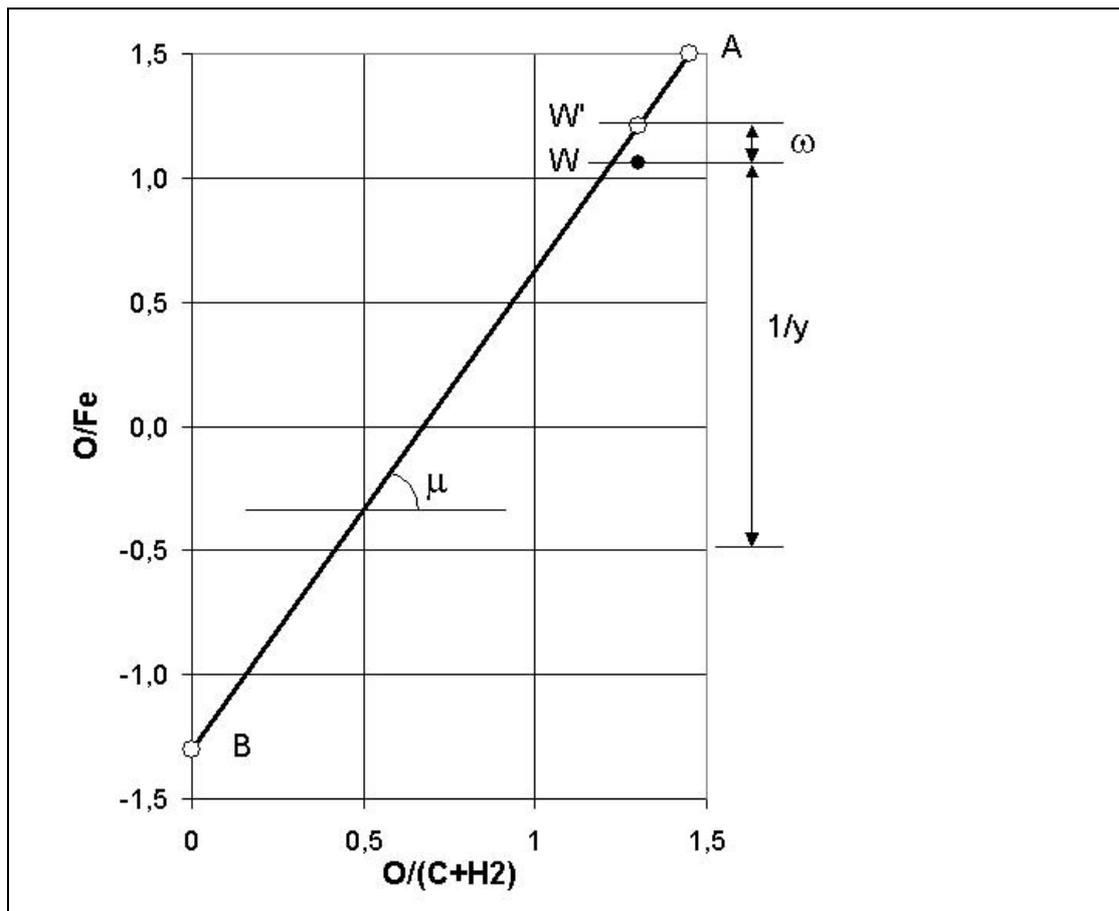


Figura 2.6 - Diagrama de Rist

A Figura 2.6 exemplifica a aplicação do modelo de Rist a um alto-forno através de um diagrama genérico. No eixo horizontal é representado o grau de oxidação do gás de processo através da relação atômica O/(C+H<sub>2</sub>). No eixo vertical é indicada a relação entre kg-átomos de oxigênio e de ferro. Na seção positiva do eixo está a parcela de oxigênio fornecida pela carga e na seção negativa, o oxigênio do ar soprado e dos outros elementos reduzidos da carga, com silício, manganês e fósforo. O coeficiente angular da reta operacional, (C+H<sub>2</sub>)/Fe, representa o consumo de redutor do processo,  $\mu$ . O ponto W indica as condições de equilíbrio químico entre o gás e a carga à temperatura em que começa a reação de Boudouard. Esse ponto representa o limite inferior da chamada zona de reserva química. O ponto W' pertence à reta operacional AB e representa a composição da carga em contato com o gás quando este atinge a composição de equilíbrio. A distância W'-W, ou  $\omega$ , quantifica o excesso de oxigênio da carga e portanto o desvio da idealidade da operação. Este modelo, dada a sua engenhosidade, ainda é um dos mais usados na prática. A interseção entre a reta AB e o eixo horizontal indica a relação entre redução direta e indireta na zona inferior, chamada por Michard e Rist, de zona de elaboração, e define o consumo térmico do processo. A reta operacional AB retrata o desempenho termo-químico do processo.

No rastro do diagrama de Rist, surgiu o diagrama C-TRD ou R-TDR, representando também em plano cartesiano a relação entre taxa de redução direta na zona de elaboração e o consumo de carbono ou de redutor (C+H<sub>2</sub>) (Melgaço et alli, 1982; Chaigneau et alli, 1997).

### **Modelos fluidodinâmicos**

Os modelos fluidodinâmicos foram os que menos penetraram no meio industrial. Muitos são baseados na equação original de Ergun (1952) e Tek (1957).

$$-\Delta P/\Delta z = 150 (1 - \varepsilon)^2 \mu / g_c (\Phi d_p)^2 \varepsilon^3 \rho_g G_g^2 + 1,75 (1 - \varepsilon) / g_c (\Phi d_p) \varepsilon^3 \rho_g G_g^2 \quad (2.1)$$

em que:

$P$	pressão total dos gases
$z$	coordenada vertical
$\varepsilon$	índice de vazios
$\mu_g$	viscosidade dos gases
$g_c$	fator gravitacional
$\Phi$	fator de forma
$d_p$	tamanho das partículas
$\rho_g$	densidade da fase gasosa
$G_g$	velocidade modular mássica dos gases

Ajustada para o sistema de coordenadas cilíndricas a equação é aplicada ao espaço interno de um alto-forno. A equação pode ser bastante simplificada adotando a forma que tecnólogos japoneses introduziram na cultura de alto-fornistas do Japão e de muitos outros países:

$$V_{gv}^{1,7} = K (P_s^2 - P_t^2) \varepsilon_g^{0,3} \mu_g^{0,7} \quad (2.2)$$

em que:

$V_{gv}$	vazão de gás de ventaneiras
$P_s$	pressão absoluta de sopro
$P_t$	pressão absoluta de topo
$K$	resistência fluidodinâmica do forno
$\varepsilon_g, \mu_g$	mesmo significado a expressão anterior

Na maioria dos casos os dois últimos fatores não são considerados e ficam embutidos na constante K, como aparece na Figura 2.7.

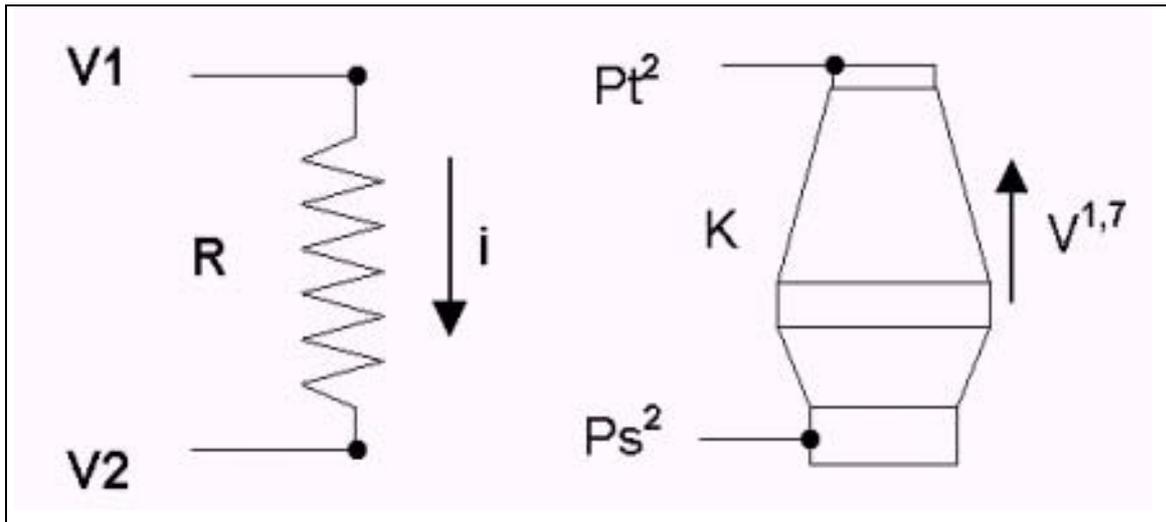


Figura 2.7 Similaridade entre um circuito elétrico e a fluidodinâmica do alto-forno

Os métodos descritos por Szekely (1979) e por Patankar (1979) são comumente referidos na literatura. A sua aplicação exige o conhecimento de diâmetros de partícula, fatores de forma, índices de vazios e temperaturas nas diferentes camadas e zonas no interior do forno. Uma vez que essas variáveis não são conhecidas, mas estimadas a partir de características dos materiais medidas fora do reator, antes de sua degradação física e química, esse tipo de modelo tem tido pouco uso. É justamente a identificação dos parâmetros-problema, entre os já citados, o que o modelo deve fazer para explicar qualquer disfunção do processo.

A região de maior impacto sobre o campo de velocidade e distribuição de pressão é a *zona coesiva*, também chamada de *zona de amolecimento e fusão*. A sua posição e forma geométricas dentro do reator dependem fortemente da composição mineralógica dos elementos da carga e de sua interação com o campo interno de temperatura. As tentativas de modelamento do escoamento gasoso nesta zona têm tido pouco sucesso, embora haja na literatura muitas referências à caracterização dos fenômenos que ocorrem ali. Destacam-se os trabalhos de Sugiyama et alii (1980) e de Barnaba (1982).

## **Métodos de avaliação integrada**

Em 1970 Yagi e colaboradores apresentaram um modelo de fluxo potencial para o escoamento de sólidos e pela primeira vez se utilizou o método de elementos finitos para a solução das equações de transporte. Trabalhos anteriores utilizaram o método de diferenças finitas, entre estes, os de Lahiri e Seshadri (1969), de Radestock e Jeschar (1970), e de Kuwubara e Muchi (1976).

Em um trabalho posterior Yagi (1993) apresentou uma revisão do modelo anterior, agora chamado de modelo dos quatro fluidos, isto é: gás, partículas grossas, partículas finas e líquidos. De Castro (2000) apresentou tese de doutoramento na Universidade de Tohoku, Japão, adotando esta técnica.

Às equações de continuidade são acrescentadas parcelas relativas a cada um dos fluidos gerados por reações químicas, fusão de sólidos ou degradação física. A cinética das reações e sua termicidade ocorrem no modelo através dessas equações.

Um modelo desenvolvido na Nippon Steel Corporation tem nome comercial: BRIGHT. Na verdade é uma espécie de gerenciador de dados obtidos por outros modelos como os já descritos anteriormente neste trabalho. A função-objetivo da sua solução está associada à convergência das soluções dos modelos para a forma da zona coesiva.

Uma série de trabalhos surgidos na década passada aplica ferramentas matemáticas e técnicas experimentais concentrando-se, porém, em alguns sub-processos específicos. Entre esses estão os trabalhos sobre a segregação da carga nas zonas superiores dos altos-fornos de Cheng e alli (1993) e de Steeghs (1997). A pesquisa mais uma vez demonstra a complexidade do processo e magnitude do número de variáveis influentes.

## 2.4 REDES NEURAIS ARTIFICIAIS (RNA)

### Aplicação de Redes Neurais Artificiais ao problema presente

A análise crítica dos modelos de processo realizada no item anterior evidencia que a maior dificuldade na simulação numérica do processo reside no elevado número de variáveis incidentes. Além disso, raramente o equilíbrio químico é alcançado nos sub-processos. A eficiência das trocas térmicas e as perdas de calor também variam significativamente. O problema aponta para uma técnica que seja capaz de estimar os desvios químicos e térmicos a partir de um grande número de variáveis de processo ajustadas pelos operadores ou determinadas nos laboratórios da usina e da mina. A eficácia de uma rede neural artificial para realizar esta tarefa é prevista pelo sucesso obtido em uma série de aplicações e será verificada neste caso nos capítulos de se seguem. (Agrypoulos, 1990; Fileti, A 1995; da Cunha, 2001)

### Evolução histórica dos conceitos em RNA

A *inteligência artificial* busca solucionar problemas através da reprodução do comportamento do cérebro. Foi demonstrado que muitas tarefas podem ser realizadas com maior eficiência através de técnicas tais como:

- Redes Neurais Artificiais, inspiradas nos princípios de processamento paralelo e distribuído neurofisiológico e tendo como principal característica o aprendizado através de exemplos e a capacidade de generalização.
- Algoritmos genéticos, onde possíveis soluções de um problema são combinadas e alteradas por mecanismos inspirados na seleção natural como as mutações.
- Sistemas especialistas que organizam, avaliam e encadeiam regras a serem adotadas na solução de problemas;
- Sistemas híbridos que combinam diversas técnicas de *inteligência artificial*.

No cérebro animal cada neurônio possui um corpo celular, ou soma, dentro do qual encontra-se a maioria dos seus orgânulos. Cada neurônio possui dendritos e, normalmente, um axônio. Através dos dendritos, o neurônio recebe impulsos elétricos, processa-os no soma, e transmite novo impulso através de seu axônio para os dendritos de outros neurônios.

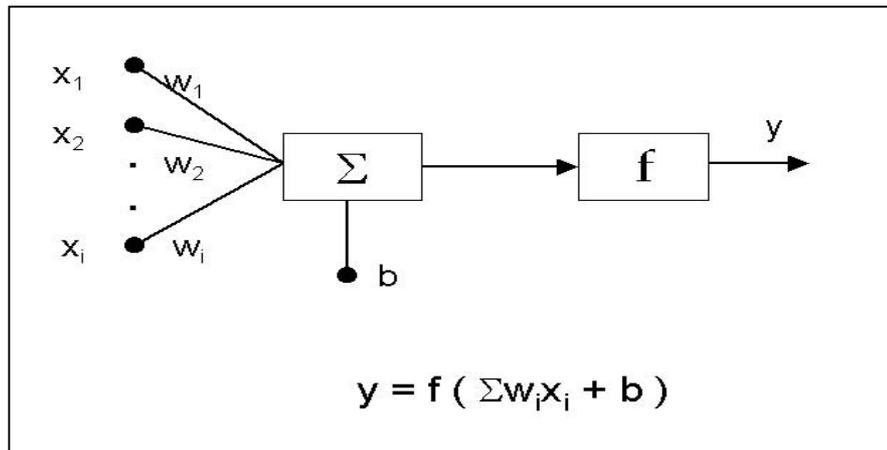


Figura 2.8 – Estrutura de um neurônio artificial

O neurônio artificial é a unidade fundamental da rede neural. A Figura 2.8 mostra a estrutura matemática de um neurônio artificial. As sinapses conectam as variáveis de entrada,  $x_i$ , ao integrador,  $\Sigma$ , através dos respectivos pesos  $w_i$ . Um *parâmetro escalar, ou bias*,  $b$ , é também adicionado ao somatório  $x_i \times w_i$ , e o valor integrado é transformado na saída do neurônio após ser modificado por uma função de ativação  $f$ .

A função de ativação pode ser uma função degrau, ou seja,  $y = 0$  se  $(\Sigma w_i x_i + b) < 0$  ou  $y = 1$  se  $(\Sigma w_i x_i + b) \geq 0$ . Esta é a função de ativação típica de um *perceptron*. A função pode ser ainda linear,  $y = (\Sigma w_i x_i + b)$ , ou sigmoideal. A função sigmoideal exponencial corresponde à equação abaixo:

$$y_i = 1 / ( 1 + \exp ( - (\Sigma w_i x_i + b) ) ) \quad (2.3)$$

Uma propriedade importante desta função é a chamada saturação, que pode ser observada na Figura 2.9. O domínio da função é intervalo real (0;1) e valores muito altos ou muito baixos do argumento não alteram a função. Assim variações intermediárias das variáveis de entrada é que influem efetivamente as variáveis de saída. Nos modelos lineares dá-se o contrário: uns poucos valores extremos podem definir o coeficiente angular subestimando a informação contida em valores intermediários que estão presentes em maior proporção.

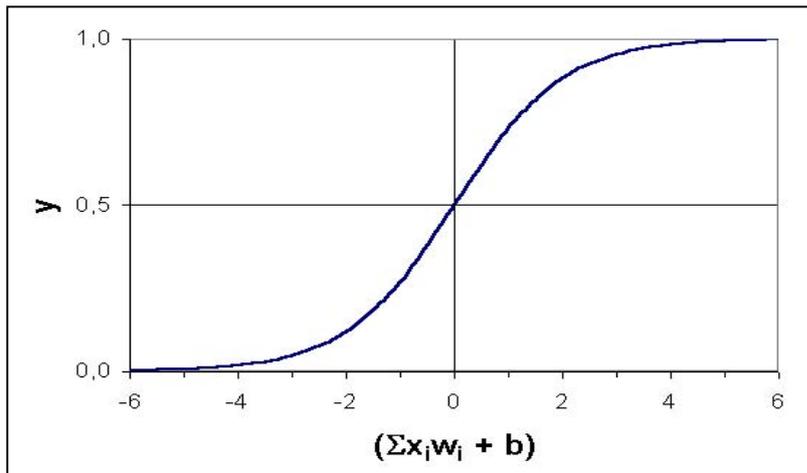


Figura 2.9 – Aspecto de uma função de ativação sigmoide logarítmica.

A estrutura de uma rede neural artificial está calcada na organização dos neurônios em camadas e na conexão entre um axônio de um neurônio e um dendrito de outro neurônio por uma sinapse, responsável pela comunicação entre estes neurônios. A Figura 2.10 mostra a estrutura de uma rede com uma camada de entrada, uma intermediária e uma de saída.

As Redes Neurais Artificiais são baseadas no modelo matemático simplificado de um neurônio, como formulado por McCulloch & Pitts (1943). O neurônio artificial possuía apenas uma saída na forma de uma função degrau. Aqueles pesquisadores criaram redes interligando os neurônios e demonstraram que estes eram capazes de executar ações complexas.

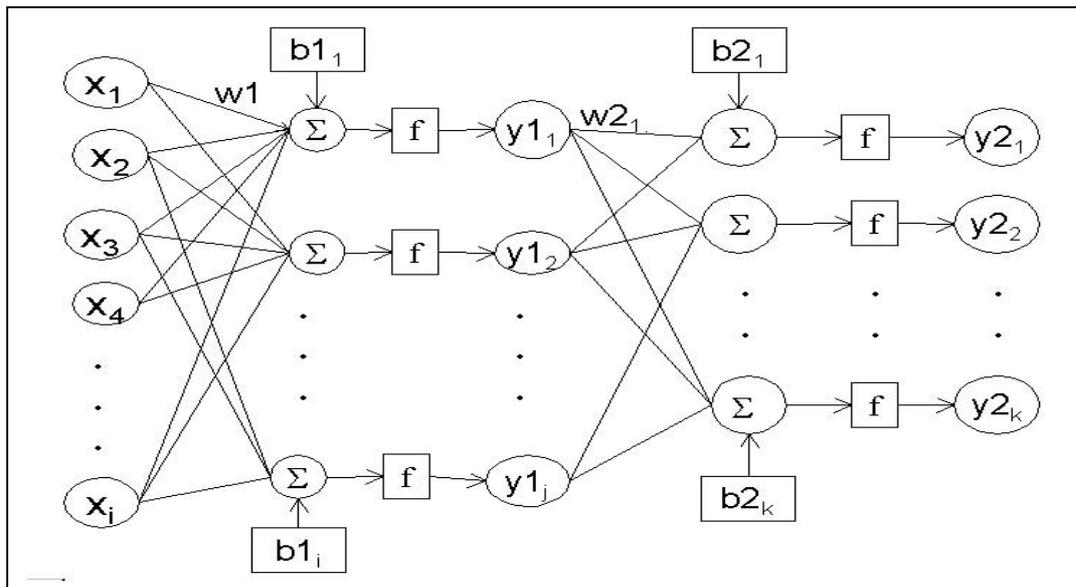


Figura 2.10 – Aspecto de uma rede neural com uma camada intermediária

As técnicas de treinamento surgiram em 1949 com Hebb que sugeriu que a alteração da eficiência sináptica seria a base do aprendizado. Assim estabeleceu o postulado: “Quando uma célula A está suficientemente próxima para excitar uma célula B e repetida e persistentemente toma parte no disparo desta, algum processo de crescimento ou mudança metabólica ocorre em uma ou ambas as células de modo que a eficiência de A em excitar B é aumentada”. Portanto, segundo Hebb, uma determinada conexão somente é reforçada se tanto as células “pré-sinápticas” quanto as “pós-sinápticas” estiverem excitadas.

Em 1958 Rosenblatt apresentou o modelo dos *Perceptrons*. Nele, os neurônios estão organizados em duas camadas, uma de entrada e outra de saída, totalmente conectadas entre si. Estas conexões são direcionadas da camada de entrada para a camada de saída. As saídas dos neurônios desta rede produziam zero ou um, conforme a classificação dada ao valor calculado; saída da RNA maior ou igual a zero classificada como um e saída menor que zero classificada como zero. Baseado no treinamento de Hebb, Rosenblatt conseguiu criar um algoritmo que alterava os pesos entre os neurônios de modo a realizar

a computação desejada, que foi denominada algoritmo de aprendizagem ou de treinamento.

Na mesma época que Rosenblatt trabalhava no *Perceptron*, em 1962, Widrow criou a *Adaline* (*Adaptative linear neuron*), um elemento processador linear cuja saída é a soma ponderada das entradas pelos respectivos pesos. Mais tarde Widrow criou o *Madaline* (*Multilayer adaline*), uma arquitetura com duas camadas de neurônios. O modelo *Adaline* é similar ao *Perceptron*, mas os elementos processadores variam entre  $-1$  e  $+1$ .

### **Algoritmo de treinamento**

A etapa de treinamento é uma das mais importantes na formulação de um modelo com base em RNA. Há basicamente dois tipos de treinamento:

- Autônomo, auto-organizado ou não-supervisionado. Neste não há intervenção de um supervisor e nem há um vetor projetado de saída. O algoritmo de treinamento necessita apenas vetores de entrada, modificando os pesos para produzir vetores de saída consistentes. O processo de treinamento extrai propriedades estatísticas do conjunto treinado e grupo de vetores similares entre classes.
- Supervisionado, associativo ou por correção do erro. Neste caso há um supervisor que avalia o comportamento do sistema e providencia as modificações subseqüentes. Pares são formados de cada vetor de entrada com um vetor alvo de saída. Juntos são chamados “par treinado”. Um vetor de entrada é aplicado, a saída da rede é calculada e comparada ao vetor alvo. O erro adquirido é retropropagado através da rede e os pesos são mudados conforme o algoritmo de treinamento escolhido que tende a minimizar o erro. Os vetores do treinamento são aplicados seqüencialmente, os erros são calculados para cada par treinado e os pesos ajustados para cada vetor até que o erro de toda a rede atinja um valor aceitável.

A maior contribuição de Widrow foi a invenção da “Regra Delta” ou “Método do Gradiente”, para treinamento das redes com processadores *adaline*, que se baseavam em um processo de iteração local para obtenção de um ponto mínimo utilizando um exemplo do conjunto de treinamento por vez.

O treinamento por gradiente descendente é a técnica mais utilizado para agilizar o treinamento e o algoritmo denominado “regra de delta generalizada” (RDG) O método iterativo por gradiente, que minimiza o quadrado do erro, se utiliza de um peso extra atribuído aos pesos a serem ajustados. A RDG calcula o erro de saída e gera os novos valores para os pesos, propagando de volta para as camadas anteriores.

A RDG utiliza a função *bias* ao invés de valores internos residuais de ativação (*threshold*). São adicionadas e não subtraídas, como nos valores *threshold*, ao somatório do neurônio e não devem ser atualizados no treinamento. Na RDG, estas funções devem assumir o valor *zero*, na primeira camada, e *um* nas seguintes.

A RNA tipo *backpropagation* é um dos algoritmos de treinamento mais conhecidos e utilizados. Pode ser aplicada a topologias com qualquer número de camadas em conexões *feedforward* ou *feedback*. É uma generalização da RDG.

### **Algoritmo Levenberg-Marquardt com regularização bayesiana**

A regularização bayesiana minimiza uma combinação linear de erros e pode ser aplicada a parametrização de redes neurais artificiais como descrito por MacKay (1992) e por Foresee & Hagan (1997).

Esse conceito pode ser utilizado dentro de um algoritmo Levenberg-Marquardt, um método de aproximação matricial não-linear que será descrito a seguir (Cunha, 2003).

Seja  $y(x; a_1, a_2, \dots, a_n)$  a função da variável  $x$  e dos parâmetros  $a_i$  que será usada no conjunto de dados  $y_1, y_2, \dots, y_m$ , sendo  $m \gg n$ . Se  $\mathbf{a}$  for um vetor com os parâmetros de ajuste, o resíduo do ponto  $x_i$  é:

$$f_i(\mathbf{a}) = y_i - y(x_i, a_1, a_2, \dots, a_n) \text{ para } i = 1 : m \quad (2.4)$$

A função a minimizar é:

$$F(\mathbf{a}) = \frac{1}{2} \sum f_i^2(\mathbf{a}) \text{ para } i = 1 : m \quad (2.5)$$

O fator  $\frac{1}{2}$  visa facilitar desenvolvimentos posteriores.

Os três primeiros termos da série de Taylor em torno de  $\mathbf{a}$  associado a um vetor incremento  $\mathbf{p} = (p_1, p_2, \dots, p_m)$ , portanto  $F(a_1+p_1, \dots, a_n+p_n)$  são:

$$F(\mathbf{a}) + p_1 \frac{\partial F}{\partial a_1} + \dots + p_n \frac{\partial F}{\partial a_n} + \frac{1}{2} (p_1^2 \frac{\partial^2 F}{\partial a_1^2} + \dots + p_n^2 \frac{\partial^2 F}{\partial a_n^2}) \quad (2.6)$$

As derivadas de primeira ordem desta função são:

$$\frac{\partial F}{\partial a_i}(\mathbf{a}) = \sum f_i(\mathbf{a}) \frac{\partial f_i}{\partial a_i}(\mathbf{a}) \text{ para } i = 1 : m \quad (2.7)$$

Em notação matricial as derivadas das componentes de  $f(\mathbf{a})$ , calculada em  $\mathbf{a}$  formarão a matriz jacobiana  $J(\mathbf{a})$  :

$$J(\mathbf{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial a_1}(\mathbf{a}) & \frac{\partial f_1}{\partial a_2}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial a_m}(\mathbf{a}) \\ \frac{\partial f_2}{\partial a_1}(\mathbf{a}) & \frac{\partial f_2}{\partial a_2}(\mathbf{a}) & \cdots & \frac{\partial f_2}{\partial a_m}(\mathbf{a}) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial a_1}(\mathbf{a}) & \frac{\partial f_m}{\partial a_2}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial a_m}(\mathbf{a}) \end{pmatrix} \quad (2.8)$$

O produto da matriz jacobiana transposta,  $J^T$ , pelo vetor  $\mathbf{f}$ , é uma matriz cujas componentes correspondem às expressões (2.7).

Os termos de primeira ordem da série de Taylor (2.6) podem ser escritos na notação matricial:

$$p_1 \frac{\partial F}{\partial a_1}(\mathbf{a}) + \dots + p_m \frac{\partial F}{\partial a_m}(\mathbf{a}) = \mathbf{p} J^T(\mathbf{a}) \mathbf{f}(\mathbf{a}) \quad (2.9)$$

O resultado dessas operações é um número real que será somado a  $F(\mathbf{a})$ .

Para os termos de segunda ordem, pode-se usar a regra do produto para derivar (2.5) gerando:

$$\begin{aligned} \frac{\partial^2 F}{\partial a_i \partial a_j} = \sum \frac{\partial f_k}{\partial a_i}(\mathbf{a}) \cdot \frac{\partial f_k}{\partial a_j}(\mathbf{a}) + \\ + \sum f_k(\mathbf{a}) \cdot \frac{\partial^2 f_k}{(\partial a_j \partial a_k)}(\mathbf{a}) \quad \text{para } i, j, k = 1 : m \end{aligned} \quad (2.10)$$

Os somatórios correspondentes às primeiras parcelas de (2.8) são elementos da matriz simétrica  $J^T(\mathbf{a}) J(\mathbf{a})$ . Os produtos pelos incrementos  $p_i$  levam a :

$$\mathbf{p}^T J^T(\mathbf{a}) J(\mathbf{a}) \mathbf{p} \quad (2.11)$$

Os segundos somatórios são derivadas de segunda ordem que irão compor uma matriz  $Q(\mathbf{a})$ . Os últimos termos de (2.6) podem ser re-escritos :

$$\mathbf{p}^T (J^T(\mathbf{a})J(\mathbf{a})\mathbf{p} + Q(\mathbf{a})\mathbf{p}) \quad (2.12)$$

Na notação matricial a série de Taylor será:

$$F(\mathbf{a} + \mathbf{p}) \approx F(\mathbf{a}) + \mathbf{p}J^T f(\mathbf{a}) + \frac{1}{2}\mathbf{p}^T (J^T(\mathbf{a})J(\mathbf{a}) + Q(\mathbf{a}))\mathbf{p} \quad (2.13)$$

que é uma função quadrática dos incrementos  $p_i$ .

O passo seguinte do processo consiste em encontrar um incremento que diminua a soma dos resíduos, ou seja, encontrar  $\mathbf{p}^k$  tal que:

$$F(\mathbf{a}^k + \mathbf{p}^k) < F(\mathbf{a}^k) \quad (2.14)$$

em cada passo de  $k$ . Se isso for possível  $\mathbf{a}^k + \mathbf{p}^k$  estará mais próximo do mínimo do que  $\mathbf{a}^k$ . Para isso, em cada passo de  $k$  deve-se resolver o sistema linear:

$$(J_k^T J_k + Q_k) \mathbf{p}_k = -J_k^T f_k \quad (2.15)$$

Resolvido o sistema a nova aproximação será  $\mathbf{a}^{k+1} = \mathbf{a}^k + \mathbf{p}$ . A matriz  $\mathbf{p}$  envolve os erros de pesos e *bias* da rede.

Diversos métodos existem para resolver o sistema entre eles o de Gauss-Newton, usado por Foresee & Hagan (1997).

O método Levenberg-Marquardt será usado neste trabalho através de vantagem do pacote computacional *Matlab* (R).

No método a matriz  $Q_k$  é substituída por  $\lambda_k I$ , onde  $\lambda_k$  é um número real não negativo e  $I$  é a matriz identidade  $n \times n$ .

No *Matlab(R)* a expressão assume a forma:

$$D = - (J_k^T J_k + \lambda_k I) \setminus J_k p_k \quad (2.16)$$

em passos sucessivos. O operador  $\setminus$  no *Matlab(R)* pode ser entendido da seguinte maneira: Se  $A$  é uma matriz quadrada,  $A \setminus B$  é aproximadamente igual a  $inv(A) \times B$ , mas computado de uma forma diferente. Se  $A$  é uma matriz  $n \times n$  e  $B$  é um vetor coluna tamanho  $n$  ou uma matriz com várias colunas tamanho  $n$ , então  $X = A \setminus B$  é a solução da equação  $A X = B$ .

### **Aspectos Avançados em RNA**

Dificuldades durante o processo de treinamento, tais como: mínimos locais, paralisia da rede ou instabilidade temporal, levaram muitos pesquisadores a investigar as propriedades das RNA e buscarem modificações no algoritmo de treinamento para redes *backpropagation*. (Koacs, 1996)

Abelém e colaboradores (1995) citam fatores que determinam o desempenho das redes artificiais como: convergência; generalização e a escalabilidade entre os dois.

A convergência é função do aprendizado das redes e indica quando o procedimento de treinamento é capaz de aprender a classificação dos vetores de entrada e saída apresentados. Quais condições são necessárias na rede e quais condições computacionais são necessárias. Alguns critérios são adotados para se avaliar o processo de treinamento: número de iterações, erro mínimo aceitável de toda a rede ou por padrão treinado ou ainda tempo de treinamento.

A generalização é a habilidade da RNA de reconhecer os padrões fora do conjunto treinado e é a principal propriedade que deve ser exigida em uma RNA.

A escalabilidade é a propriedade de se ajustar os parâmetros do sistema neural projetado para se obter boa convergência e se extrair a generalização desejável. É o ponto ótimo entre uma grande rede que aprende com facilidade e uma pequena rede que generaliza com sucesso.

Esses fatores são determinados pela definição dos parâmetros da RNA:

- Função de ativação: as mais usadas são as funções sigmoideal e tangente hiperbólica, com domínios entre  $0$  e  $1$  e entre  $-1$  e  $1$ , respectivamente. A utilização destas funções tem a ver com o fato de saturarem nos extremos no contra-domínio, como já discutido;
- Função erro: informa o grau de convergência das RNA e o quanto a rede está assimilando os padrões em treinamento.
- Parâmetros de controle do gradiente e do tempo de treinamento:
  - Momentum: utilizado para diminuir o tempo de treinamento e o risco de mínimos locais, sem adicionar oscilações ao processo de treinamento;
  - Razão de aprendizado: influencia a intensidade da mudança dos pesos. Valores muito altos podem introduzir oscilações ou saturação indesejadas ao treinamento.
- Condições de inicialização do treinamento (Atiya, 1997);
- Arquitetura da rede: o tamanho e a configuração influenciam a escalabilidade da RNA, pois pode-se aumentar ou diminuir a capacidade de treinamento e generalização pela variação do número de camadas ocultas, de neurônios, etc.

## Arquitetura da rede

O número de camadas ocultas em redes neurais artificiais é um ponto importante a ser definido e foi estudado por Villiers e Barnard (1992). Baseado na idéia do número de pesos ser um indicador aproximado da complexidade da RNA, os autores compararam as redes com três e quatro camadas com o mesmo número de pesos, para dimensões de entrada da RNA variando de pequenas,  $D = 2$ , a grandes,  $D = 21$ . As redes com quatro camadas que apresentaram melhores desempenhos na qualidade de generalização foram as ajustadas com o mesmo número de neurônios das duas camadas ocultas. Resultados experimentais indicaram que as redes neurais com quatro camadas são mais propensas a caírem em um mínimo local e que por isto as redes com três camadas apresentam um desempenho melhor tanto no treinamento quanto na generalização. Entretanto, não foi encontrada diferença entre o desempenho ótimo, melhor rede ajustada, entre redes com três e quatro camadas. Outra observação importante: não se verificou influência no desempenho das redes de acordo com a mudança do número de entradas.

Na mesma linha do tratamento anterior, um artigo publicado por Tamura e Tateishi (1997) deu ênfase à comparação entre a capacidade de redes neurais artificiais de três e de quatro camadas com um número finito de unidades ocultas. Os autores concluem neste trabalho que a RNA *feedforward* de três camadas com  $n - 1$  neurônios ocultos pode dar qualquer relação exata entre  $n$  relações entrada-saída e que a RNA com quatro camadas dará a mesma relação entre entrada-saída usando somente  $n/2 + 3$  unidades ocultas. Desta forma, a RNA com quatro camadas é superior em termos de número de parâmetros necessários para o treinamento dos dados e ainda está associada a erros menores durante o treinamento.

Ainda sobre a definição da arquitetura da RNA, outro ponto importante é a definição do número de neurônios das camadas ocultas. Hecht-Nielsen (1990) concluiu que uma rede neural com três camadas, sendo uma oculta, pode modelar qualquer função matemática contínua através de uma rede com  $2n+1$  neurônios ocultos, sendo  $n$  o número de neurônios da camada de entrada. Marem (1990), por sua vez, concluiu que redes com quatro

camadas, duas ocultas, apresentam melhor desempenho quando são treinadas com uma proporção de 3:1 entre o número de neurônios da primeira camada oculta e o número de neurônios da segunda camada oculta. Já Lippmann (1987) anunciou que para redes com quatro camadas a segunda oculta deverá conter o dobro do número de neurônios da camada de saída e que para redes com três camadas, a camada oculta deverá conter  $s(n+1)$  onde  $s$  é o número de neurônios da camada de saída e  $n$  o número de neurônios da camada de entrada.

Mirchandani e Cao (1989) estudaram a determinação do número ótimo de neurônios ocultos em RNA de três camadas, considerando um conjunto de vetores de entrada em um espaço Euclidiano  $R^d$ . Um neurônio oculto age como um hiperplano de dimensão  $d - 1$  que forma duas regiões de decisão. Foi definido que o espaço de entrada de dimensão  $d$  é linearmente separável em  $m$  regiões, se elas existem em  $m$  regiões distintas limitadas compondo porções do hiperplano. A associação das regiões com classes é a tarefa dos neurônios de saída conectados aos neurônios ocultos. O número de regiões separáveis identifica o número mínimo de padrões de entrada  $t$ , necessários para o treinamento do *perceptron*. A situação ideal seria fazer  $t = m$ , ou seja, o emprego de cada padrão de treinamento gerando uma região separável para assegurar a separação do espaço de entrada em  $m$  regiões necessárias. Na prática,  $t > m$ , pois o algoritmo *backpropagation* não assegura a convergência a um mínimo global.

Os autores criaram na ocasião um teorema: No espaço  $d$ , o número máximo de regiões linearmente separáveis,  $m$ , usando  $h$  neurônios ocultos, seria dado por:

$$m(h,d) = \sum_{k=0}^d C_{h,k} \quad (2.17)$$

sendo que para  $d = 1$ ,  $m = h + 1$  e para  $h < d$ ,  $m = 2^h$

Por conseqüência, tendo dois dos três parâmetros,  $m$ ,  $h$  e  $d$ , pode-se determinar um mínimo de padrões de treinamento necessários para o sucesso da RNA.

Hagiwara (1992) publicou um artigo sobre a seleção/redução do tempo de treinamento e eliminação de unidades ocultas. Sabe-se que redes que usam *backpropagation* possuem preocupações de:

- encurtar o tempo de treinamento através do uso de uma função-erro íngreme, da mudança da forma de apresentação das padrões, do emprego do erro total durante o treinamento e da variação dos parâmetros  $\alpha$  e  $\eta$ ;
- reduzir o número de neurônios ocultos, pois resultaria em um aumento na capacidade de generalização, em uma interpretação facilitada da RNA e na redução dos custos computacionais.

Foi proposto introduzir a capacidade de seleção para as unidades ocultas, para a redução do tempo de treinamento e das piores unidades ocultas:

- priorizando a convergência pelo ajustes dos pesos conectados aos neurônios selecionados como piores para pequenos valores aleatórios, acelerando o treinamento e escapando do mínimo local;
- eliminando os piores neurônios selecionados quando a convergência é atingida e reduzindo a número de unidades ocultas.

Para a camada oculta  $k - 1$ , com neurônios  $i$  e camada de saída  $k$ , com neurônios  $j$ , indicou-se o fator má-qualidade,  $Q$ , para representar a má qualidade dos neurônios ocultos selecionados, em função do erro e do neurônio oculto  $i$ , para o padrão  $p$ :

$$Q_{k-1} = \sum (e_i^{k-1,p})^2 = \sum (\sum w_{ij}^{k-1} \delta_j^k)^2 \quad (2.18)$$

Assim, se o fator  $Q$  é pequeno, a correção para aquela unidade de processamento também é pequena. Logo,  $Q$  indica a condição de convergência dos pesos conectados a cada neurônio da camada oculta. Desta forma, o neurônio oculto com o maior  $Q$  é considerado o pior, pois está longe de convergir.

Baseado na seleção dos neurônios pelo fator  $Q$ , foi desenvolvido um método para escapar do mínimo local e reduzir o tempo de aprendizado. Ao se cair em um mínimo local é necessário inserir algum tipo de distúrbio:

- monitorando o erro total na camada de saída;
- selecionando as unidades piores quando a mudança no erro total na camada de saída tornar-se muito pequena;
- reajustando os pesos das unidades selecionadas para pequenos pesos aleatórios;
- aumentando o número de camadas ocultas, se a rede não convergir após um certo número de iterações.

Para os casos em que a rede atinge a convergência, pode-se reduzir o número de unidades ocultas.

### **Inicialização do treinamento**

No campo das condições iniciais do treinamento e a generalização, Atiya e Ji (1997) deram sua contribuição em um estudo sobre o desempenho de redes grandes. A generalização depende de três fatores:

- graus de liberdade da rede ;
- quantidade de ruído no conjunto de treinamento;
- tamanho do conjunto de treinamento.

Embora muitos pesquisadores tenham concentrado esforços no estudo do efeito do tamanho da rede na generalização, os autores optaram pela investigação da distribuição inicial dos pesos. Quando um algoritmo de gradiente descendente é

usado, ele tende a convergir para o mínimo local das funções-objetivo encerradas para as condições iniciais. No caso de uma rede sobreparametrizada, o mínimo da função corresponde não a um ponto, mas a uma superfície da solução (como quando o erro tende a zero). O algoritmo de treinamento geralmente aterrará os pesos em um ponto da superfície de solução fechada às condições iniciais. Se os pesos iniciais são pequenos, então a tendência é que os pesos finais alcançados sejam pequenos. Isto pode resultar em muita suavidade a ser justificada pela complexidade do problema dado ou pela quantidade de ruído. (Wann et alli, 1989)

Os estudos empíricos dos autores resultaram no seguinte:

- Existe uma faixa ótima inicial de pesos que conduzem a uma melhor generalização. Isto contradiz a crença de que pesos iniciais pequenos geram uma melhor generalização. (Lehmen et alli, 1998)
- O efeito na generalização é mais forte, mas não é limitado à RNA sobreparametrizada;
- Visto que a condição inicial seja somente um dos fatores que afetam os algoritmos de treinamento de gradiente descendente, é natural esperar que ela somente afete o desempenho da generalização de certas redes, mas não de todas.

Três simulações com distribuição de pesos iniciais entre  $-0,05$  e  $0,05$ ,  $-0,5$  e  $0,5$  e  $-10,0$  e  $10,0$  mostraram que os pesos iniciais geraram a pior generalização e o melhor caso ocorreu com pesos iniciais entre  $-0,05$  e  $0,5$ .

### **Critério de finalização do treinamento**

Schmidt e colaboradores (1993) provaram, através de muitas simulações computacionais de treinamento usando diferentes pesos iniciais, que a regra para se selecionar uma rede dita ótima através do conceito de menor erro quadrático atingido pelo treinamento não é válida. A primeira observação importante nas simulações foi o fato de que, ao aumentar-se o número de iterações que regem o

processo de treinamento, o erro quadrático torna-se cada vez menor. Isto, porém, não significa que o ótimo foi encontrado, pois durante o processo de treinamento algumas regiões com erros pequenos apresentaram dois ou mais mínimos locais.

Outras 1000 simulações foram realizadas considerando números fixos de iterações e mesmo banco de dados, diferindo somente nos pesos iniciais. Na seleção das melhores redes, a que apresentou o menor erro quadrático durante o treinamento não resultou na melhor generalização. A maior proporção de acerto na generalização, principal propriedade da RNA, concordou com um resultado de erro quadrático um pouco maior. Desta forma, na hora da escolha da rede ótima o que define é a obtenção da melhor generalização.

## **2.5 APLICAÇÕES DAS RNA NA SIDERURGIA**

Nesta pesquisa, além da biblioteca da própria UNICAMP foram acessadas as bibliotecas da COPPE/UFRJ e da Escola de Engenharia da UFMG. Em todas as bibliotecas foram levantados os temas de tese de mestrado e doutorado na área de metalurgia. Também foram analisados artigos afins de periódicos dedicados à metalurgia e aplicações de técnicas de *inteligência artificial*.

Na área de engenharia as aplicações são raras e recentes no Brasil. Na Universidade Federal de Minas Gerais, de 1993 a 1996, 16 teses de doutorado e 39 teses de mestrado foram apresentadas. Em nenhuma delas foram encontradas aplicações de tratamentos de dados de processos siderúrgicos por qualquer técnica de *inteligência artificial*.

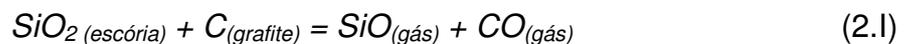
Aplicações na indústria siderúrgica são encontradas na área de aciaria, lingotamento contínuo e laminação. (Ping et alli, 2003)

Da Cunha, em 2001, apresentou tese de doutorado na Universidade de Campinas versando sobre o controle do processo de refino dos aços e na predição de características metalúrgicas do sinter de minério de ferro.

Uma interessante tese de Mestrado em Engenharia Elétrica e Ciência Computacional apresentada no Massachusetts Institute of Technology, Boston e um artigo publicado um ano depois relatam aplicações de inteligência artificial em altos-fornos. Ambos descrevem um modelo baseado em uma rede neural multi-camada visando a previsão da temperatura do gusa, o teor de silício do gusa e a basicidade da escória. (Ge, 1999; Radhakrishnan & Mohamed, 2000)

De fato, a temperatura do gusa é um dos parâmetros de processo visados pelos operadores. Obviamente a manutenção deste parâmetro dentro dos limites acordados entre os gestores do alto-forno e da aciaria indica o adequado controle térmico do processo. Neste resultado influem quase todos os parâmetros manipuláveis pela operação e também características físicas e químicas das matérias primas e insumos.

O teor de silício e a temperatura do gusa estão intimamente ligados. O silício se incorpora ao gusa na chamada *zona de gotejamento*, após a sílica ter sido reduzida a SiO gasoso. Às temperaturas reinantes nesta zona, 1800 a 2200 K, só os poros do coque, ou de outro redutor sólido presente, podem fornecer o potencial de O<sub>2</sub> suficientemente baixo para esta reação. O cálculo do valor da velocidade para a reação:



foi demonstrado por diversos autores que chegaram à equação:

$$(Ws/M_{SiO_2}) (d[\%SiO_2]/dt) = k_{SiO} A_{s-g} A_{SiO_2} \quad (2.20)$$

onde  $A_{s-g}$  e  $A_{SiO_2}$  denotam a área superficial entre escória e grafita. Biwas (1987) calculou a média obtida por diversos autores e chegou aos seguintes valores para  $k_{SiO}$  :

$$k_{SiO} = 9,0 \times 10^4 \exp(-86\,000/RT) \text{ (kgmol/m}^2 \cdot \text{s.atm)} \quad (2.21)$$

se  $T < 2000\text{K}$  e

$$k_{SiO} = 2,3 \times 10^4 \exp(-180\,000/RT) \text{ (kgmol/m}^2 \cdot \text{s.atm)} \quad (2.22)$$

se  $T > 2000\text{K}$

Em condições próximas às de equilíbrio, a velocidade da reação



é também extremamente dependente da temperatura. O processo ainda depende do tempo de residência do gusa nessas condições, que por sua vez depende da estrutura interna da carga e da taxa de produção de gusa. Esta última, além de interferir no tempo de residência, influi na área de contato entre metal e gás. Em uma infinidade de trabalhos publicados é demonstrada a forte dependência do teor de silício e temperatura de gusa para um mesmo forno. As outras variáveis que influem neste processo são mais estáveis e estão associadas a características da carga, principalmente a reatividade e teor de sílica nas cinzas do redutor sólido. Pelo dito fica clara também a influência da geometria do forno.

Dentro da zona de gotejamento, a região coronal que se apoia na zona de combustão é onde predominam as maiores temperaturas. Sendo a seção horizontal da zona de combustão uma coroa de aproximadamente, 1,5 m de largura, independentemente do diâmetro do forno, é evidente que a proporção

ocupada por esta seção na seção total do forno é maior quanto menor for o forno, como pode ser observado na Figura 2.11.

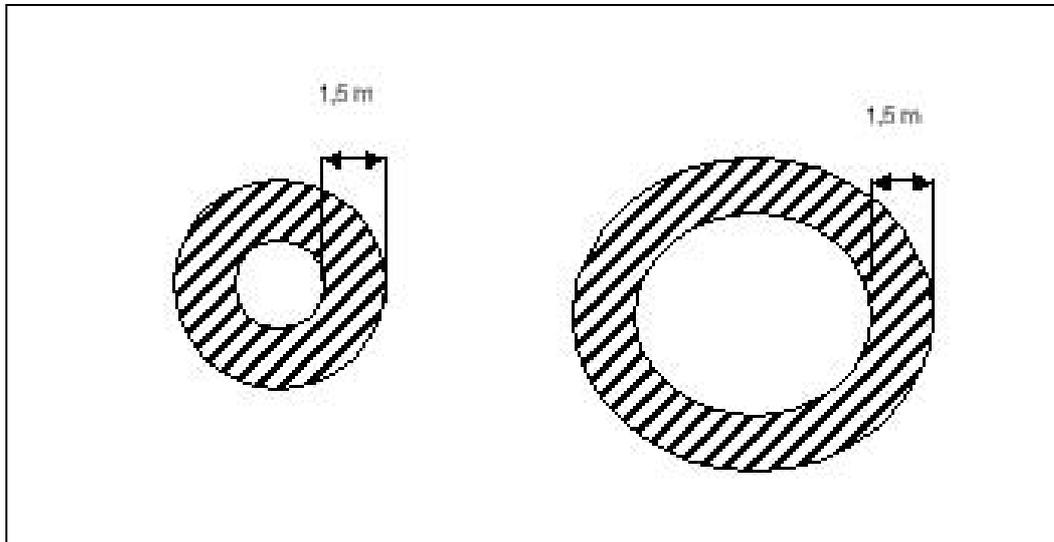


Figura 2.11 – Comparação entre a relação entre a área da seção quente sobre a área total de dois altos-fornos de diâmetros de cadinho distintos

O fato de altos-fornos produzirem gusa de teores de silício mais baixos é uma evidência da forte relação entre temperatura e este parâmetro.

A escória do alto-forno é composta majoritariamente por 4 óxidos:  $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{CaO}$  e  $\text{MgO}$ . O teor de outros constituintes é da ordem de 3% do peso total. A basicidade da escória, na sua versão binária,  $\text{CaO} / \text{SiO}_2$ , ou quaternária,  $(\text{CaO} + \text{MgO}) / (\text{SiO}_2 + \text{Al}_2\text{O}_3)$ , são variáveis-meta na operação de um alto-forno. O valor almejado é definido com vistas à adequada temperatura de *liquidus*, a viscosidade e o poder dessulfurante da escória. O poder dessulfurante é quantificado pela relação entre o teor de enxofre na escória e teor de enxofre no gusa. O valor necessário a se obter um gusa com teor de enxofre máximo aceito pela aciaria é definido com base na carga de enxofre e na taxa de geração de escoria. A existência de quatro variáveis, os teores de  $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{CaO}$  e  $\text{MgO}$ , já condicionada pela soma fixada em 97%, rendem ao sistema a três graus de liberdade. A combinação mais conveniente dessas variáveis é  $B_2, x\text{Al}_2\text{O}_3, x\text{MgO}$ ,

onde  $B_2$  é a basicidade binária e  $x_i$  o teor do óxido  $i$  na escória. Mas  $xAl_2O_3$  e  $xMgO$  dependem inteiramente da carga e de sua composição química já que esses óxidos não são reduzidos no alto-forno. O óxido de cálcio também não é reduzido, de forma que resta como variável dependente das condições de processo a basicidade binária, que por sua vez depende fundamentalmente de  $xSiO_2$ . Como a massa de  $SiO_2$  na escória é a diferença entre a massa carregada e a reduzida, esta última convertida inteiramente em silício dissolvido no gusa, tem-se a clara relação entre basicidade e temperatura do gusa.

Portanto, as três variáveis: temperatura do gusa, teor de silício e basicidade, em condições de controle de processo estão intimamente ligadas, sendo no caso das duas últimas variáveis praticamente equivalentes.

Singh e colaboradores, em 1995, também publicaram uma aplicação de redes neurais artificiais à predição do teor de silício do gusa. Obtiveram um coeficiente de correlação de 0,86 através de uma rede com arquitetura 18x3x1.

Outro trabalho nesta linha foi o de Waller e Saxen, da Universidade Abo da Finlândia (2002), que utilizaram dados dos altos-fornos da Rautaruukki e descreveram uma análise de série temporais não-lineares para previsão de silício do gusa. Foram denominados redes de função de base radial e modelos de séries temporais lineares. Os autores realçam as dificuldades de tratar dados industriais entre elas a necessidade de utilizar resultados de análises químicas de amostras que não estão realmente em série temporal regular.

Mauhart e colaboradores, em 2003, descreveram um pacote desenvolvido pela Voest-Alpine Industrialanlagebau e aplicado em altos-fornos da Áustria, na usina da própria Voest-Alpine, e da África do Sul, em um forno da Iscor. O pacote, chamado de VAlron, apesar de anunciar o uso de *inteligência artificial* e sistema especialista em circuito fechado, não envolve as ferramentas objetivadas neste estudo. A verdadeira base são os modelos determinísticos tradicionais.

Um grupo de Sidmar, Companhia siderúrgica da Bélgica e do CRM, Centre de Recherche Metalurgique, do mesmo país, desenvolveram e aplicaram um modelo de alto-forno. O principal objetivo foi o de caracterizar a geometria da carga dentro do reator. Nenhuma técnica matemática foi mencionada no artigo publicado. O modelo pode ser um módulo de sistema global. Como esperado foi observado uma forte correlação entre a altura da zona de gotejamento prevista e o teor de silício do gusa. O modelo é chamado de MOGADOR.

Zaimi e outros pesquisadores do Institute de Recherche Siderurgique, IRSID, da França (2003), aplicaram algumas técnicas de mecânica dos solos à carga no interior dos altos-fornos. Usaram técnicas de processamento em elementos finitos e teoria da hipoplasticidade. A técnica se mostrou sensível a características da carga, mas sua eficácia como ferramenta de avaliação global do processo não foi comprovada.

## **2.6 JUSTIFICATIVA DA PROPOSTA DE TESE**

Pelo que foi visto e discutido verifica-se que:

- O desenvolvimento de tecnologia em siderurgia é vital para o crescimento do Brasil, maior exportador de minério de ferro do mundo.
- Os altos-fornos são e permanecerão sendo os equipamentos dominantes no setor siderúrgico de grande porte.
- A operação de altos-fornos é normalmente constituída de ajustes em reação a alterações observadas no processo decorrentes de variações nas características dos materiais empregados, dependendo fortemente da experiência dos operadores.
- Existe na indústria siderúrgica uma demanda por uma ferramenta capaz de apoiar a operação de altos-fornos a condições de matéria prima propiciando o planejamento o processo a curto prazo.

- Modelos mais objetivos, como os de Reichardt ou Rist-Michard-Maysson tendem a tratar isoladamente os fenômenos termodinâmicos e fluido-dinâmicos.
- Os modelos fenomenológicos globais de simulação citados no item 2.3 trabalham com elevado número de relações e variáveis e exigem longos períodos de processamento.
- As redes neurais artificiais têm se mostrado eficientes na obtenção de correlações entre grande número de variáveis, mas seu potencial tem sido pouco explorado neste ramo tecnológico.

Os modelos de análise de processo relatados apresentam concepção interessante e se mostram eficientes quando aplicados *a posteriori*. Se os parâmetros de equilíbrio químico mais complexos, desvio de equilíbrio dependentes de grande número de fatores e a intensidade das perdas térmicas pudessem ser conhecidos *a priori*, a partir de variáveis de processo, esses conceitos poderiam ser combinados em um modelo de previsão de processo. A tarefa de prever esses parâmetros-chave pode ser entregue a um modelo neural, cuja vocação para esse tipo de problema foi demonstrada.

Portanto é proposta da presente tese a formulação e parametrização de um modelo híbrido, com base em uma rede neural para prever parâmetros que alimentam módulos fenomenológicos. Como resultado, obtém-se uma ferramenta de planejamento operacional capaz de avaliar padrões de processo alternativos indicando os mais adequados aos objetivos empresariais. A ferramenta torna a operação dos altos-fornos uma prática pró-ativa.

A Figura 2.12 ilustra a aplicação do modelo híbrido na previsão das condições operacionais do alto-forno definidas pelos parâmetros  $[Si]$ , teor de silício do gusa.  $[S]$  teor de enxofre do gusa,  $CR$ , *coke-rate* e  $\pi$ , produtividade.

A taxa de consumo de redutores, ou *fuel-rate* ( $FR$ ) é a soma do *coke-rate* ( $CR$ ), *pulverized coal-rate* ( $PCR$ ) e *gás-rate* ( $GR$ ). Na prática as taxas de consumo de carvão e de gás são fixadas fazendo-se o ajuste térmico através apenas da taxa de consumo de coque, o *coke-rate*. Daí apenas esta aparecer como variável de saída. As outras taxas incluem-se entre os parâmetros de entrada, no plano de carga e sopra.

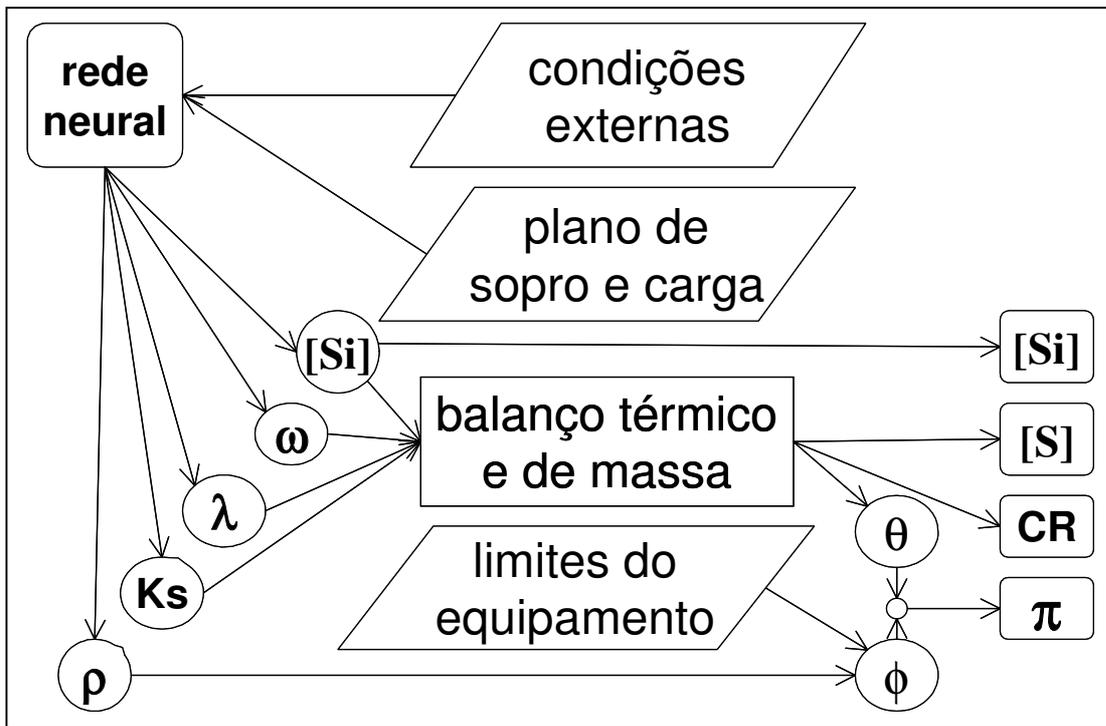


Figura 2.12 – Fluxo esquemático de processamento do modelo proposto



### CAPITULO 3 - FORMULAÇÃO DO MODELO HÍBRIDO

Ao longo das últimas décadas tornou-se muito comum na siderurgia mundial indicar a produtividade dos altos-fornos pela relação entre a massa de metal produzido por dia e o volume total do reator ou das zonas mais ativas. O modelo que se desenvolve neste trabalho propõe uma outra abordagem: sendo o alto-forno um reator contínuo em contracorrente parece mais razoável associar a produtividade à sua seção transversal a uma determinada altura. Não se trata de uma novidade mas o resgate de uma melhor concepção, já que era assim que se indicava a produtividade na Alemanha (Hartig, 1996) e Holanda (Toxopeus, 1996) antes dos japoneses se firmarem como a grande referência mundial na tecnologia. Numericamente as duas formas de medida são praticamente equivalentes porque há uma relação linear quase proporcional entre as duas grandezas quando se considera os altos fornos do mundo, como se vê na Figura 3.1 . Do ponto de vista conceitual, no entanto, a forma defendida neste trabalho aproxima a visão do processo aos fenômenos reais.

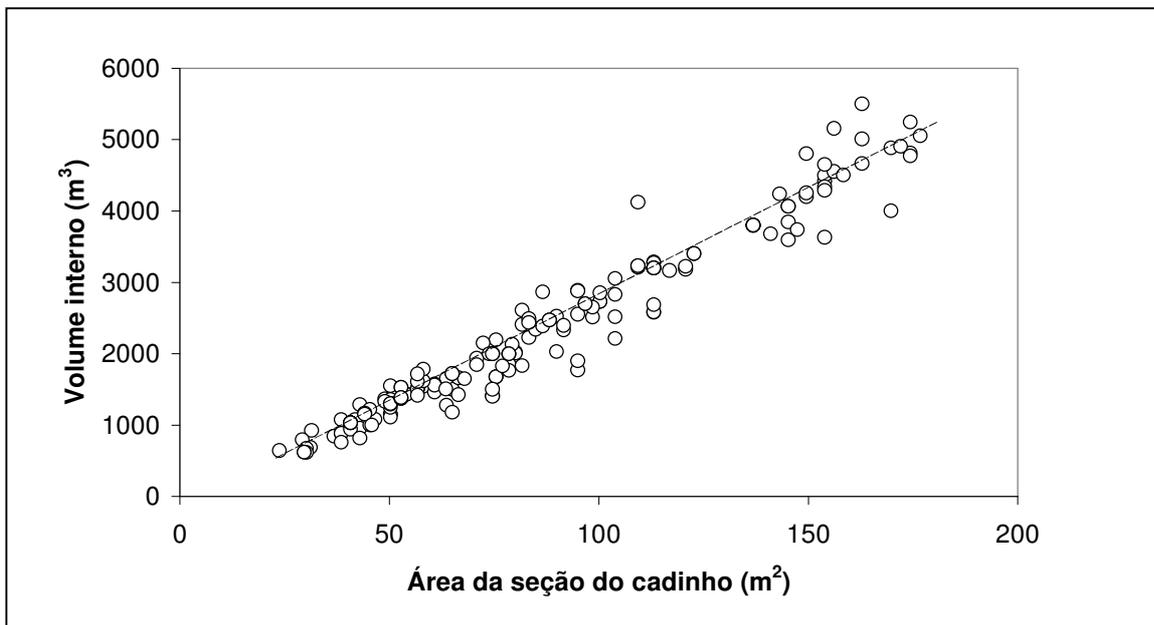


Figura 3.1 – Relação entre volume interno e área da secção transversal do cadiño de altos-fornos do mundo (VdEh)

A Figura 3.2 mostra a concepção básica do modelo proposto.

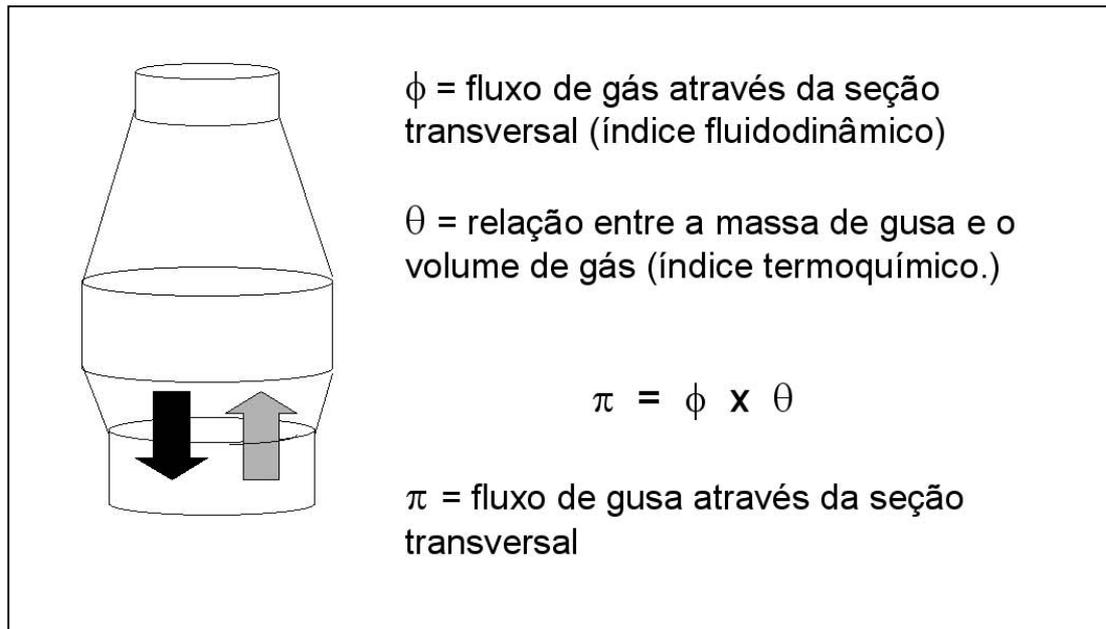


Figura 3.2 – Esquema básico do modelo proposto

A formulação bi-dimensional possibilita a representação do regime operacional em um plano cartesiano onde as dimensões são  $ML^{-3} \times L^{-1}T^{-1}$ .

Ao longo do eixo horizontal do plano cartesiano se representa a variável  $\phi$  (*Índice de Desempenho Fluidodinâmico*) que quantifica o volume de gás que atravessa uma seção transversal do reator na unidade de tempo.

Ao longo do eixo vertical representa-se a relação entre massa de gusa e gás de processo associado. Essa variável, que passa a ser designada por  $\theta$ , é o *Índice de Desempenho Termoquímico* do processo.  $\theta$  é o inverso da geração específica de gás que deixa a zona de elaboração e representa a produtividade do gás de processo. Está, portanto, associado ao consumo térmico específico daquele regime operacional, uma vez que o gás carrega os elementos que geram a maior parte do calor do processo.

A cada momento o processo apresenta um par ordenado  $(\phi, \theta)$  cujo produto quantifica a produtividade global  $ML^{-2}T^1$ , e será indicado neste trabalho por  $\pi = \phi \times \theta$ .

Na Figura 3.3 apresenta-se as condições do Alto-Forno #3 da CSN ao longo dos 12 meses de 2003. Os eixos estão em escala logarítmica de forma a explicitar as isométricas para  $\pi$  em forma de retas isométricas. Na legenda os valores foram multiplicados pela área da seção do cadinho gerando as condições para 9,5; 10,0; 10,5 e 11,0 mil toneladas de gusa por dia. Pode-se fazer duas observações:

- Os índices variam significativamente ao longo do tempo;
- É possível obter-se determinada produtividade global ajustando  $\phi$ ,  $\theta$  ou ambos. Em outras palavras: há mais de uma forma de ajustar a operação para se obter a mesma produtividade. Algumas estarão associadas a custos menores.

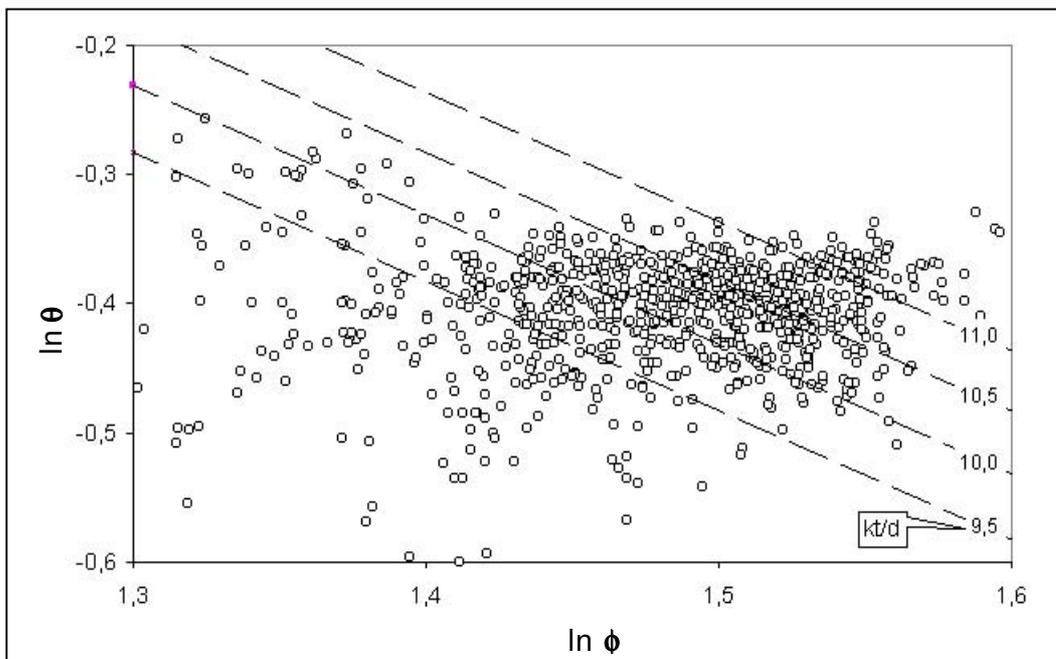


Figura 3.3 - Diagrama  $\phi \times \theta$  para o Alto-Forno 3 da CSN ao longo de 2003.

### 3.1 ÍNDICE DE DESEMPENHO FLUIDODINÂMICO, $\phi$

Tomando-se como ponto de partida a equação de Ergun simplificada, Equação 2.2, é possível identificar os seguintes fatores:  $V_{gv}^{1,7}$ , que é a vazão de gás de ventaneira à potência 1,7; o fator  $(P_{as}^2 - P_{at}^2)$ , que é um parâmetro totalmente imposto pelo conjunto equipamento+operador e a resistência fluidodinâmica  $K$ . Re-ordenando a expressão (2.2) e associando o fluxo a uma seção unitária tem-se:

$$V_{gv} / A = (P_{as}^2 - P_{at}^2)^{0,59} / (K^{0,59} \mu g^{0,18} \rho g^{0,41} A) \quad (3.1)$$

Ou seja: um parâmetro de vazão, fisicamente semelhante ao que foi chamado de  $\phi$ , como resultado da razão entre um parâmetro de potência fluidodinâmica e um parâmetro de resistência ao fluxo gasoso.

Esse parâmetro pode ainda ser modificado para maior sensibilidade e conforto dimensional, se a vazão de gás de ventaneira for substituída pela vazão do gás de rampa, ou seja, do gás que atravessa a linha de Boudouard, e a potência do fator de pressão for ajustada para 0,5.

A vazão do gás de rampa é mais representativa para o processo que a de gás de ventaneira, porque incorpora o gás gerado entre a zona de combustão e a linha de Boudouard. A primeira é normalmente usada por ser de cálculo mais imediato pois depende apenas das condições de sopro.

A expressão pode ser re-escrita então como:

$$\phi = \delta / \rho \quad (3.2)$$

Espera-se que  $\rho$ , a nova resistência fluidodinâmica, dependa das características dos materiais, como parâmetros granulométricos, fator de forma e fusibilidade. O

fator  $\delta$ , o novo diferencial de pressão, corresponde a  $(P_s^2 - P_t^2)^{1/2}$  e tem as dimensões simples de pressão: atm, bar ou Pascal.

À primeira vista parece estranho que o volume de gás seja expresso em condições normais de temperatura e pressão. Isto é possível porque, na verdade:

$$(P_s^2 - P_t^2) = (P_s + P_t) (P_s - P_t) \quad (3.3)$$

Há portanto dois fatores contidos em  $\delta$ . O primeiro, soma das pressões quantifica a média das pressões e corrige o volume. O segundo é o verdadeiro potencial de pressão. O outro fator de correção do volume, a temperatura, é parcialmente influenciada pela temperatura de chama, que estará dentro de  $\tilde{\rho}$ . A temperatura de chama será discutida mais adiante.

Há algum tempo o circuito vem sendo apresentado como uma série de resistências correspondentes a seções do forno entre o nível das ventaneiras e do topo (Kuwubara e Muchi, 1976). A divisão é feita de acordo com a posição de tomadas de pressão que atravessam as paredes a diferentes níveis e, às vezes, em diferentes setores. Nesta proposta, porém toda a coluna é tomada como um só segmento de resistência, considerando o grande número de variáveis envolvidas e a natureza não linear das relações. Constitui um caso típico de aplicação das técnicas de análise através de redes neurais artificiais (Haykin, 1980).

O índice  $\rho$  leva outra vantagem em relação ao índice normalmente usado na prática industrial,  $K$ : independe do tamanho do forno. Adotando-se  $\rho$  é possível comparar a resistência de fornos de tamanhos diferentes. Com  $K$ , compara-se apenas condições distintas de um mesmo forno.

É comum os especialistas na área de altos-fornos também falarem em permeabilidade e não em resistência. Onde for conveniente referir-se à permeabilidade, entenda-se que esta será igual a  $1/\rho$ .

### Temperatura de chama

A temperatura de chama é um dos índices tecnológicos de uso mais generalizados na tecnologia de altos-fornos. Advém do cálculo de um balanço termoquímico adiabático da zona de combustão do forno assumindo que a chama é uma mistura gasosa contendo apenas CO, H<sub>2</sub> e N<sub>2</sub>. Gusa e escória líquidos, que atravessam a zona, são ignorados. Considera-se que o coque que entra no volume de controle vindo da zona de elaboração está a uma temperatura fixa ou proporcional à própria temperatura de chama.

Embora essas premissas sejam universais, com exceção da questão da temperatura do coque, cada empresa desenvolve a sua equação que é o resultado de um balanço padrão. Essa equação tem, em geral, a seguinte forma:

$$T_{ch} = a_0 + a_1 T_s + a_2 U_s + a_3 E_s + a_4 t_i \quad (3.4)$$

onde:

$T_s$	temperatura do sopro (°C)
$U_s$	umidade do sopro (g/m <sup>3</sup> )
$E_s$	enriquecimento do sopro com O <sub>2</sub> (m <sup>3</sup> /m <sup>3</sup> de sopro)
$t_i$	taxa de injeção auxiliar (gás, óleo ou carvão) (kg/m <sup>3</sup> de sopro)

A Figura 3.4 ilustra o balanço termoquímico que permite o cálculo da temperatura de chama e que conduz à Equação 3.4.

É evidente que o valor da temperatura de chama depende dos valores das entalpias das reações envolvidas entre elas e de decomposição

dos hidrocarbonetos injetados como combustíveis auxiliares na ventaneiras. Para efeitos práticos a maioria dos fornos considera um valor de  $a_4$  para gás natural, outro para óleo e outro para carvão. Para o estudo em pauta se faz necessário conhecer a entalpia de decomposição dos carvões usados. Quando se injeta gás natural, as características mudam muito pouco o valor relativo a uma mistura de 95% de metano e 5% de etano será, com certeza suficiente.

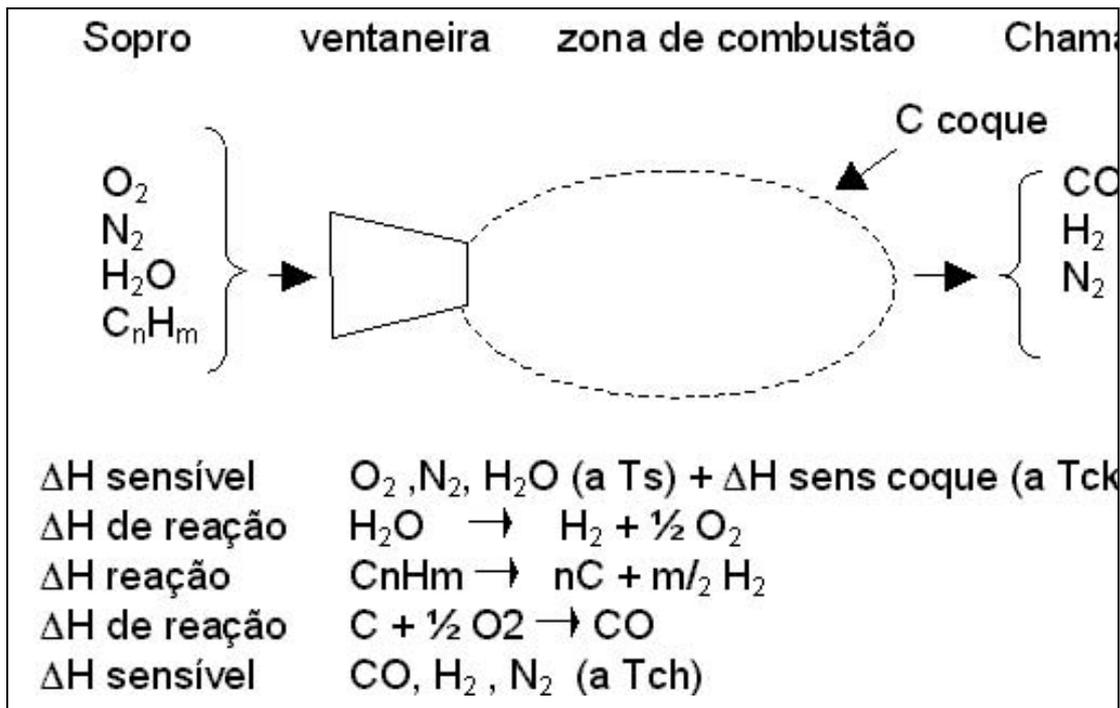


Figura 3.4 – Esquema do balanço termoquímico para cálculo da temperatura de chama,  $T_{ch}$ .

Conhecendo-se o poder calorífico e a composição química dos carvões é possível determinar a sua entalpia de decomposição.

Um pequeno estudo foi desenvolvido com vistas na obtenção da entalpia de decomposição de diversos carvões conhecidos através de seus poderes caloríficos superiores e suas composições químicas resultando em uma

correlação útil para carvões que não passaram pela avaliação de seus poderes caloríficos. Na Tabela 3.1 encontram-se dados dos carvões estudados. Os valores de Hdec, entalpia de decomposição do carvão, correspondem à diferença entre a energia de formação calculada a partir da composição química e o poder calorífico superior (PCS), foi determinado em laboratório. Esse parâmetro é relacionado ao teor de oxigênio na Figura 3.5. De posse desse valor é possível calcular o balanço térmico na zona de combustão diretamente da composição química do carvão.

Tabela 3.1 – Características de carvões usados no estudo

Carvão:	Pittsburg	Elkhorn	Lewiston	Midle Kitt.	Kentucky	Illinois	Indiana
Rank	AV-A	AV-A	AV	AV-B	AV-B	AV-C	AV-C
C	76,6	79,2	66,2	68,4	71,5	62,8	63,4
H	5,2	5,7	4,2	5,6	5,80	5,90	5,7
N	1,6	1,5	1,3	1,4	1,6	1,0	1,3
O	6,2	10,0	8,65	16,4	14,3	17,4	18,6
S	1,3	0,60	0,69	1,20	2,60	4,30	2,3
Cinzas	9,10	3,0	19,0	7,0	7,2	8,6	8,7
MV	30,0	35,0	29,4	36,1	39,8	40,2	36,6
MV (bslmm)	33,0	36,1	36,3	38,8	42,9	44,0	40,1
PCS (MJ/kg)	31,66	33,24	26,80	28,28	30,12	26,70	26,56
Hfor (MJ/kg)	32,57	34,09	27,71	30,43	31,8	29,18	28,99
Hdec (MJ/kg)	0,91	0,85	0,91	2,15	1,68	2,46	2,43

O *rank* do carvão é um indicador do seu grau de metamorfismo. AV significa que os carvões são classificados como alto-voláteis. C, H, N, O e S indicam a composição elementar da matéria orgânica do carvão. Cinzas e MV, índice de matérias voláteis, são resultados da análise imediata. MV(bslmm) corrige o índice de matérias voláteis para a base seca livre de matéria mineral. PCS corresponde

ao poder calorífico superior do carvão medido em bomba calorimétrica, Hdec é calor de decomposição calculado.

Obeve-se assim uma correlação entre a composição química e o teor de oxigênio do carvão, e a sua entalpia de decomposição. Essa relação é útil nos casos em que não se tem o poder calorífico do carvão, o que é muito frequente.

Uma vez que a temperatura de chama dependerá de variáveis presentes no banco de dados, optou-se por não considerá-la como variável de entrada da RNA, tomando-se o cuidado, no entanto de garantir a presença do teor de oxigênio do carvão injetado. Assim será como se a RNA calculasse a temperatura de chama internamente, nos liberando de considerações subjetivas, como a temperatura do coque que entra na zona de combustão e o caráter adiabático do balanço e dos erros resultantes dessas considerações.

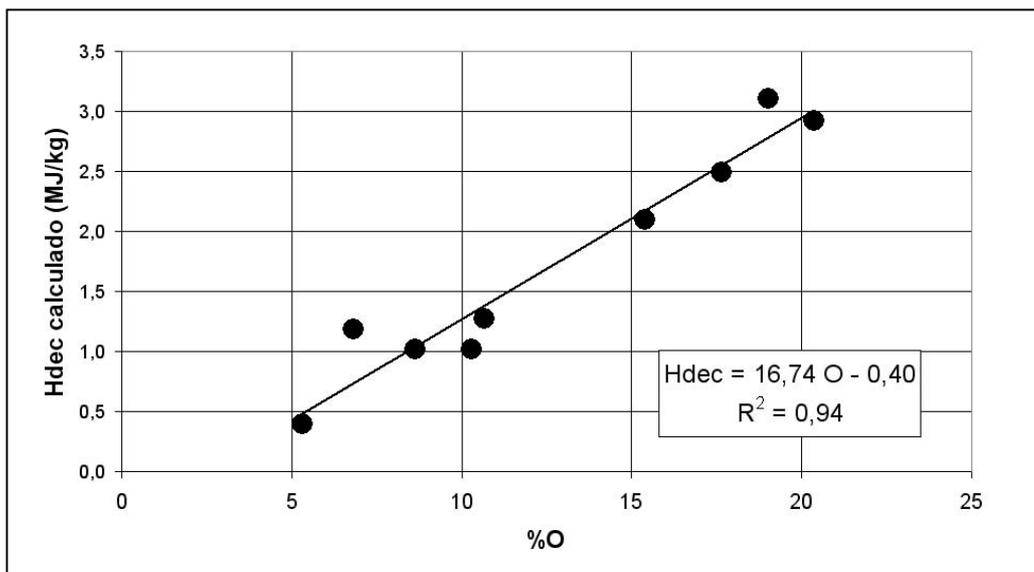


Figura 3.5 – Relação entre o teor de oxigênio e o calor de decomposição de carvões.

### 3.2 ÍNDICE DE DESEMPENHO TERMOQUÍMICO, $\theta$

A estequiometria associada a condições de equilíbrio fisicoquímico e a cinética das reações comandam o índice  $\theta$ . Uma vez que o gás, formado na zona de elaboração se modifica em composição e temperatura, as condições no limite da zona de elaboração podem ser tomadas como referência, sendo as transformações físicas e químicas tratadas através dos fatores de cada zona. O cálculo envolve a solução de um balanço de massa simultâneo de ferro, carbono, hidrogênio e oxigênio, levando em conta também condições de equilíbrio químico.

O modelo de Rist., Michard & Meysson (1967) discutido no Capítulo 2 e esquematizado na Figura 2.6 permite um cálculo racional do balanço de massa, conduzindo ao índice  $\theta$ .

De acordo com aquele trabalho, um único parâmetro pode definir as condições estequiométricas no sistema Fe-O-C-H : trata-se do índice  $\omega$ , que mede o desvio do equilíbrio entre wustita, ferro metálico e o gás redutor na fronteira entre a zona de elaboração e de preparação, também chamada linha de Boudouard.

O uso do gás de rampa, em lugar do gás de ventaneira permite a harmonização com o modelo proposto com o modelo de Rist-Michard-Meysson. O índice  $\mu$  indica a relação entre o número de átomos-kg de carbono gaseificado e hidrogênio e o número de átomos-kg de ferro. O número de átomos-kg de hidrogênio aparece sempre multiplicado pelo fator dois porque esse elemento aparece em pares nas moléculas que o contém e que participam do balanço. Como o volume específico de gás de rampa contém exatamente o volume de  $CO+CO_2+H_2+H_2O$ , resta apenas somar o volume de nitrogênio. Chamando de  $\nu$  o número de moléculas-kg de nitrogênio por átomo-kg de ferro e de  $\gamma$  o número de átomos-kg de ferro em uma tonelada de gusa, tem-se:

$$\theta = ((\mu + \nu) \gamma \times 22,4)^{-1} \quad (3.5)$$

Os fatores relativos ao calor sensível e à entalpia de reações estão contidas no cálculo de  $\mu$ . O índice se associa a um balanço térmico estagiado considerando também condições de trocas térmicas entre gás, sólidos e líquidos.

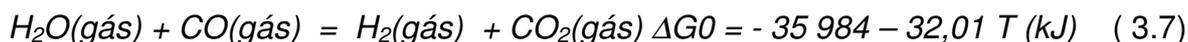
A divisão do processo em zonas como apresentado na Figura 2.2 permite a estratificação do balanço químico e térmico do processo.

O valor de  $\square$  depende do valor de  $\square$  chamado por Rist-Michard-Maysson (1967) em seu modelo de “desvio da idealidade”. Indica o excesso de oxigênio ligado ao ferro na fronteira entre zona de preparação e zona de elaboração. É medido em moles de O/moles de Fe. Idealmente essa relação é próxima a 1,06, que representa o valor estequiométrico da wustita, FeO em equilíbrio com ferro metálico na temperatura da linha de Boudouard. Ou seja, na temperatura em que a reação do CO<sub>2</sub> com o coque (3.1) dispara.



Esta relação pode ser melhor compreendida quando associada ao diagrama de equilíbrio Fe-O mostrado na Figura 3.6. O valor de  $y$ , em FeO corresponde à wustita na temperatura de equilíbrio. Essa sua relação com o equilíbrio de Boudouard fica mais clara no diagrama de Choudron, mostrado na Figura 3.7.

O diagrama de Choudron permite observar a relação entre a temperatura de equilíbrio e a composição do gás na mistura CO+CO<sub>2</sub>+H<sub>2</sub>+H<sub>2</sub>O. Neste sistema o equilíbrio associado à reação (3.7) é também relevante.





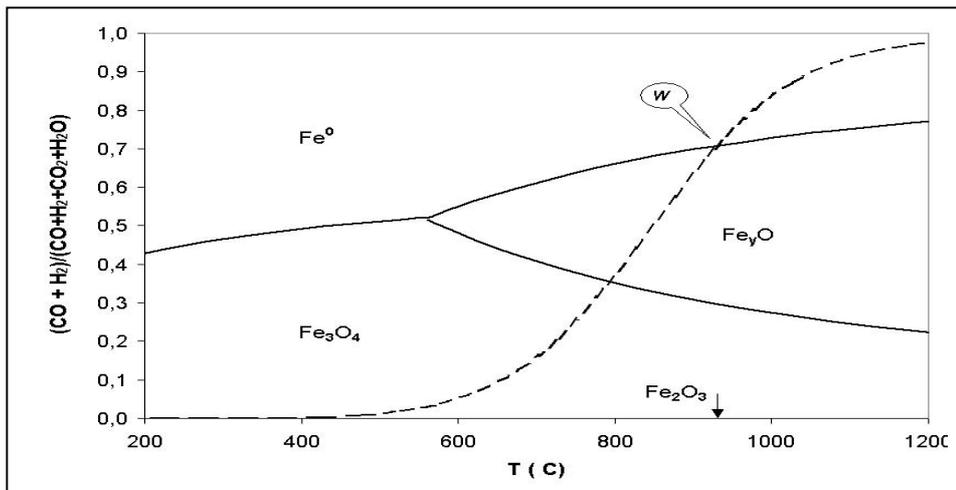


Figura 3.7 – Diagrama de Choudron para o sistema Fe-O-C-H

Já considerados os elementos que quantificarão temperatura e tempo de residência, resta envolver as variáveis ligadas ao minério que determinam o rendimento da própria reação de redução, ou seja, a redutibilidade dos minérios, RI, e sua distribuição granulométrica, ou frações relevantes. A proporção das diversas fontes de óxido de ferro na carga é naturalmente importante.

As incógnitas que restam para o cálculo preciso de um balanço para cada umas das zonas, de preparação e de elaboração são o teor de silício do gusa,  $[Si]$ , e a perda térmica na zona de elaboração  $\lambda$ .

Pelo que foi discutido no Capítulo 2, as variáveis que devem estar presentes na rede neural para predição do teor de silício do gusa,  $[Si]$ , são aquelas ligadas ao equilíbrio químico no sistema  $Si-O-C$ , ou seja, temperatura dos gases, pressão de  $O_2$ , atividade do  $SiO_2$  nas fases relevantes, cinza do coque e escória primária e final e o tempo de residência, ou seja, a velocidade do processo.

A temperatura dos gases relacionada ao fenômeno está associada à fronteira entre a zona de combustão e o núcleo de coque, onde ocorre a formação de  $SiO$  (gás), chamada temperatura de chama,  $T_{ch}$  e que calculável através de variáveis do sopro. Essa variável será discutida em detalhe mais adiante.

Como já relatado, o tamanho do forno influi no teor do silício. Essa variável não será adotada nesta análise estatística uma vez que os dados são de um só forno neste trabalho.

### Balanco térmico na zona de redução indireta

Como discutido anteriormente o valor de  $\Delta H$ , que com  $\Delta T$  irá definir  $Q$ , está associado ao balanço de massa (envolvendo ferro, oxigênio, carbono e hidrogênio) e ao balanço térmico da zona de elaboração.

Tabela 3.2 – Dados termodinâmicos,  $HT - H_{298K}$  (MJ/kg-mol)

Calor sensível do sopro	
$O_2$	$0,0362 T_s - 13,5$
$N_2$	$0,0344 T_s - 13,0$
$H_2O$	$0,0474 T_s - 22,2$
Calor sensível da carga	
$FeyO$	$0,0586 T_c - 21,4$
$Fe_3O_4$	$0,2010 T_c - 53,2$
$SiO_2$	$0,0714 T_c - 24,3$
$Al_2O_3$	$0,1280 T_c - 50,0$
$CaO$	$0,0542 T_c - 19,8$
$MgO$	$0,0522 T_c - 19,3$
$MnO$	$0,0569 T_c - 21,8$
$CaO.SiO_2$	$0,1280 T_c - 49,5$
$Ccoque$	$0,0235 T_c - 11,8$
Calor sensível do gás de rampa	
$CO$	$0,0353 T_{gr} - 14,0$
$CO_2$	$0,0586 T_{gr} - 26,0$
$H_2$	$0,0325 T_{gr} - 12,2$
$H_2O$	$0,0474 T_{gr} - 22,2$
$N_2$	$0,0344 T_{gr} - 13,0$
Calor sensível dos líquidos (MJ/t)	
Gusa	$0,786 T_{gu} - 0,104$
Escória	$1,103 T_{es} - 0,431$

A massa e composição das escórias primária e final derivam de um balanço de CaO, MgO, SiO<sub>2</sub> e Al<sub>2</sub>O<sub>3</sub>. O enxofre, que se distribui entre gusa e escória, tem partição associada a uma variável de saída da rede como discutido no início deste Capítulo. Os outros elementos relevantes, manganês e fósforo, têm partição igual à média histórica.

Depois de fechado o balanço térmico, o modelo procede ao balanço térmico que utiliza os dados apresentados na Tabela 3.2.

O calor sensível do gusa inclui o calor de fusão no caso do gusa e da escória.

Os dados foram obtidos de Stull et alii (1971), linearizados para os intervalos de variação e convertidos a unidades ISO.

As temperaturas indicadas são:

$T_s$  – temperatura de sopro (K) – variável;

$T_c$  – temperatura da carga na linha de Boudouard – fixa em 1250 K;

$T_g$  – temperatura do gás na linha de Boudouard – fixa em 1300 K;

$T_{gu}$  – temperatura do gusa (K) – variável;

$T_{es}$  – temperatura da escória (K) – variável e igual a  $T_{gu} + 50K$ .

As temperaturas que foram fixadas não são medidas. Suas estimativas dependem de algumas variáveis que estarão na rede neural. Espera-se que as consequências da fixação desses valores sejam absorvidas por  $\lambda$ , o valor das perdas térmicas.

A Tabela 3.3 completa os dados do balanço com os valores de entalpia de reação adotados.

Para compor as parcelas do balanço térmico os valores de entalpia e calor sensível são multiplicados pelos respectivos números de mol-kg ou pela massa no caso do gusa e da escória. As substâncias que deixam a zona de elaboração, ( gás de rampa, gusa e escoria) têm indicadores de massa negativos enquanto que para os que entram na zona de elaboração tem medidas de massa positivas. A base de cálculo é uma tonelada de gusa.

Tabela 3.3 – Dados termodinâmicos,  $\Delta H_{298K}$  (MJ/kg-mol)

$\Delta H_5$ – Entalpia de reação	
$H_2 O (g) \rightarrow H_2 + \frac{1}{2} O_2$	+ 242
$CH_4 \rightarrow C + 2H_2$	+ 75
$C_2H_6 \rightarrow 2C + 3H_2$	+ 242
$C(\text{coque}) \rightarrow C$	+ 13
$C \rightarrow CO$	- 111
$C \rightarrow CO_2$	- 394
$FeyO \rightarrow yFe + \frac{1}{2} O_2$	+ 266
$Fe_3O_4 \rightarrow 3Fe + 2O_2$	+ 1121
$SiO_2 \rightarrow Si + O_2$	+ 911
$MnO \rightarrow Mn + \frac{1}{2} O_2$	+ 385
$Si \rightarrow Si(\text{gusa})$	+ 10
$Mn \rightarrow Mn(\text{gusa})$	+ 9
$C \rightarrow C(\text{gusa})$	+ 8

As massas ligadas as reação são negativas se os reagentes são consumidos ou positivas quando são gerados. O somatório geral resulta em:

$$\sum \Delta H_i - \lambda = 0 \quad (3.5)$$

onde  $\lambda$  é o índice de perdas térmicas em MJ/t de gusa.

### 3.3 ARQUITETURA DA REDE NEURAL

Pelo que foi apresentado nos últimos parágrafos, verifica-se que é necessário alimentar os balanços de massa e térmico com 5 variáveis que dependem das matérias-primas e do regime operacional do forno. Essa tarefa é possível se a rede for capaz de estimar esse dados. A estrutura global do modelo híbrido pode ser apreciada na Figura 3.8.

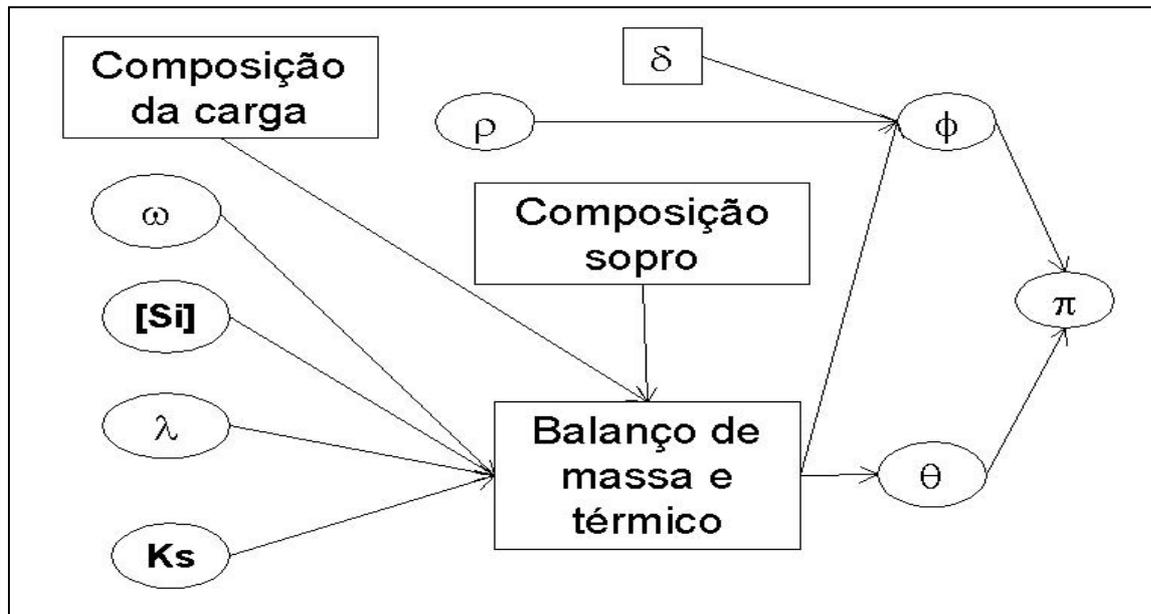


Figura 3.8 – Esquema global do modelo híbrido.

Considerando a disponibilidade de dados, 3 anos de operação do Alto-Forno 3 em vetores equivalentes a um dia de operação cada, optou-se por uma arquitetura com mínimo número de parâmetros capaz de garantir os objetivos.

Com esse objetivo a rede deve ter uma camada de entrada com número de neurônios equivalentes as variáveis que forem definidas, uma camada intermediária com número de neurônios que garanta mínimo erro na generalização e uma camada de saída com 5 neurônios equivalente às variáveis definidas neste capítulo. A estrutura dessa rede foi apresentada na Figura 2.7.

É importante usufruir da propriedade da função log-sigmoidal que ameniza a influência de valores extremos entre a camada de entrada e a intermediária. (Equação 2.3).

A definição das variáveis de entrada será descrita na próxima seção. O número de neurônios da camada intermediária será consequência de estudo descrito no Capítulo 4.

### **3.4 DEFINIÇÃO DE VARIÁVEIS E DO BANCO DE DADOS**

O banco de dados em que se baseia este trabalho foi obtido do sistema digital de automação do Alto-Forno #3 da CSN, desde o início de sua atual campanha, em 8 de agosto de 2001, até 31 de dezembro de 2004, complementado com os registros de operação e de relatórios de análises dos laboratórios da Usina Presidente Vargas e da Mina de Casa de Pedra.

Cada vetor corresponde a um dia de operação. Os dados relativos a matérias primas foram tomados com a defasagem de 1 dia, considerando o tempo de trânsito e residência nos silos.

Vários arquivos dão origem ao banco de dados utilizado neste estudo. Cada um tem uma base temporal diferente, algumas horárias, outras diárias, outras baseadas em eventos como a chegada de uma carga de minério ou uma corrida de gusa. Constituiu parte fundamental deste trabalho a consolidação deste banco de dados final. Algumas variáveis são índices secundários calculados a partir das variáveis primárias do banco.

As variáveis do banco de dados final encontram-se descritas na Tabela 3.4 com suas médias e desvios-padrão.

O período de novembro de 2001 a dezembro de 2004 possibilitou a obtenção de 820 vetores com as 23 variáveis de entrada e 5 de saída. Os dias em que a produtividade foi reduzida por parada de equipamentos ou que não tinham os dados suficientes para o cálculo das variáveis não foram consideradas.

Tabela 3.4 – Variáveis de entrada e de saída da rede neural

	média	desvio	descrição da variável
Variáveis de entrada - Sopros			
X1	29,26	1,26	N2 no sopros/gusa (kgmol/t)
X2	16,82	3,62	H2O no sopros/gusa (kgmol/t)
X3	1056	24	temperatura do sopros (°C)
X4	147	14	carvão pulverizado/gusa (kg/t)
X5	77	22	teor de cinzas no carvão pulverizado (kg/t)
X6	71	16	teor de oxigênio no carvão pulverizado (kg/t)
Variáveis de entrada - Carga			
X7	260	13	escória/gusa (kg/t)
X8	0,99	0,03	(CaO+MgO)/(SiO2+Al2O3) da escória final (kg/kg)
X9	1,44	0,06	(CaO+MgO)/(SiO2+Al2O3) da escória primária (kg/kg)
X10	24	45	pelotas/gusa (kg/t)
X11	417	44	hematita/gusa (kg/t)
X12	3,3	3,2	fundentes/gusa (kg/t)
X13	134	561	sínter do estoque/gusa (kg/t)
X14	53	13	small-coke/gusa (kg/t)
X15	142	91	coque externo/coque total (kg/t)
X16	85,3	0,8	Índice de resistência médio do coque
X17	21,6	1,4	índice de reatividade médio do coque
X18	56,6	3,4	tamanho médio do coque (mm)
X19	110	40	fração < 6,35 mm na hematita (kg/t)
X20	111	31	fração > 38 mm na hematita (kg/t)
X21	5,5	1,6	índice de crepitação da hematita
Variáveis de entrada – Equipamento/ Ambiente			
X22	0,15	0,36	Índice de idade do furo de gusa (0-1)
X23	4,1	10,1	índice de precipitação pluviométrica (mm/d)
Variáveis de saída			
$\omega$	0,18	0,02	desvio estequiométrico da wustita (kgmol/kgmol)
$\rho$	0,83	0,03	resistência fluidodinâmica média (bar/m)
[Si]	3,8	0,9	silício no gusa ( $10^{-3}$ )
$\lambda$	867	247	perdas térmicas /gusa (MJ/t)
Ks	34	10	teor enxofre na escória/teor de enxofre no gusa (-)

No que se refere às variáveis de entrada o banco se divide em 3 blocos: variáveis do sopro, variáveis da carga e variáveis do equipamento ou do ambiente.

No bloco das variáveis ligadas ao sopro  $X1$  quantifica o volume de nitrogênio no sopro e, com uma pequena contribuição do nitrogênio contido no carvão pulverizado, cerca de 0,2 kg-mol/t de gusa, será também o nitrogênio do gás de rampa. Embora não participe significativamente das reações químicas, o nitrogênio serve de portador de calor da zona de elaboração para a zona de preparação ajudando a pré-aquecer a carga. Com certeza, a operação com pequeno volume de nitrogênio estará associada a um histórico térmico da carga mais desfavorável para a redução indireta e conseqüentemente maior valor de  $\omega$ , daí a sua inclusão na lista de variáveis de entrada.

A variável  $X2$ , é importante por vários aspectos: influi na temperatura de chama, portanto na variável,  $\rho$ ; altera a relação  $(H_2 + H_2O)/(CO + CO_2)$  da qual depende condições de redução indireta e o valor de  $\omega$ .

A temperatura de sopro,  $X3$ , importa no calor introduzido no reator pelo ar quente. Esse aspecto é considerado no balanço térmico. O outro aspecto que justifica a sua inclusão na lista é sua influência na temperatura de chama, assim como  $X1$ .

A variável  $X4$ , taxa de injeção de carvão, além de influir na temperatura de chama determina muitas alterações no processo não só do ponto de vista térmico como fluidodinâmico, uma vez que é um redutor que substitui coque atuando de forma bem distinta deste nas zonas inferiores.

O teor de cinzas do carvão pulverizado,  $X5$ , é responsável pela maior ou menor permeabilidade da camada de coque que envolve a zona de combustão em frente a cada ventaneira. Por esta região deve passar todo o gás que se transformará em gás de rampa, correspondendo portanto a uma parte da resistência fluidodinâmica do forno,  $\rho$ ;

O teor de oxigênio do carvão,  $X_6$ , como foi discutido anteriormente, é um indicador do calor de decomposição deste material na zona de combustão, influenciando significativamente na temperatura de chama e, por consequência, na resistência fluidodinâmica.

Na lista das variáveis ligadas à carga, uma das mais importantes é a taxa de geração de escória,  $X_7$ , ou o volume de escória como é falado entre os altos-fornistas. Influi definitivamente no balanço térmico, mas está também associada à dinâmica de trocas térmicas na zona de elaboração, influenciando em  $\omega$  e em  $\rho$ , uma vez que escoar em contracorrente com o gás.

$X_8$  e  $X_9$  representam, respectivamente, as basicidades da escória final e da escória primária. A primeira está intimamente ligada ao equilíbrio químico do enxofre entre metal e escória e certamente terá papel importante na estimativa de  $K_s$ , o coeficiente de partição deste elemento no processo. A segunda está ligada à fusibilidade da carga e à viscosidade da escória resultante, ou seja, à maior ou menor facilidade com que a carga passará do estado sólido poroso a pastoso e impermeável, e depois a líquido percolando o coque em direção ao cadinho em contra corrente ao gás. Claramente  $\rho$  dependerá desta variável. A escória primária difere da final por não conter ainda a matéria mineral do coque e do carvão que é incorporada depois.

$X_{10}$ ,  $X_{11}$  e  $X_{12}$  definem a composição da carga em termos de materiais. Vários aspectos do processo, entre eles a cinética de redução, que influi em  $\omega$ , a resistência,  $\rho$ , e as trocas térmicas,  $\lambda$ , dependem dessas variáveis.

O sínter que passou pelo estoque ficou exposto às intempéries por tempo indeterminado. As suas fases ficaram expostas à hidratação e à carbonatação, sobretudo se há bases livres, CaO e MgO. Esse fato justifica a inclusão da variável  $X_{13}$  entre as que estimarão as 5 variáveis de saída. O sínter é estocado

e carregado em função de paradas de máquina de síter. Preferencialmente o síter é transportado diretamente das máquinas para os silos do forno.

O *small-coke*, cuja taxa é quantificada por  $X14$ , é carregado junto com os componentes da camada de carga ferrífera, substituindo o coque de tamanho maior que é carregado individualmente. Esse material tem valor menor que o coque, portanto ajuda a reduzir o custo da operação, mas pode redundar em aumento da resistência do forno. Na quantidade adequada, no entanto ajuda na redução direta e pode implicar em menor  $\omega$ . É um caso típico de relação não-linear no alto-forno.

A Usina Presidente Vargas não é auto-suficiente em coque. Uma proporção variada de coque externo, produzido em outras usinas, com características diferentes do coque doméstico é carregada nos altos-fornos. Alterações no processo são observadas pelos operadores mas poder-se-ia até julgar que essa influência já estaria representada na rede nas médias de índices de características do coque que compõem as próximas variáveis de entrada. Entretanto, sabe-se que o coque desempenha vários papéis dentro do forno: estrutura do leito, regenerador do CO do gás, combustível sólido, etc. O forno tende a segregar porções do coque nos diferentes sub-processos, portanto a média não dá informação suficiente sobre o coque total. Daí a necessidade dessa variável  $X15$ .

Como dito no parágrafo anterior, as variáveis  $X16$ ,  $X17$  e  $X18$  indicam as características médias do coque. A resistência é medida pela degradação granulométrica após tamboramento. Neste caso o indicador se chama Drum Index e o ensaio é normalizado nos sistemas JIS (Japanese Industrial Standards) e ISO. A resistência influi na granulometria do coque dentro do forno e está obviamente ligada a  $\rho$ .

A reatividade é uma medida indireta da atividade do carbono no coque. É medida pela perda de massa de uma amostra de coque aquecida em um reator onde passa CO<sub>2</sub>. Portanto altera as condições do equilíbrio de Boudouard influenciando  $\omega$  e  $\lambda$ . *X18*, medida da granulometria do coque influi na resistência fluidodinâmica e na reatividade real do coque.

*X19*, *X20* e *X21* caracterizam a hematita. A fração fina dificulta a passagem dos gases enquanto a grossa tem redutibilidade real menor. A crepitação é um fenômeno que ocorre na hematita entre 250 e 500 °C. A decomposição de fases hidratadas gera finos e essa característica é medida em laboratório gerando o índice adotado. Seria importante ter um índice de redutibilidade, mas o ensaio não é realizado em rotina. O resultados disponíveis não representam as variações no material enornado.

*X22* é uma variável de Bernoulli, que assume o valor 1 se pelo menos um dos furos de gusa em operação tem menos de 1 dia de uso e 0 se não for o caso. O Alto-Forno 3 altera os furos de gusa a cada 28 dias de operação. Nos dois primeiros dias de um furo o escoamento não é tão bom prejudicando a permeabilidade e a transferência de enxofre do gusa para escória no cadinho.

A precipitação pluviométrica, *X23*, aumenta a umidade do minério dificultando a remoção de finos por peneiramento antes do carregamento. O índice refere-se ao dia anterior ao da operação, no qual a hematita passou pelo pátio e recebeu a chuva.

O comportamento de cada variável, de entrada ou de saída, foi verificado através das distribuições de frequência acumulada apresentadas nas Figuras 3.9 e 3.10. A padronização foi realizada segundo a Equação (3.6).

$$Z_i = (X_i - \text{média } X_i) / (\text{desvio padrão } X_i) \quad (3.6)$$

Em todas as figuras no eixo horizontal estão as variáveis de saída da rede normalizadas com base no desvio-padrão.

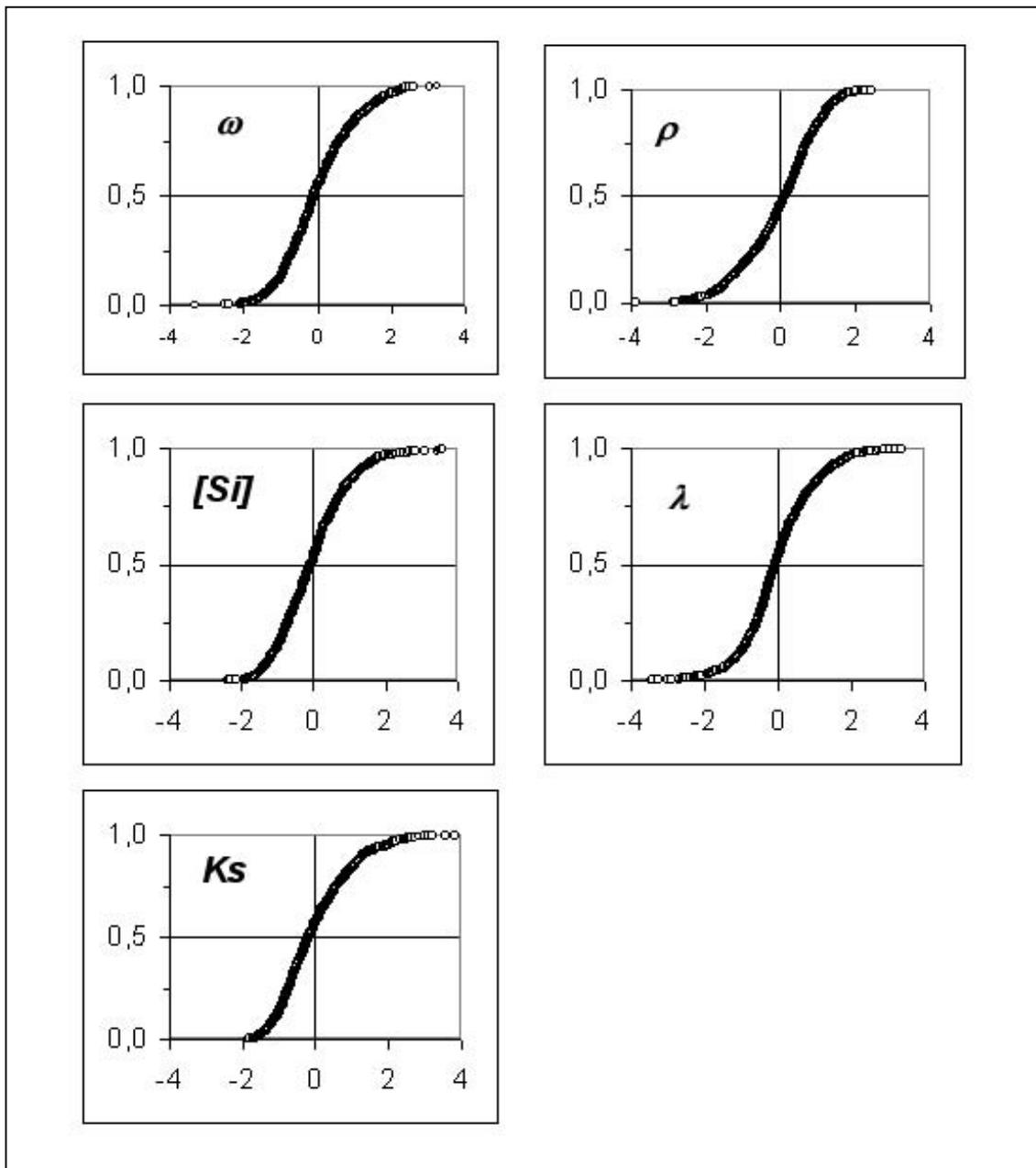


Figura 3.9 – Distribuição de freqüência acumulada das variáveis de saída

As variáveis de saída apresentam distribuições aproximadamente gaussianas.

As distribuições das variáveis de entrada dependem da sua natureza. Algumas variáveis são produtos de processos e medidas com grande frequência como o teor de finos da hematita apresentado distribuição também gaussiana. Outras são impostas pelo operador, com X3, a temperatura do sopro, a taxa de adição de pelotas, de coque externo etc. e resultam em distribuições atípicas. Todas, no entanto, se mostraram coerentes e passíveis do tratamento matemático proposto.

### **Análise de vizinhança (consistência dos vetores)**

Em virtude do tamanho e da natureza do banco de dados, que envolve diversas fontes, algumas automáticas, algumas manuais, foi necessário criar um método para análise da consistência dos dados que será descrito a seguir.

O método proposto pode ser chamado de *análise de vizinhança* e consiste em estabelecer uma identificação para uma vizinhança no espaço  $R^n$ , domínio do vetor  $(x_1, x_2, \dots, x_n)$ . Identificada a vizinhança do vetor verificamos todos os valores de  $y$  correspondentes aos vetores individuais que pertencem a esta vizinhança. A partir daí realiza-se uma análise de variância dos valores de  $y$  nas vizinhanças em relação ao total de valores de  $y$ .

A vizinhança é identificada por um número inteiro de ordem  $n$ , ao qual chamaremos de código da vizinhança, formado pela seguinte função:

$$C = (int \ x \ i + 5) 10^{(i-1)} \quad (3.7)$$

Claramente  $(int \ x \ i + 5)$  será um número inteiro entre 0 e 9, correspondente a um das 10 classes em que foi dividida a distribuição de  $x_i$ .

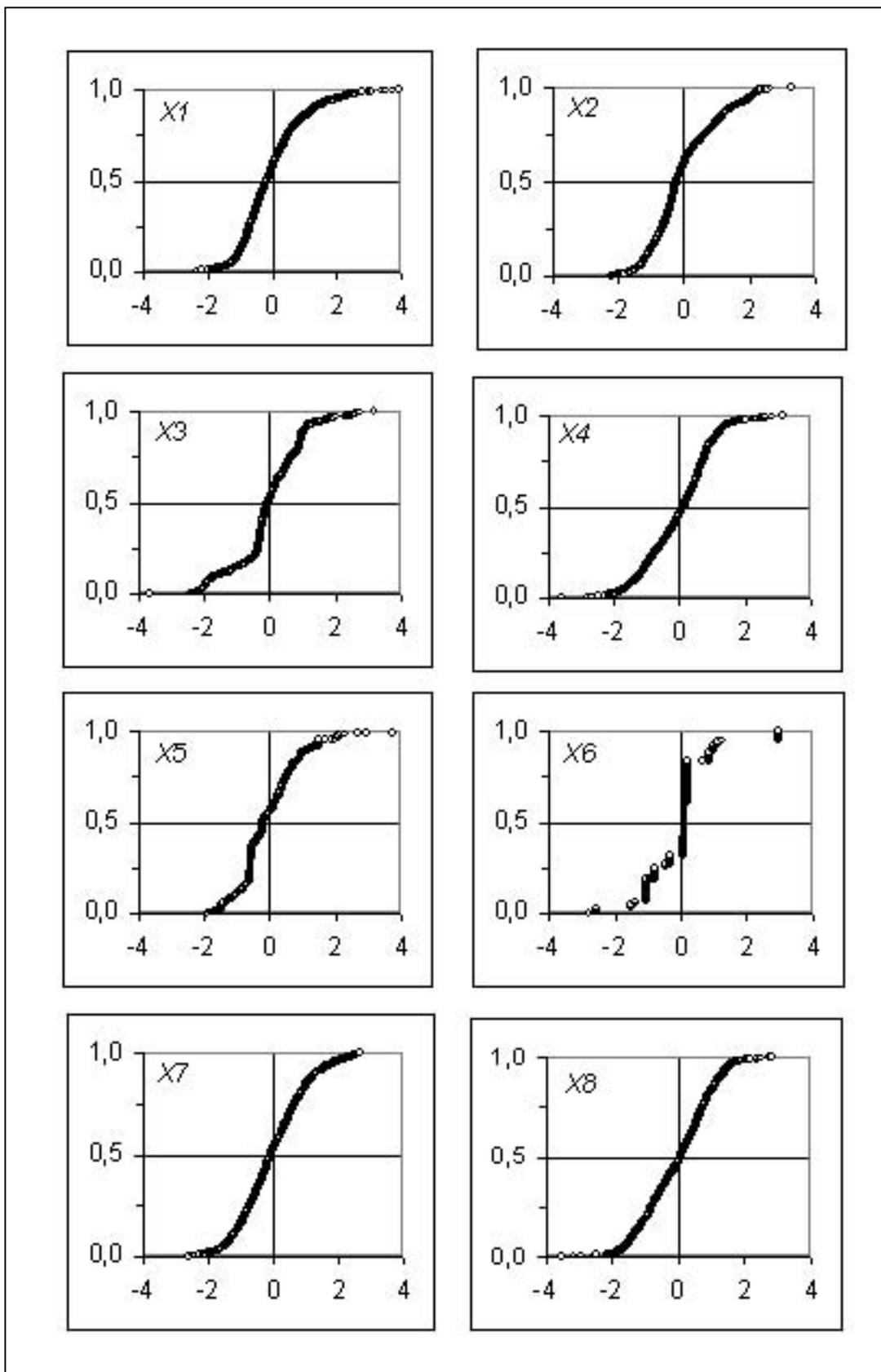


Figura 3.11a – Distribuição de freqüência das variáveis de entrada

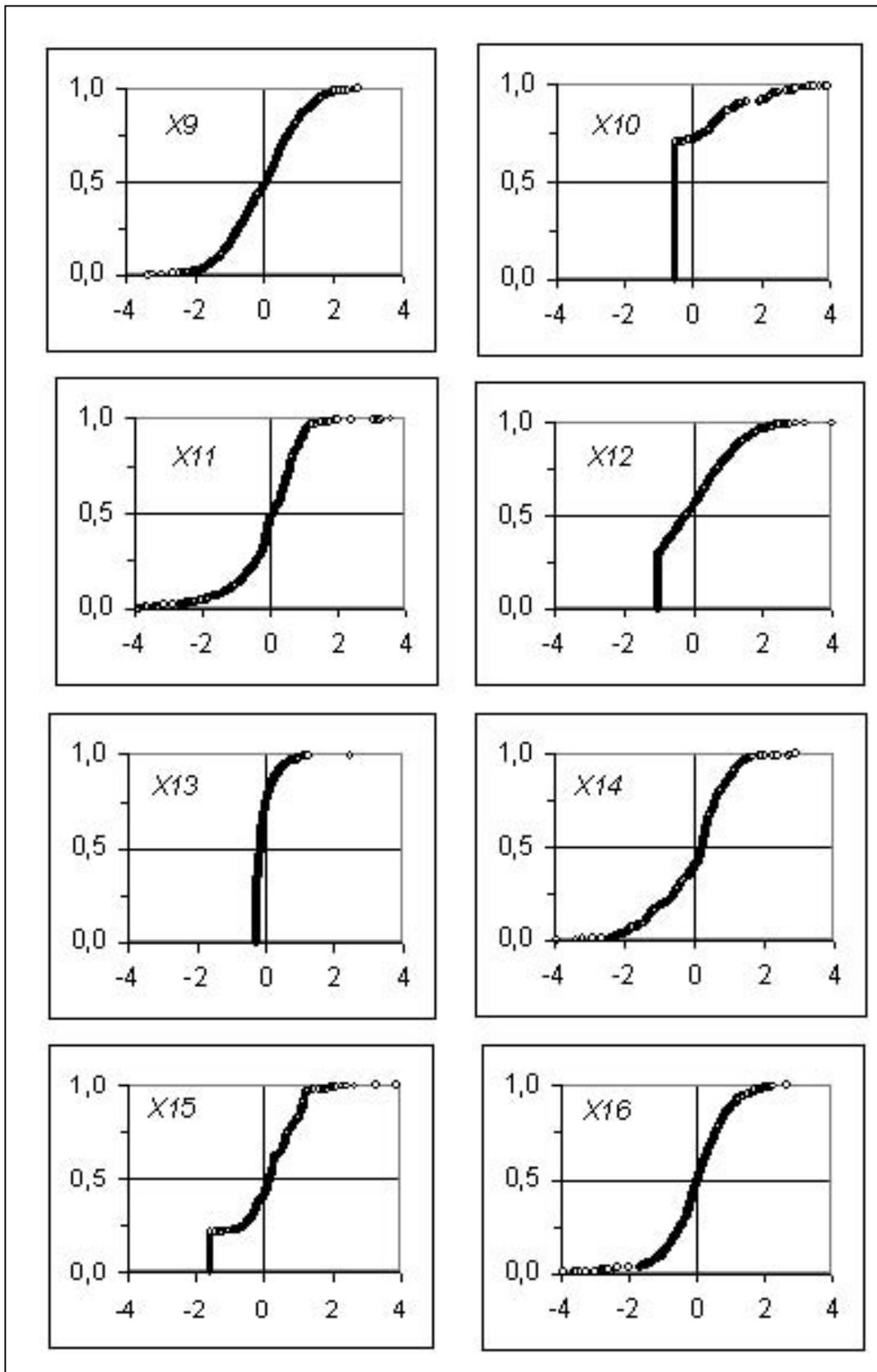


Figura 3.11b – Distribuição de freqüência das variáveis de entrada

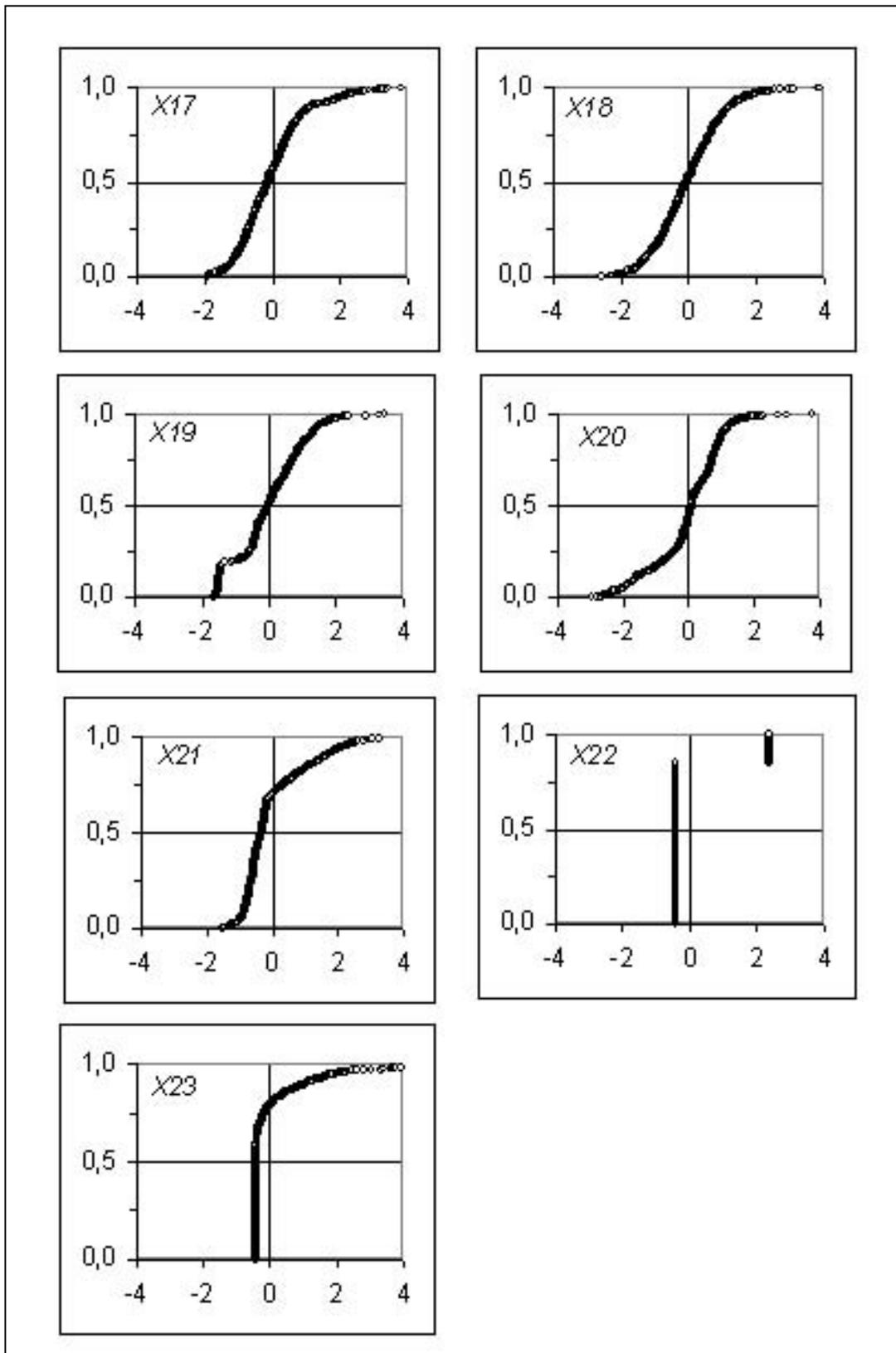


Figura 3.11c – Distribuição de freqüência das variáveis de entrada

Como não há valores de  $x_i$  fora do intervalo  $(-4; +4)$ , uma vez que esses valores foram eliminados na primeira crítica ao banco de dados, não há perigo de contaminação de qualquer dígito pelo dígito antecessor ou sucessor. Cada valor de  $C$  corresponde a uma e só uma das  $10^n$  vizinhanças em que foi dividida a distribuição dos vetores de entrada,  $(x_1, x_2, \dots, x_n)$ .

No presente caso a maior dificuldade será o elevado número de vizinhanças possíveis em função do número de variáveis de entrada, 23. Este número tornou impossível a aplicação do método em todo o banco de dados. O aplicativo empregado neste caso, *Microsoft Excel*®, só reconhece 15 dígitos. Foi possível porém examinar quanto às 15 primeiras variáveis apenas a título de ilustração .

Foram identificadas apenas 11 com mais de um vetor para o número total de 820 vetores totais.

Na Tabela 3.5 compara-se a variância dentro das vizinhanças com mais de 1 “residente”, comparada à variância global. As vizinhanças que têm variância maior que a global podem conter valores anormais. Porém, no presente caso, existem 8 variáveis ainda a controlar o valor das variáveis de saída, portanto o método não se mostrou efetivo. Além disso, a relação entre o número de vetores e de vizinhanças possíveis é muito grande tornando pouco provável a presença de mais de um vetor por vizinhança.

As ferramentas usadas nesta análise foram a ordenação e a consolidação de valores segundo a coluna que continha o código da vizinhança.

O método pode ser útil em redes de menor dimensão.

Tabela 3.5 – Resultado da análise de vizinhança.

código da vizinhança	$\omega$	$\rho$	$[Sij]$	$\lambda$	$Ks$
445666455646455	0,403	1,698	1,408	0,023	0,130
456675645446545	0,104	0,017	0,138	0,047	0,022
54555546453555	0,007	0,013	0,104	0,006	0,489
545776565445555	0,000	0,031	0,058	0,026	1,076
554443656454466	0,032	0,006	0,189	0,059	0,694
555445665463644	0,010	0,023	0,420	0,000	0,182
555553645455465	0,668	0,051	0,120	0,071	0,184
556564565455456	0,028	0,077	0,006	0,117	1,172
564445546453554	0,082	0,352	1,408	1,418	1,861
56455546443552	0,000	0,003	0,862	0,000	0,028
654445565473454	0,001	0,215	1,882	0,005	0,001
<b>Variância global</b>	<b>0,962</b>	<b>0,972</b>	<b>0,983</b>	<b>0,974</b>	<b>0,990</b>

### 3.5 APLICAÇÃO DO MODELO À PRÁTICA INDUSTRIAL

Na aplicação do modelo o que se quer é poder avaliar propostas de regimes operacionais antes de praticá-los. A Figura 3.11 possibilita visualizar o lugar do modelo na operação dos altos-fornos.

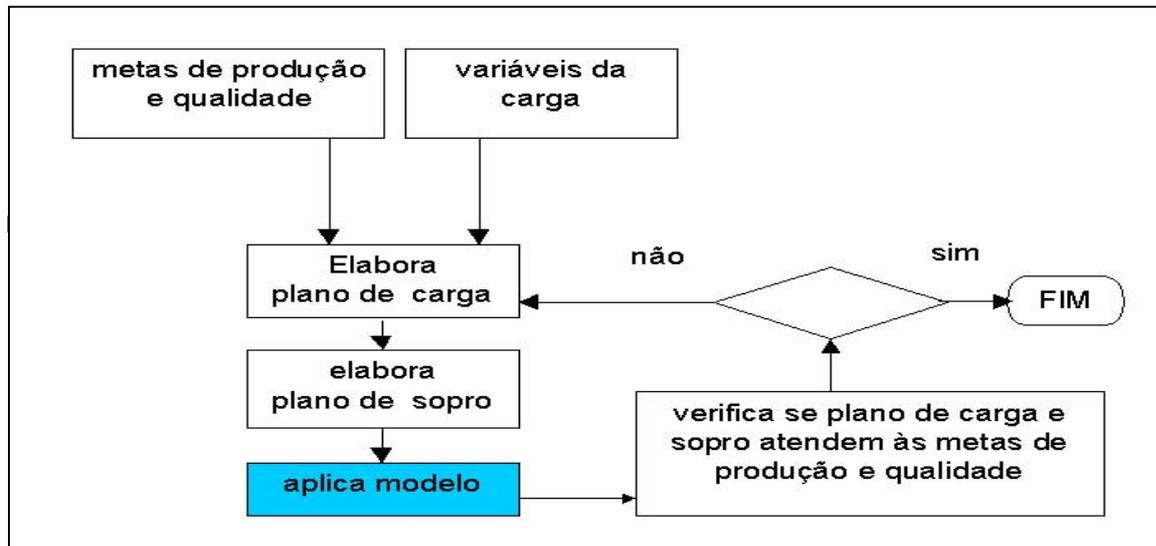


Figura 3.11 – Fluxograma do modelo proposto dentro do padrão operacional dos altos-fornos.

Pelo que foi visto, o modelo será capaz de trazer para avaliação sistêmica a adequação dos planos operacionais, de sopro e carga, às condições das matérias-primas, do ambiente e do equipamento, verificando a viabilidade do plano. O uso do modelo antecipa as dificuldades e evita o ajuste por reação, obviamente menos eficaz que o pré-ajuste do processo.

## CAPÍTULO 4 - RESULTADOS E DISCUSSÃO

### 4.1 TREINAMENTO E GENERALIZAÇÃO DA REDE

Tendo sido definida a rede com uma camada de entrada, uma camada intermediária e uma camada de saída, resta definir o número de neurônios da camada intermediária. Naturalmente a camada de entrada terá 23 neurônios correspondentes às 23 variáveis de entrada. Da mesma forma, a camada de saída terá 5 neurônios correspondentes a  $\omega$ ,  $\rho$ ,  $[S]$ ,  $\lambda$  e  $Ks$ .

Para isso os dados foram submetidos a em uma série de treinamentos de rede nas quais o número de neurônios na camada intermediária tomou os valores de 23, 25, 30 e 35. O processamento foi realizado em triplicata para cada número de neurônios em ambiente *Matlab 6.1* ®. Aplicou-se o programas listado no Anexo I.

Dentro do programa foi feita mais uma normalização linear para manter as variáveis dentro do espaço amostral (0 ; 1)

$$x_i = (X_i - X_i \text{ min}) \times 0,8 / (X_i \text{ máx} - X_i \text{ min}) + 0,1 \quad (4.1)$$

A mesma operação foi aplicada às variáveis de saída.

Como função de ativação nos neurônios da camada intermediária foi adotada a função sigmoideal, denominada em ambiente *Matlab* de *logsig*, conforme Equação 2.3.

Na ativação dos neurônios da camada de saída foi utilizada a função *purelin*, no ambiente *Matlab*, que é linear:

$$y_{2k} = b_{2k} + \sum y_{1j} \times w_{2,j,k} \quad (4.2)$$

Os parâmetros  $b_{1j}$  e  $b_{2k}$ , são os *bias* enquanto  $w_{1ij}$  e  $w_{2jk}$  são os pesos da entre camada de entrada e intermediária e entre a camada intermediária de saída, respectivamente.

A função de treinamento adotada foi aquela denominada *trainbr*, que atualiza pesos e *bias* pelo método Levenberg-Marquardt com Regularização Baysiana, conforme descrito no Item 2.4.

### **Efeito do número de neurônios na camada intermediária**

A relação entre o número de neurônios na camada de entrada e a camada intermediária,  $n_i/n_e$ , foi tomada como parâmetro no estudo cujos resultados serão discutidos a seguir.  $n_s$  será o numero de neurônios na camada de saída.

Pelo critério de Hagan e Foresee (1977) deve-se aumentar o número de parâmetros iniciais da rede, pesos e *bias*, aumentando-se o número de neurônios intermediários até que o número de parâmetros sobreviventes ao treinamento não acompanhe mais o número inicial. O algoritmo Levenberg-Maquardt usado elimina parâmetros supérfluos ao longo das iterações. Em uma rede com uma camada intermediária o número de parâmetros iniciais  $P$  será dado por:

$$P = n_i (n_e + 1) + n_e(n_s + 1) \quad (4.4)$$

Na Equação 4.4,  $n_e$ ,  $n_e$  e  $n_s$  indicam o número de neurônios das camadas de entrada, intermediária e de saída, respectivamente.

A Figura 4.1. mostra o que acontece com a relação  $p/P$  em função do número de neurônios na camada intermediária,  $n_i$ .

Observa-se na Figura 4.1 que, de fato o treinamento tende a eliminar parâmetros indicando que a rede não deve ser aumentada além da condição  $n_i = n_e$ .

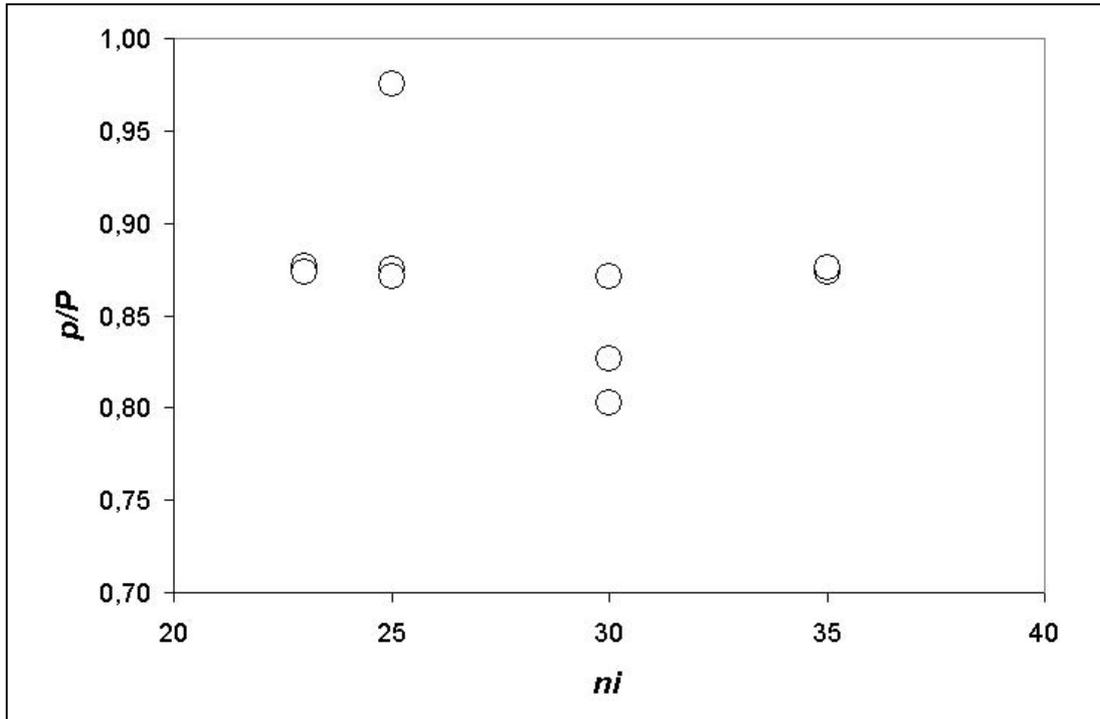


Figura 4.1 – Evolução da relação entre número de parâmetros sobreviventes/parâmetros iniciais e número de neurônios na camada intermediária

Outro critério para definir o número de neurônios camada intermediária é o que verifica a evolução do erro quadrático médio com esta variável. A Figura 4.2 apresenta esta relação para o grupo de vetores de treinamento e de generalização. No cálculo do erro quadrático médio os valores das variáveis foram padronizados com a equação (3.6).

Observa-se que o erro decresce com o aumento da relação para o grupo de treinamento, mas cresce para o grupo de generalização. Esse critério recomenda a adoção do valor de 23 para  $n_i$ , ou seja, uma rede na qual o número de neurônios na camada intermediária seja igual ao número de neurônios na camada de entrada.

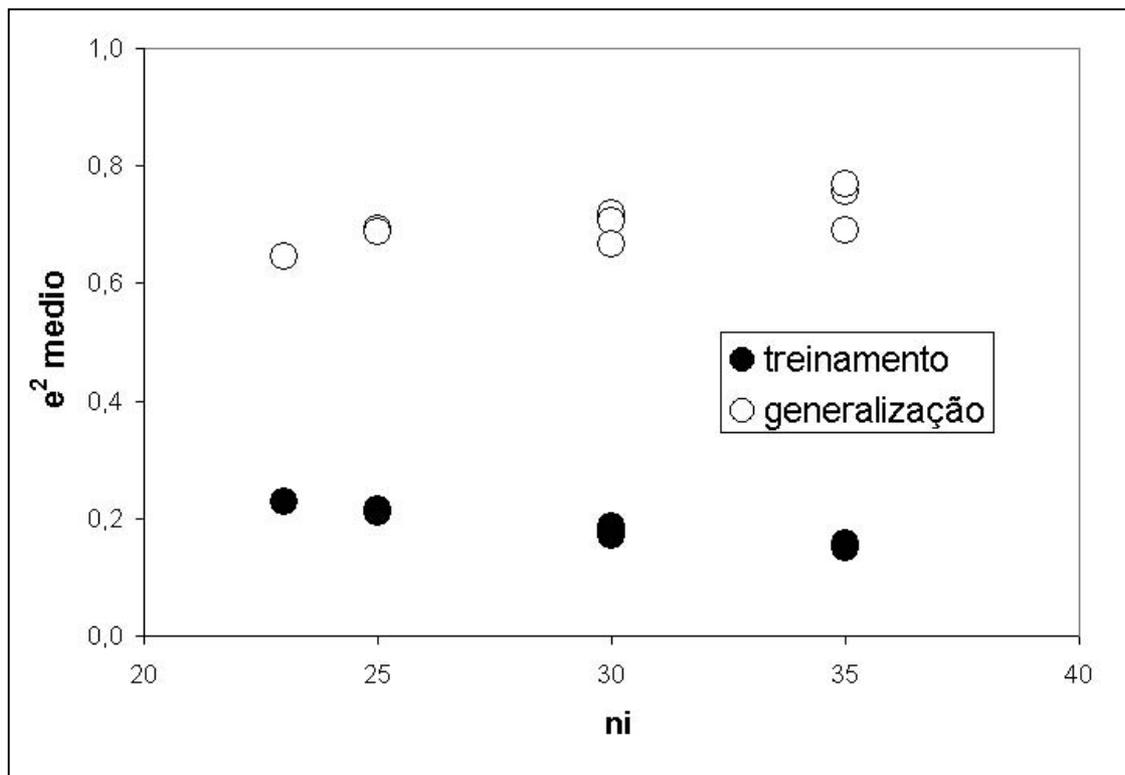


Figura 4.2 – Relação entre o erro quadrático médio e o número de neurônios na camada intermediária

Esses resultados indicam que a arquitetura  $23 \times 23 \times 5$  deve ser a adotada para o modelo neural.

### Resultados com a rede de arquitetura adotada

Os três resultados com a arquitetura  $23 \times 23 \times 5$  foram bem consistentes, como pode ser visto na Tabela 4.1 que contém os valores que geraram as Figuras 4.1 e 4.2.

No Anexo II encontram-se os pesos e *bias* determinados pelo treinamento da rede. Com estes parâmetros e algoritmos de balanços de massa e térmicos foi construído um

aplicativo em Excel ( Anexo III) para a orientação da equipe técnica para tomada de decisão na operação do alto-forno conforme Figura 3.8 e 3.11.

Tabela 4.1 – Resultados do treinamento com diversas arquiteturas de rede.

Ni	P	p	$e^2_{\text{medio}_t}$	$e^2_{\text{medio}_g}$
23	672	589	0,228	0,646
		588	0,228	0,646
		587	0,228	0,646
25	730	639	0,210	0,693
		639	0,215	0,687
		712	0,211	0,687
30	875	762	0,171	0,713
		723	0,187	0,718
		702	0,184	0,707
35	1020	893	0,150	0,757
		891	0,152	0,770
		891	0,157	0,691

Os erros quadráticos médios para cada variável de saída encontram-se na Tabela 4.2. Observa-se que as variáveis melhor estimadas no treinamento o são também na generalização

Tabela 4.2 – Erros quadráticos médios para as variáveis de saída.

	$\omega$	$\rho$	[Si]	$\lambda$	Ks
treinamento	0,253	0,161	0,314	0,131	0,282
generalização	0,956	0,514	0,716	0,413	0,613

Para a rede com a configuração selecionada foram traçados gráficos que comparam os valores das variáveis reais com os valores estimados. Neste caso a normalização foi a esma adotada no treinamento, equação (4.1). Os gráficos

podem ser vistos na Figura 4.3. Observa-se uma boa aderência entre dos pontos à diagonal dos diagramas.

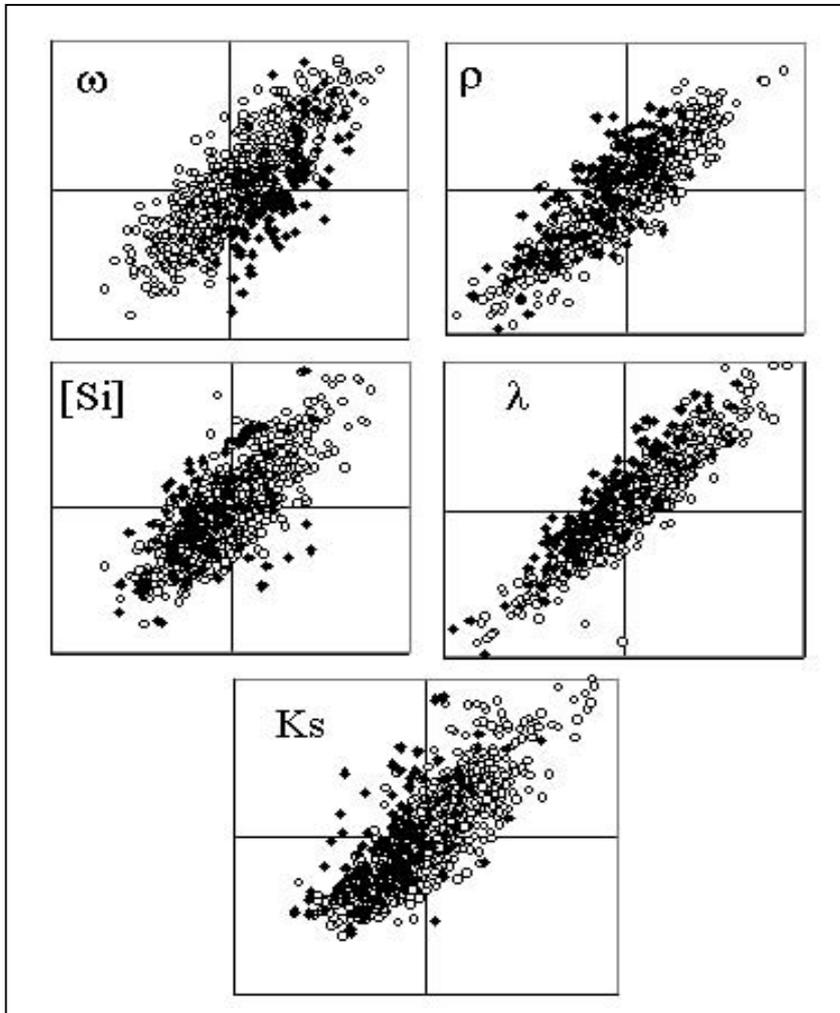


Figura 4.3 Relação entre o valor estimado no eixo horizontal, e valor real no eixo vertical, para dados do treinamento (o) e da generalização (◆).

Em todos os eixos a escala é de 0 a 1.

### **Análise da contribuição de cada variável**

A contribuição de cada variável de entrada na variação das variáveis de saída foi quantificada através do diferencial de cada variável de saída e o intervalo  $X_i \pm 1$ , ou seja 2 desvios-padrão de  $X_i$ .

Tabela 4.3 – Diferenciais das variáveis de saída em função  $X_i$ 

	$\omega$	$\rho$	[Si]	$\lambda$	Ks
Variáveis do sopro					
X1	0,162	-0,915	0,813	0,617	0,634
X2	0,090	0,084	0,544	-0,460	0,684
X3	0,033	-0,100	-0,328	0,212	-0,178
X4	-0,636	0,217	0,174	0,555	0,439
X5	-0,129	-0,104	0,105	-0,192	-0,043
X6	-0,434	-0,442	0,460	0,327	0,565
Variáveis da carga					
X7	0,208	-0,056	-0,084	-0,028	-0,057
X8	1,396	5,649	-6,441	6,505	-5,440
X9	-1,365	-4,274	5,319	-5,817	4,809
x10	-0,225	0,089	1,186	-1,639	1,063
X11	0,154	-0,295	0,909	-0,819	0,120
x12	0,232	1,406	-1,412	1,363	-1,108
x13	-0,184	-0,163	-0,092	0,010	-0,170
X14	0,440	0,304	-0,091	-0,049	0,395
X15	-0,255	-0,337	-0,036	0,065	-0,182
x16	0,027	-0,185	0,191	0,192	0,142
x17	0,051	0,355	0,126	-0,057	0,161
x18	0,209	0,236	-0,087	-0,072	-0,227
X19	0,022	-0,029	0,224	0,056	0,074
X20	-0,357	0,022	0,221	0,029	-0,143
X21	0,100	-0,173	-0,132	0,056	0,074
Variáveis do equipamento / ambiente					
x22	0,132	0,104	0,051	-0,057	-0,117
x23	0,138	-0,297	0,324	-0,255	0,039

Destacam-se como variáveis de forte influência  $x8$  e  $x9$ , basicidades da escória final e primária, respectivamente. Na prática, no entanto, essas variáveis não terão alto potencial de controle uma vez que há uma forte covariância entre elas, embora teoricamente seja possível fazê-las variar de forma independente. Neste caso deve-se observar o diferencial médio.

A maior concentração de nitrogênio,  $X_1$ , deve ser interpretada operacionalmente como uma menor taxa de enriquecimento do sopro com oxigênio. Como esperado aumenta a perda térmica e o teor de silício do gusa.

O desvio estequiométrico da wustita,  $\omega$ , é beneficiado pela redução do nitrogênio no sopro, pela injeção de carvão pulverizado e pelo maior teor de oxigênio no carvão. Das variáveis da carga, as que ajudam nesta variável são: mais pelotas e menos hematita e fundentes na carga. O sínter do estoque ajudou neste aspecto. Furo de gusa novo e chuva são prejudiciais assim como o maior tamanho médio do coque o uso de *small-coque*.

A resistência fluidodinâmica,  $\rho$ , foi maior para um menor enriquecimento com oxigênio. A maior basicidade tende a elevá-lo assim como o uso de *small-coke* e maior taxa de injeção de carvão. A maior reatividade do coque ajudou na permeabilidade do forno. Mais hematita na carga significou menor permeabilidade enquanto o de pelotas significa o contrário. Mais quartzito também indicou elevação dessa variável, enquanto o coque externo significou sua redução. O maior teor de oxigênio no carvão injetado também reduziu a permeabilidade. A maior fração de finos na hematita reduziu também a permeabilidade.

O teor de silício do gusa,  $[Si]$ , deve estar dentro de uma faixa aceita pela aciaria. Em geral a fuga acontece para valores maiores, o que representa maior consumo de energia nos altos-fornos. Mais nitrogênio, ou seja, menos oxigênio no sopro, aumenta o teor de silício no gusa enquanto a umidade do sopro o reduz. A maior basicidade também tende a reduzir o silício, mas o uso de hematita o eleva assim como o uso de pelotas. Mais quartzito significa, na verdade, menos silício no gusa. O maior teor de oxigênio no carvão eleva o teor de silício enquanto a menor resistência do coque o diminui. A fração maior que 38 mm na hematita eleva esse teor.

As perdas térmicas na zona de elaboração do alto forno,  $\lambda$ , foram aumentadas com o menor enriquecimento do sopro com oxigênio e reduzidas pela maior umidade do sopro. A maior basicidade aumenta as perdas térmicas mas o uso de small-coke tende a diminui-las. As perdas térmicas são maiores para maiores taxas de injeção de carvão e menores para cargas com mais hematita e mais pelotas. Mais quartzito aumenta as perdas térmicas assim como o teor de oxigênio do carvão injetado.

O coeficiente de partição do enxofre, ou seja, o menor teor de enxofre no gusa,  $K_s$  é beneficiado por menor enriquecimento e maior umidade no sopro. A maior basicidade da escória reduz o enxofre do gusa. Mais *small-coke*, mais carvão pulverizado, assim com mais hematita e mais pelotas associam-se a menor enxofre no gusa. O uso de quartzito tende a aumentar o enxofre do gusa. O maior teor de oxigênio no carvão e a maior resistência do coque ajudam a reduzir o teor de enxofre do gusa. Coque de menor tamanho aumenta o enxofre do gusa.

Todos os diferenciais observados tem a direção esperada com exceção da fração maior que 38 mm e o desvio estequiométrico da wustita,  $\omega$ . Essa variável depende da redutibilidade do minério e de sua granulometria. A discrepância pode estar relacionada à mineralogia do minério. Minérios de mesma composição mineralógica têm maior facilidade para serem reduzidos se tiverem tamanho médio menor mas o resultado indica que há um fator mineralógico associado à granulometria. Lotes de minério de maior redutibilidade parecem estar associados a minérios com maior fração de grossos.

Se comprovada esta hipótese este trabalho terá realizado outra contribuição inesperada, um dos méritos da aplicação das redes neurais artificiais como discutido no capítulo 2.

## 4.2 SIMULAÇÃO OFF-LINE

Para verificar a capacidade preditiva do modelo e sua aplicabilidade como ferramenta de planejamento operacional e de assistência à operação industrial foi conduzido no mês de setembro/2005 no AF3 da CSN, utilizando o aplicativo em Excel elaborado (Anexo III).

Quatro variáveis do processo foram selecionadas para a avaliação: o *coke-rate* ( $CR$ ), o índice de resistência de Ergun ( $K$ ), o teor de silício [ $Si$ ] e o teor de enxofre do gusa, [ $S$ ].

É prática comum nos altos-fornos fixar-se a taxa de injeção de combustível auxiliar pelas ventaneiras e ajustar o nível de aporte de energia variando a carga específica de coque. Como a resposta do forno a variações do *coke-rate* necessitam de cerca de 4 horas para que o forno sinta o seu efeito, faz-se simultaneamente o ajuste no coque e outro equivalente na taxa de injeção. Quatro horas depois a taxa de injeção é trazida de volta ao seu valor inicial. Mas de qualquer modo prever o nível de coque significa prever a demanda térmica total, daí a escolha dessa variável como indicador do desempenho do planejamento.

Como discutido no Capítulo 3, o índice  $\rho$  apresenta algumas vantagens em relação ao tradicional  $K$ . Como ainda é este último o mais usado na prática resolveu-se prever  $\rho$ , achar o correspondente  $K$  e comparar ao valor real.

O teor de silício, [ $Si$ ], é previsto diretamente pelo modelo.

O teor de enxofre, [ $S$ ], é o que realmente se deseja prever ao se buscar o coeficiente  $K_s$ . Com a posse desse último e da composição da carga é possível realizar o balanço de enxofre e calcular seus teores no gusa e na escória.

Os valores de coke-rate estimado e real no período de teste podem ser observados na Figura 4.4.

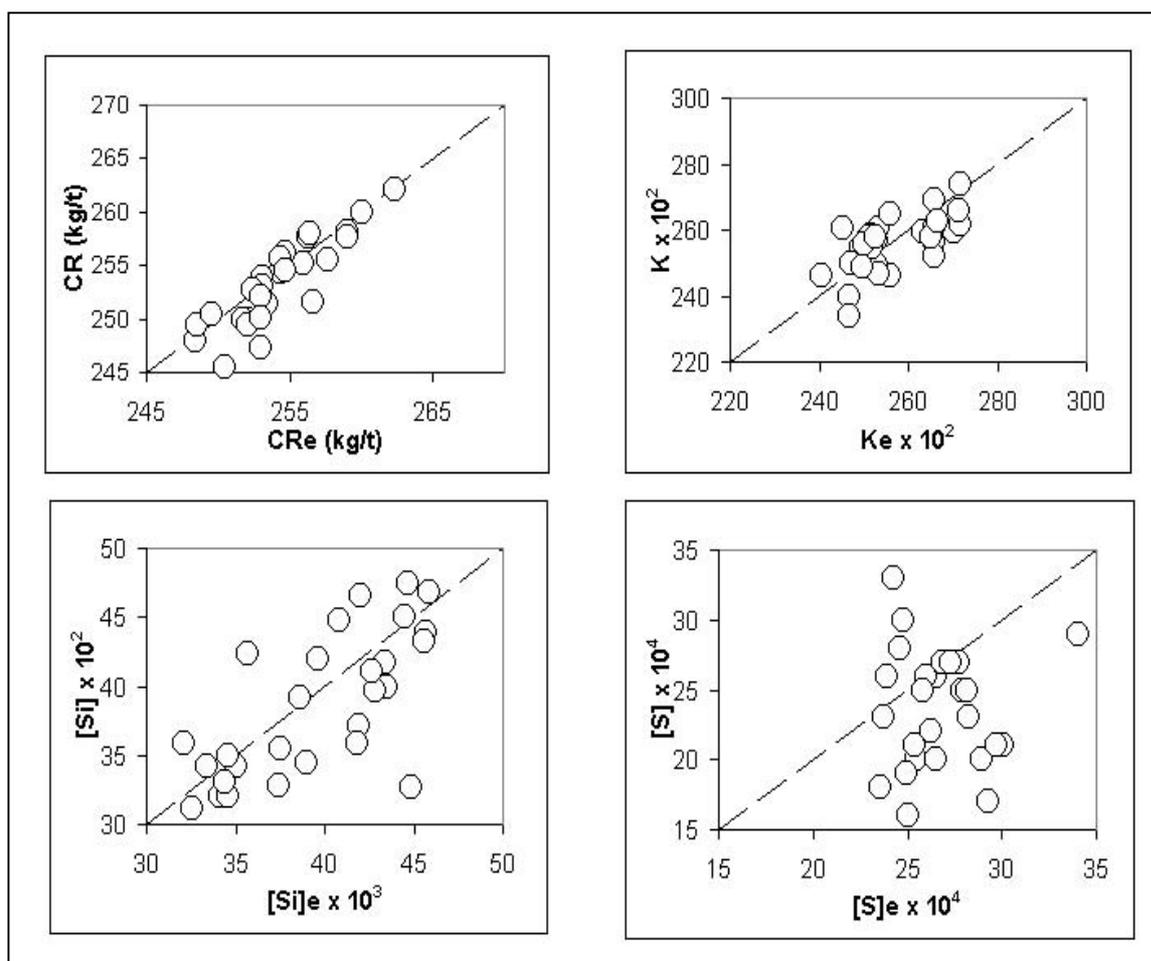
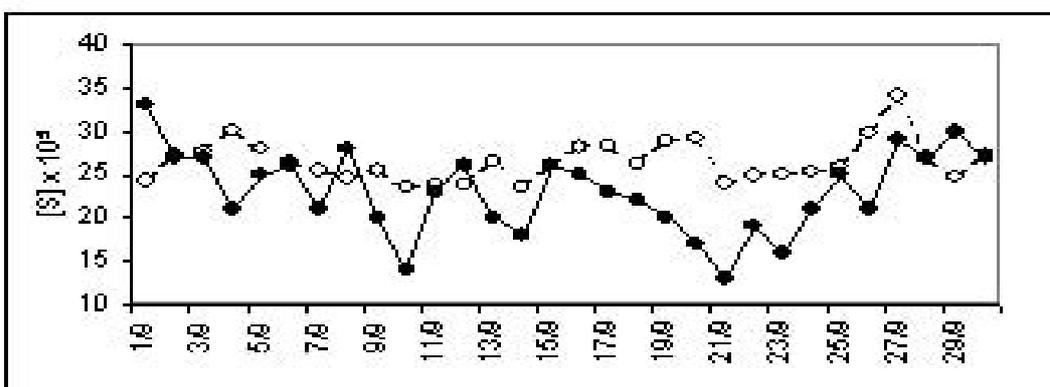
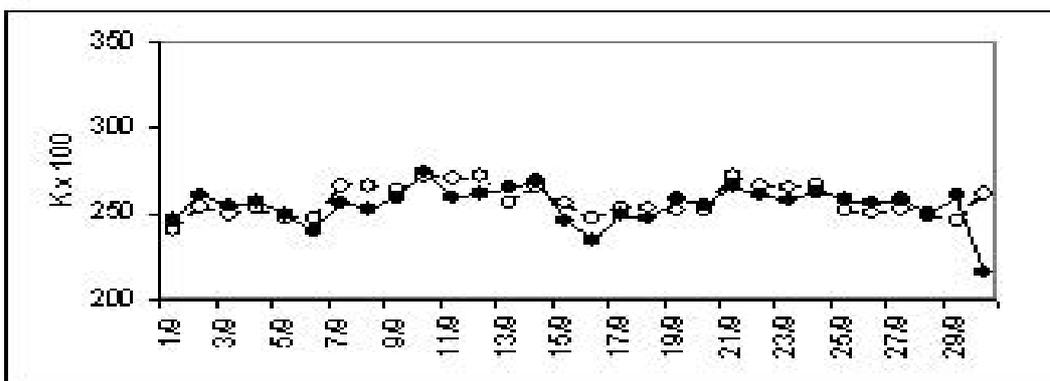
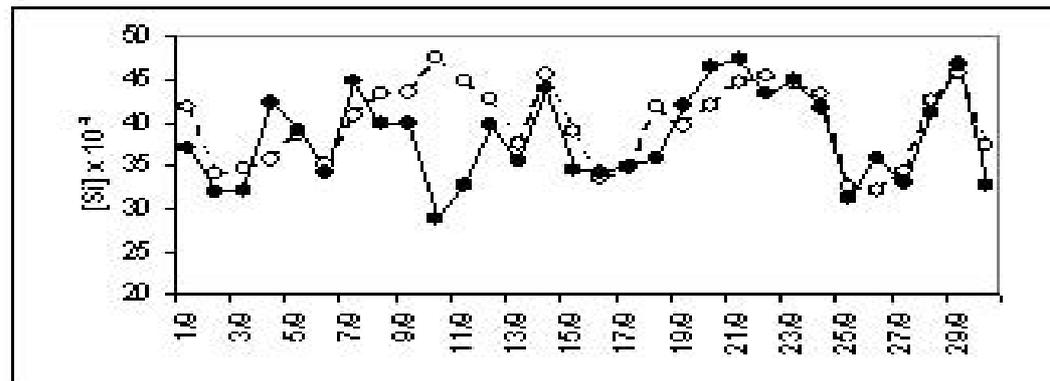
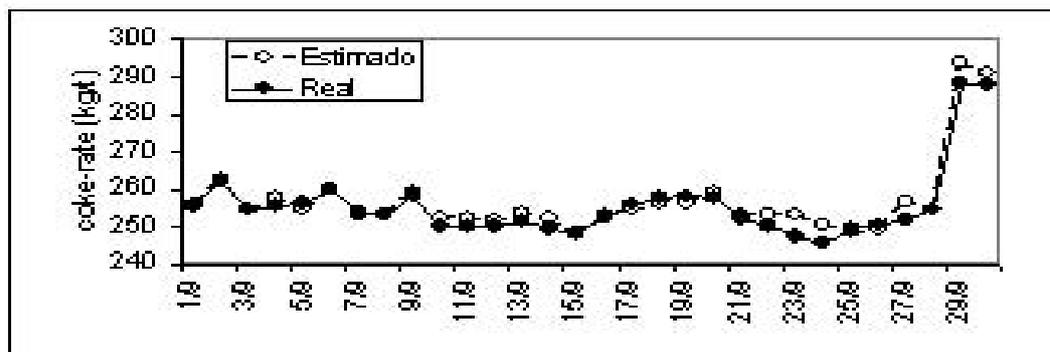


Figura 4.4 - Comparação entre valores estimados e reais para as variáveis de controle no teste industrial.

Observa-se boa concordância no caso do *coke-rate*, da resistência e do teor de silício e discrepância no caso enxofre.



Na Figura 4.5 esses resultados são observados ao longo do período do teste..

No caso do teor de enxofre dois períodos com comportamentos distintos são observados: na primeira quinzena do mês houve grande flutuação e outro no qual o valor real se mostra bem inferior ao estimado. Esse fato levantou a hipótese de algum problema com a representatividade da variável. De fato, foi detectada uma alteração nos procedimentos do laboratório neste período em função da elevada variação que o teor de enxofre vinha apresentando. Na tentativa de corrigir este problema um método alternativo de análise foi introduzido, o que levou a uma tendência para valores inferiores aos reais. Aqui o modelo serviu a mais um propósito: indicar problemas na geração de dados.

### 4.3 TESTE INDUSTRIAL

Findo o período de avaliação da capacidade preditiva do modelo iniciou-se um período de avaliação dos resultados de sua aplicação no planejamento semanal da operação do alto-forno. Foram definidas oito semanas com início em 2 de outubro de 2005. Nas quatro semanas iniciais o modelo não influenciou nas decisões operacionais, servindo de referência para avaliação dos resultados. Nas quatro semanas seguintes, período que começou em 30 de outubro de 2005, o modelo orientou o planejamento, visando menor necessidade de ações corretivas na operação.

Seis parâmetros diários foram observados: produção de gusa,  $PR$ , em toneladas/dia, o *fuel-rate*,  $FR$ , ou consumo específico de redutores em kg/tonelada de gusa, o desvio-padrão da vazão de sopro,  $sVs$ , em  $m^3$ /minuto, e dos teores de silício e de enxofre do gusa. e o índice  $K$ , resistência fluidodinâmica de Ergun, na Equação 2.2.

O *Fuel-rate* é a soma das taxas de consumo de todos os redutores, coque, *small-coke*, carvão e gás natural. No caso de planejamento fixam-se as taxas de injeção de carvão e de carga de *small-coke* sobrando o *coke-rate* como variável a ser estimada. Neste caso, em que se verifica o desempenho do processo como um

todo é mais conveniente adotar o *fuel-rate* como indicador de consumo de energia e de custo.

Desses parâmetros, os dois primeiros são os mais importantes. Os três seguintes, estão relacionados à regularidade do processo. São parâmetros de controle secundário em relação aos dois primeiros. O valor do terceiro segundo parâmetro, *K*, não é manipulável pelo modelo, depende da qualidade das matérias-primas. O modelo deve apenas prever o seu valor para que as contra-medidas tomadas sejam adequadas e não redundem em prejuízo do custo e da produtividade. Neste caso, esse parâmetro serve para avaliar o grau de dificuldade que a operação teve no período.

A Tabela 4.4 verifica a significância da hipótese nula para médias, (produção, *fuel-rate* e índice *K*) e dos desvios-padrão (sopro, teor de silício e teor de enxofre) através da distribuição de Student e de Fisher, respectivamente.

Tabela 4.4 – Teste de significância da hipótese nula para os parâmetros comparativos entre os períodos sem e com suporte do modelo

	Produção (t/dia)	Fuel-rate (kg/t)	Sopro (m <sup>3</sup> /min)	Teor de Si (10 <sup>-4</sup> )	Teor de S (10 <sup>-5</sup> )	K
Sem	10277	506	111	11,7	5,6	2,34
Com	10554	496	99	10,3	7,3	2,40
s médio	53	0,9	-	-	-	0,01
t	5,26	-11,3	-	-	-	4,45
s <sub>1</sub> <sup>2</sup> /s <sub>2</sub> <sup>2</sup>	-	-	1,49	1,07	1,24	-
P(H0)	<0,001	<0,001	0,85	0,57	0,71	<0,001

A figura 4.6 apresenta os resultados em forma de gráfico.

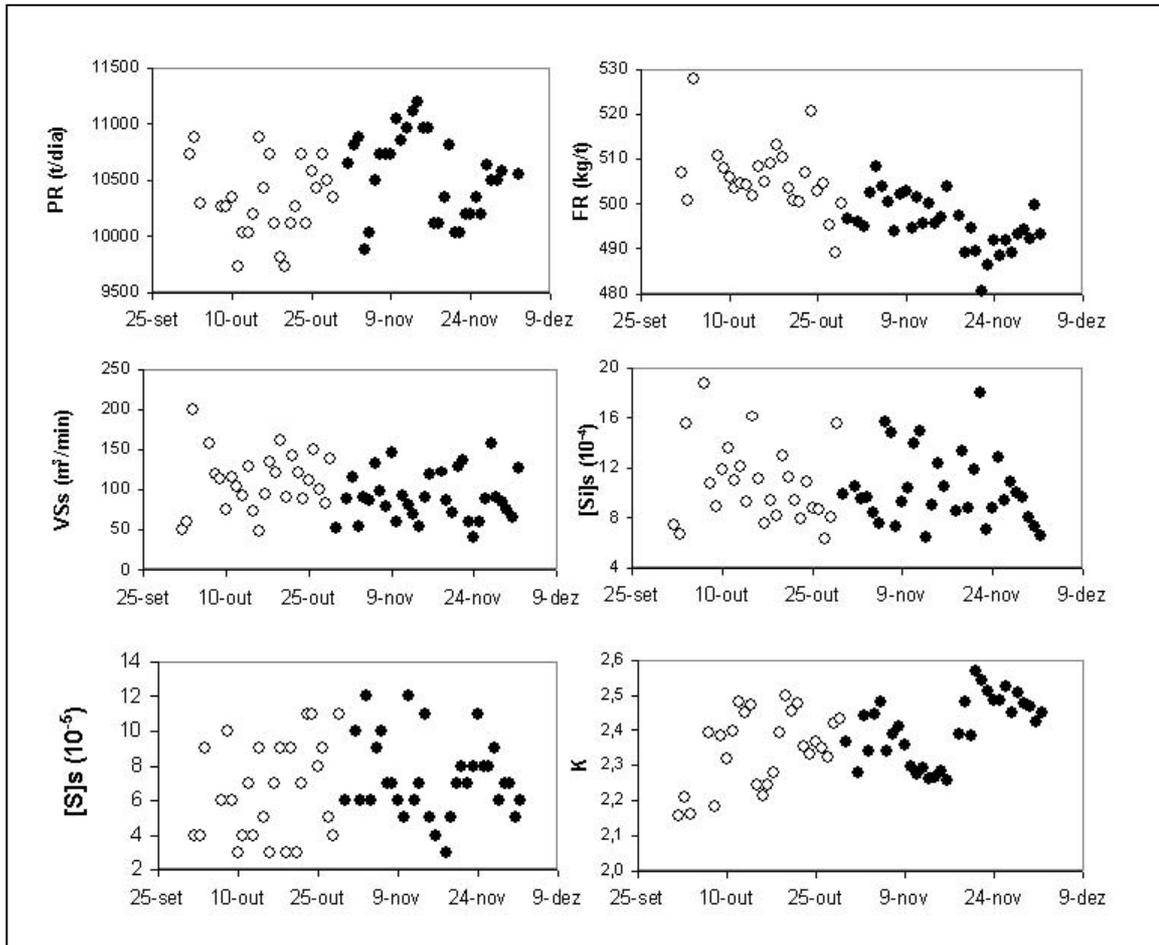


Figura 4.6 – Comparação entre resultados dos períodos sem suporte do modelo (o) e com suporte do modelo (●).

É possível concluir que, no teste industrial onde foram comparadas 4 semanas de operação sem suporte do modelo e 4 semanas com suporte do modelo, no período com suporte do modelo o forno mostrou maior produtividade com menor consumo específico de redutores, apesar de operar com matérias primas que produziram menor permeabilidade média do leite. Do ponto vista do sopro, a operação foi mais estável com uma pequena tendência a menor consistência de qualidade química. Esta última tendência mostrou pequena significância estatística e irrelevância diante da especificação do produto.

## CAPÍTULO 5 - CONCLUSÕES E SUGESTÕES

Este trabalho aplicou redes neurais artificiais ao processo de produção de ferro-gusa em alto-forno com o objetivo de compor um modelo híbrido de processo capaz de prever o seu desempenho fluidodinâmico e termo-químico a partir de características das matérias primas e do sopro de ar quente.

A componente fenomenológica do modelo consiste de um balanço de carbono hidrogênio, oxigênio e ferro e um balanço térmico que dependem de índices de desvio de equilíbrio químico e térmico. Esses índices constituem variáveis de saída de uma rede neural. A mesma rede neural permite prever o teor de silício do gusa e o coeficiente de partição do enxofre entre metal e escória. Somados ao índice de resistência fluidodinâmica, a rede gera a estimativa para 5 parâmetros capazes de definir o estado do processo.

Vários bancos de dados dão origem ao banco utilizado neste estudo, entre eles os arquivos do sistema digital de automação e controle de processo do Alto-Forno #3 e dos laboratórios da Usina Presidente Vargas e da Mineração de Casa de Pedra, de onde vem o minério bitolado. Cada vetor de entrada corresponde a um dia de operação. Como cada um dos bancos de dados tem uma base temporal diferente, constituiu parte fundamental deste trabalho a consolidação do banco de dados final. As variáveis de entrada foram 23.

A fim de encontrar a arquitetura de rede associada à melhor estimativa, a somatória dos erros quadráticos entre valores reais e estimados foi avaliada para redes neurais com as seguintes características: intervalo de normalização de  $[0;1]$ ; funções de ativação log-sigmoidal e linear entre as camadas intermediária e final, respectivamente; e o algoritmo de treinamento de *Levenberg-Marquardt* com regularização bayesiana.

O algoritmo utilizado apresentou bom desempenho evitando o sobre-ajuste na estimativa, comprovado pelos bons resultados na generalização.

Os resultados indicaram a adoção de uma rede 23 x 23 x 5 que apresentou o menor erro na generalização.

Dos cinco índices, três,  $(\omega, [Si], \lambda)$  são combinados com o índice  $\mu$ , calculado através na componente fenomenológica do modelo, e outras variáveis de entrada para gerar dois índices:  $\theta$  e  $\phi$ . O primeiro indica o desempenho fluidodinâmico do forno e o segundo o seu desempenho termoquímico. O Produto dos dois,  $\pi$ , cuja dimensão é massa de ferro-gusa, por unidade de tempo por unidade de área da seção do cadinho do forno constitui o índice de produtividade global do processo.

Simulações “off-line” conduzidas no mês de setembro/2005 no AF3 da CSN mostraram bom desempenho preditivo do modelo para as variáveis: consumo de coque, teor de silício, resistência ao fluxo gasoso. Para o teor de enxofre a predição não teve menor intensidade de variação que o valor real na primeira metade do período e apresentou desvio da média na segunda. Os fatores que prejudicaram parte das simulações relativas ao teor de enxofre do gusa chamaram a atenção da equipe de processo que detectaram não-conformidades na análise do gusa neste item. Essa irregularidade foi corrigida logo após de relatada. Esse fato realça outra vantagem de se aplicar um modelo de processo dessa natureza na rotina de operação.

No teste industrial, onde foram comparadas 4 semanas de operação sem suporte do modelo e 4 semanas com suporte do modelo, o forno mostrou maior produtividade com menor consumo específico de redutores no período com suporte do modelo, apesar de operar com matérias primas que produziram menor permeabilidade média do leito. Do ponto vista do sopro, a operação foi mais estável com uma pequena tendência a menor consistência de qualidade química. Esta última tendência mostrou pequena significância estatística diante da especificação do produto.

A aplicação da técnica de redes neurais artificiais mostrou-se efetiva para a solução do problema.

Recomenda-se aumentar o grau de automação da aquisição de dados formadores do banco de dados do modelo para maior potencial de resultados com uso do modelo.

## NOMENCLATURA

<b><i>a</i></b>	vetor dos parâmetros de ajuste de uma rede neural
<i>a1ij</i>	função de ativação entre camada de entrada e intermediária
<i>a2jk</i>	função de ativação entre camada intermediária e de saída
<i>b1ij</i>	<i>bias</i> da função de ativação entre camada de entrada e intermediária
<i>b2ik</i>	<i>bias</i> da função de ativação entre camada intermediária e de saída
<i>CR</i>	<i>coke-rate</i> ( consumo específico de coque no alto-forno) kg/t
<i>dp</i>	tamanho das partículas
<i>E</i>	matriz de erros de uma rede neural
$e^2$	perro quadrático das estimativas da rede neural
<i>FR</i>	<i>fuel-rate</i> ( consumo específico de reductores no alto-forno) kg/t
<i>GR</i>	<i>gas-rate</i> ( consumo específico de gás natural no alto-forno) kg/t
<i>Gg</i>	velocidade modular mássica dos gases
<i>gc</i>	fator gravitacional
<i>I</i>	matriz identidade
$J^T$	matriz jacobiana transposta
<i>K</i>	resistência fluidodinâmica na equação e Ergun simplificada
$k_x$	taxa de geração ou consumo da substância <i>X</i>
<i>Ks</i>	coeficiente de partição de enxofre do modelo híbrido
$\Delta H$	elemento do balanço térmico do modelo híbrido
<i>ne</i>	número de neurônios na camada de entrada da rede neural
<i>ni</i>	número de neurônios na camada intermediária da rede neural
<i>ns</i>	número de neurônios na camada de saída da rede neural
<i>P</i>	número máximo de parâmetros na rede
<i>PCR</i>	<i>coal-rate</i> ( consumo específico de carvão injetado no alto-forno) kg/t
<i>Pg</i>	pressão total dos gases
<i>Ps</i>	pressão absoluta de sopro
<i>Pt</i>	pressão absoluta de topo
<b><i>p</i></b>	vetor de incrementos no algoritmo Levenberg-Maquardt
<i>Q</i>	matriz auxiliar no algoritmo Levenberg-Maquardt
[ <i>S</i> ]	teor de enxofre do gusa (%)

$[Si]$	teor de silício do gusa (%)
$T_s$	temperatura de sopro (K) - variável
$T_c$	temperatura da carga na linha de Boudourd – fixa em 1250 K
$T_g$	temperatura do gás na linha de Boudouard – fixa em 1300 K
$T_{gu}$	temperatura do gusa (K) – variável
$T_{es}$	temperatura da escória (K) – variável e igual a $T_{gu} + 50K$
$V_{gv}$	vazão de gás de ventaneiras
$w_{ji}$	peso de ordem $j$ da camada $i$ de uma rede neural
$w_{1ij}$	pesos da função de ativação entre camada de entrada e intermediária
$w_{2jk}$	pesos da função de ativação entre camada intermediária e de saída
$X_i$	variável de entrada em dimensões originais
$x_i$	variável de entrada de uma rede neural
$x_{ni}$	variável de entrada normalizada em função da amplitude
$(X)$	teor da substância $X$ na escória
$[X]$	teor da substância $X$ no metal líquido
$x_X$	fração molar da substância $X$
$y_i$	variável intermediária de uma rede neural
$Y_j$	variável de saída em dimensões originais
$y_{nj}$	variável de saída normalizada em função da amplitude
$z$	coordenada vertical
$\varepsilon$	índice de vazios
$\delta$	potencial de pressão do modelo híbrido
$\Phi$	fator de forma
$\phi$	índice de desempenho fluidodinâmico do modelo híbrido
$\gamma$	número de kg-moles de Fe em 1 tonelada de gusa
$\lambda$	índice de perdas térmicas do modelo híbrido

$\mu$	número de kg-átomos de C + H <sub>2</sub> no gás por tonelada de gusa
$\mu_g$	viscosidade dos gases
$\nu$	número de kg-moles de N <sub>2</sub> no gás por tonelada de gusa
$\pi$	índice de produtividade do modelo híbrido
$\rho$	índice de resistência fluidodinâmica média da carga
$\rho_g$	densidade da fase gasosa
$\theta$	índice de desempenho termoquímico do modelo híbrido
$\omega$	índice de desvio estequiométrico da wustita do modelo híbrido

## SIGLAS

ABM	Associação Brasileira de Metalurgia e Materiais
IBS	Instituto Brasileiro de Siderurgia
IEEE	Institute of Electrical and Electronics Engineers
IISI	International Iron and Steel Institute
ISIJ	Iron and Steel Institute of Japan
VDEh	Verein Deutscher Eisenhüttenleute (Associação dos Siderurgistas Alemães)

## BIBLIOGRAFIA

1. **Abelém, A. J. ; Pacheco, M. A. C. & Vellasco, M. M. B. R.** – Modelagem de redes neurais artificiais para previsão de séries temporais – II Simpósio Brasileiro de Redes Neurais, p107-112. 1995.
2. **Ackley, D. H, Hinton, G. E. & Sejnowsky, T. J.** - A learning algorithm for Boltzmann machines. *Cognitive Science* 9, p 147 - 169. 1985
3. **Agryropoulos, S. A.** – Artificial intelligence in materials processing operations: a review and future directions – *ISIJ International*, vol 30, n 2, p 83-89. 1990.
4. **Anderson, J.** - A General introduction. In *Neurocomputing: Foundations of Research*. Mit Press. 1988.
5. **Atiya, A. & Ji, C.** – How initial conditions affect generalization performance in large networks – *IEEE transactions on Neural Networks*, vol 8, n 2, p 448-451. 1997.
6. **Barnaba, P.** - La qualità dell'agglomerato per i moderni altiforni. *Bolletino Tecnico Finsider*, maio-agosto, p 83-96. 1982.
7. **Barnard, E. & Holm, J. E. W.** – A compactive study of optimization techniques for backpropagation. *Neurocomputing*, n 6, p 19-29. 1994.
8. **Bell, I. L.** - Principles of the manufacture of iron and steel. George Routledge & Sons, London, p 184-195. 1984.
9. **Bogdandy, L. von & Engell H. J.** - The reduction of iron ores. Verlag Stahleisen m.b. H, Dusseldorf. 1971.
10. **Bouman, R. W.** - Historical development of the blast furnace. *Iron and Steelmaking*; Feb, pag 17-23. 1978.
11. **Chaigneau, R et alli** – Models to understand the blast furnace process. *La Revue de Metallurgie*, ATS JS 97, p 8-9.1997.
12. **Cheng, J.; Akiama, T. Nagomi, H., Takahashi, H. & Yagi, J.** - Modelling of solid flow in moving beds. *Transactions ISIJ*, March.1993.
13. **Coudurier, L ; Hopkins, D. W. & Wilkomirsky, I** – Fundamentals of Metallurgical Process – *International Series on Metallurgical and Technology* , vol 27. Pergamon, 1978
14. **da Cunha, A P.** – Redes neurais em processos siderúrgicos: analisador virtual de propriedades metalúrgicas e modelo de predição da qualidade do aço – Tese de Doutorado, FEQ/UNICAMP, 2001
15. **Dankly, G; Mignon, J; Munnix R.; Dauwels, G & Bonte, L** – Industrial Application of The Blast Furnace Model at Sidmar – 3<sup>rd</sup> International Conference on Science and Technology of Ironmaking, Dusseldorf, 2003
16. **Darken, I. S. & Gurry, R. W.** - The system iron-oxygen II. Equilibrium and thermodynamics of liquid oxide and other phases: *Journal of the American Chemical Society*, vol 68, p 798-816. 1946.
17. **de Castro, J. A .** – Modelamento matemático bidimensional do alto-forno para análise de processo. Tese de mestrado em metalurgia. Universidade Federal Fluminense, Volta Redonda. 1996.
18. **de Castro, J. A .** – A Multi-Dimensional Transient Mathematical Model of Blast Furnace Based on Multi-Fluid Model – Tohoku University, Japan. 2000

19. **de Medeiros, F.T. P. et alli.** - Injeção de gás natural nos altos-fornos da CSN - XLII Congresso Anual da Associação Brasileira de Metais, Belo Horizonte. 1987
20. **Eaton, A. C. H. & Olivier, T. L.** – Learning coefficient dependence on training set size – Neural Networks, vol 5, p 283-288. 1990.
21. **Ergun, S.** - Fluid Flow through Packed Columns. Chemical Engineering Progress, vol 48, n 2, p 89 - 94. 1952
22. **Fileti, A M. F.** – Controle em destilação em batelada. Controle adaptativo e controle preditivo com modelo baseado em redes neurais artificiais – Tese de Doutorado, FEQ/UNICAMP, 1995.
23. **Foresee, F. D. & Hagan, M. T.** – Gauss-Newton approximation to Bayesian regularization . Proceedings of the 1997 International Joint Conference on Neural Networks, 1997.
24. **Ge, A. X.** - A Neural Network Approach to the Modelling of Blast Furnace, Master of Engineering in Electrical Engineering and Computer Science Thesis, Massachusetts Institute of Technology. Boston, 1999.
25. **Geiger, G. H. & Poirer, D. R.** - Transport phenomena in metallurgy. Adison-Weley Publishing Company. 1973.
26. **Gudenau, H. W ; Kreibish, K & Peters, K. H** – Stahl und Eisen, vol 101, p 147, 1981.
27. **Hagiwara, M** – Backpropagation with selection-reduction of learning time and elimination of hidden units – Systems and computing in Japan, vol 23, n 8 , p 46-54. 1992.
28. **Hartig, W; Langner, K; Lungen; H. B. & Stringer, K. P.** - Measures for increasing the productivity of blast furnaces. MPT International, vol 4, p 42-51. 1996.
29. **Haykin, S.:** Neural Networks, A comprehensive foundation, Macmillan, New York. 1980.
30. **Hebb, D.** – The organization of behaviour . John Wiley, New York. 1949.
31. **Hernandez, C. A; Espi, J. & Nakayama, K.** – Effects of training data distribution on backpropagation geeralization – IEEE International Symposium on Circuits and Systems, vol 4 p 2148-2151. 1003.
32. **Hecht-Nielsen, R** – Neurocomputing, Addison-Wesley Publishing Company. 1990.
33. **Heynert, G. ; Peters, K. H. & Ringkloff, G.** - Five years experience with the bell-less top. Iron and Steelmaker, vol 5, n 3, p 15-24. 1978.
34. **Hinton, G. E. & Sejnowski, T. J.** - Learning and relearning in Boltzmann machines. In parallel distributed processing; Explorations in microstructure of cognition. Cambridge, MA. MIT Press. 1986.
35. **Holowaty, M. O** – Blast Furnace Theory and Practice, Chapter 1 – Gordon and Breach, 1969.
36. **Hopfield, J. J.** – Neural networks and physical sysem with emergent collective computational abilities – Proceedings of National Academy of Sciences, vol 79, 1982.
37. **Instituto Brasileiro de Siderurgia,** Anuário 2002, 2003, 2004
38. **Iida, O.; Taniyoshi, S.; Uetani, T.; Sawada, T; Hashimoto, M & Onoda, D.** 1992:

39. **Kanbara, K. et alli**; - Dissection of blast furnace and their internal state, Transactions ISIJ, vol 17, p 371-380. 1977.
40. **Kawasaki Steel Technical Report**. Application of AI techniques to blast furnace operation, n 26, June, p 30-37.
41. **Kitaev, G. I., Yaroshenko, Yu, G. & Suckov, V. D.** Heat exchange in shaft furnaces., Pergamon Press. 1967.
42. **Koács, Z. L.** – Redes neurais artificiais: fundamentos e aplicações – Edição Acadêmica. 1996.
43. **Kuwubara, M. & Muchi, I.** - Mathematical model of blast furnace with horizontally layered burdens. Transaction ISIJ, vol 17, p 271-278.1976
44. **Lahiri, A. K. & Seshadri, V.** - An analysis of blast furnace process based on a dynamic model, Journal of the Iron and Steel Institute, p 293-307. 1969.
45. **Lehmen, A. Y.; Paek, E. G.; Liao, P. F.; Marrackchi, A & Patel J. S.** – Factors influencing learning by backpropagation – IEEE Transactions on Neural Networks, p 335-341. 1998.
46. **Leonard, J. & Kramer, M. A.** – Improvement of the backpropagation algorithm for training of neural networks – Computers Chem. Engineering, vol 14 n 3 p 337-341. 1990.
47. **Lighter, J. E** – Historical Dictionary of American English Slang. Randon House, 1994.
48. **Lippamn, R. P.** – IEE ASSP Magazine. April, p 4-22. 1987.
49. **MacCulloch, W. S. & Pitts, W.** – A logical calculus of the ideas immanent in nervous activity – Bulletin of Math. Bio, n 5, p 115-133. 1943.
50. **MacKay, D. J. C.** – Bayesian interpolation – Neural Computation, vol 4 no 3, p 415-447, 1992.
51. **Maki, A et alli.** High rate coal injection of 218 kg/t at Fukuyama no 4 blast furnace. ISIJ International, vol 36, n 6, p 650-657. 1996.
52. **Maren, A** – Handbook of neural computing applications, Academic Press, San Diego. 1990.
53. **Mauhart, J.; Pillmar, G.; Lawrence, P.; Spalek, K. ; Hoerl J. & Schuerz B.** – Effective Use of Automation at Blast Furnace A Vostaltpine and Blast Furnace 5 Iscor. – 3<sup>rd</sup> International Conference on Science and Technology of Ironmaking, Dusseldorf. 2003.
54. **Melgaço, H; Franco M. & de Medeiros F.T. P.** - Análise do processo de alto-forno pelo diagrama de trocas térmicas - Simpósio Comin/Comap, Associação Brasileira de Metais, Rio de Janeiro. 1982.
55. **Meraikib, M** – Silicon Distribution Between Blast Furnace Slag And Hot Metal – Ironmaking and Steelmaking, vol 27, p 280-285. 2000
56. **Minsky, M. & Pappert, S.** – Perceptrons – MIT Press, Cambridge, 1969.
57. **Mirchandani, G. & Cao, W.** – On hidden nodes for neural nets – IEEE Transactions on Circuits and Systems, vol 36, n 5, p 661-664. 1989.
58. **Muraviev, V. M. et alli**; - Stal in English , p 591. 1970.
59. **Omuri, Y (org) et alli** – Blast Furnace Phenomena and Modelling – ISIJ/Elsevier, 1982.
60. **Otsuka, K.; Matoba, Y.; Kojima, M; Yoshida, M & Matsuoka, S.** - Expert system for blast furnace operation. The Sumitomo Search, 50 July, p 43-50. 1992.

61. **Patankar, S. V.** - Numerical heat transfer and fluid flow. Mineapolis. McGraw-Hill. 1979.
62. **Pereira, I. C; Aguiar, S. S. & Peçanha, R. P.** - Permeabilidade de Meios Porosos: Dependência da Constante de Kozeny com a Forma das Partículas.
63. **Qiu, G.; Varley, M. R. & Terrell, T. J.** – Accelerated training of backpropagation networks by using adaptative momentum step – Electronic Letters, vol 28, n 4 p 377-378. 1992.
64. **Radestock, J. & Jeschar, R.** - Stahl und Eisen, n 90, p 1249. 1970.
65. **Reichardt, P:** Arch the Iron and Steel Institute, n 200, p 389. 1927.
66. **Rist, A. ; Michard, J & Meysson, N.** - Journal Eisenhüttenw, n 1 , p 77. 1967.
67. **Ridgion, J. M.:** Journal of Metals, Abril, p 50. 1962.
68. **Roseblatt, F.** – The perceptron: a probabilistic model for information storage and organization in the brain – Psychol Ver, v 65, p 386-408. 1958.
69. **Rummelhart, D & McClelland, J** – Parallel distributed processing – MIT Press, Cambridge, 1986.
70. **Russo, A. P.** - Neuronal networks for sonar processing, Tutorial no 8 IEE, Conference on Neural Networks for Ocean Engineering, Washington. 1991
71. **Sasaki, K et alli** - Investigation of quenched number 2 Blast Furnace at Kokura Works, Transactions ISIJ, vol 17, p 252 – 261. 1976.
72. **Schmidt, W. F.; Raudy, S.; Kraaijveld, M. S. ; Skurikhimana, M & Duin, R. P. W.** – Initializations, backpropagation and generalization of feedforward classifiers – IEEE trasnsations on Neural Networks, vol 8, n 2, p 598-604. 1993.
73. **Steeghs, A. G. S ; Rausch, H & Sert, D.** - Investigation of ore segregation during blast furnace charging with bell less top systems. La Revue de Metallurgie, março, p 323-332. 1997.
74. **Stull et alli** – Thermochemical Tables, United States Department of Commerce, NSRDS-NDS\37, 1971..
75. **Sugyama, T & Yamaguchi, I** – ISIJ n 9, p 50, 1994.
76. **Szekely, J.** - Fluid flow phenomena in metals processing. Academic Press, New York. 1979
77. **Tamura, S. & Tateishi, M.** – Capabilities of a four-layer feedforward neural networks: four layer versus three – IEEE Transactions on Neural Networks, vol 8, n 2, p 251-254. 1997.
78. **Tang, C. Z & Kwan, H. K.** – Parameter effects on convergence speed and generalization capability of backpropagation algorithm – International Journal of Electronics, vol 74, n 21, p 35-46. 1993.
79. **Tek, M. R.** - Development of Generalized Darcy Equations. Trans AIME 210, p 376 - 377. 1957.
80. **Toxopeus, H. L. et alli** – High productivity blast furnace operation in Ijmuiden. Relatório interno, Koninklijke Hoogovens. 1996.
81. **Villiers, J. & Barnard, E.** – Backpropagation neural nets with one and two hidden layers – IEEE Transactions on Neural Networks, vol 4, n 1, p 136-141. 1992.
82. **von Bogdandy, L & Engell, H. J.** – The reduction of iron ores. Verlag Stahleisen mbH. 1971.

83. **Waller, M & Saxon, H.** – Application of Nonlinear Time Analysis to the Prediction of Silicon Content of Pig Iron – ISIJ International, vol 42, p 316 – 318. 2002
84. **Wann, M.; Hediger, T. & Greenbaum, N.N.** – The influence of training sets on generalization in feedforward neural networks. p 137-142. 1990.
85. **Wasserman, P. D.** – Neural computing: theory and practice, ANZA Research, Inc, 1989.
86. **Widrow, B.** – Generalization and information storage in networks of adaline neurons – Self-Organization Systems, 1962.
87. **Wartmann, P.**- Archive des Eisenhüttenwesens, n 34, p 879. 1962
88. **Yagi, J. & Muchi, I.** Theoretical estimation on longitudinal distribution of process variables in blast furnace and its productivity. Transactions ISIJ, vol 10, p 392-405. 1970.
89. **Yagi, J** - ISIJ International, n 33, p 619, 1993.
90. **Zaimi, S. A.; Sert, D; Guillot, J. B & Biaisser, H.** – Recent Advances in the Modelling of Solid Flows in the Blast Furnace – 3<sup>rd</sup> International Conference on Science and Technology of Ironmaking, Dusseldorf. 2003

## Anexo Ia – Listagens de programas em Matlab 6.1 ®

---

### %simulação e otimização dos parâmetros da rede

```
clear;
M=load('Mtz2t.prn');           %carrega arquivo de dados para o treinamento
entrada=M(:,1:23);             %define quais são os dados de entrada da rede
saida=M(:,24:28);             % define quais são os dados de saída da rede
entrada=entrada';              % calcula matriz transposta
saida=saida';
entrada_total_n=normal(entrada);
saida_total_n=normal(saida);
entradan=entrada_total_n(:,1:690);
saidan=saida_total_n(:,1:690);
inn=entrada_total_n(:,691:820);
outn=saida_total_n(:,691:820);
out=saida(:,691:820);
saida=saida(:,1:690);
%define parametros maximos e mininos das matrizes de entrada e saída e faz
normalização;
%net.numinputs=size(entrada,1); %numeros de variáveis de entrada
net.numLayers=2;                %numero de camadas, sem considerar a camada
de entrada
net=newff(minmax(entradan),[30,5],{'logsig','purelin'},'trainbr');
net.trainParam.epochs=500;      %numero de passos
net.trainParam.show=25;         %atualiza no grafico o numero de passos
net.trainParam.goal=1e-4;       %convergencia desejada
net.initFcn='initlay';          %funcao que inicia os pesos e bias
net.performFcn='sse';           %função objetivo a ser minimizada
net.trainParam.min_grad=1e-100; %minimo gradiente
net.trainParam.mu_max=1e+100;   %max MU
net=init(net); % inicializacao dos pesos da rede aleatoriamente
%net=load('rede1.mat');
[net,tr]=train(net,entradan,saidan); %realiza o treinamento da rede
% foram determinados os pesos e bias da rede e arquivados em "net"
Y=sim(net,entradan); %simula com os dados de entrada do arquivo de treinamento
X=dnormal(Y,saida);            %desnormaliza dados de saída

%net.IW{1}          % mostra pesos entre camadas entrada/intermediária
%net.LW{2}          % mostra pesos entre camadas intermediária/saída
%net.b{1}           % mostra bias da camada intermediária
%net.b{2}           % mostra bias da camada de saída
```

## Anexo Ib – Listagens de programas em Matlab 6.1 ®

-----  
função para normalizar os dados entre 0.1 e 1

```
function YN=normal(M)
n=size(M,1);           %numero de linhas
m=size(M,2);           %numero de colunas
PR=[min(M,[],2) max(M,[],2)]; %matriz com valores mínimos e máximos
for i=1:n
    for j=1:m
        YN(i,j)=((M(i,j)-PR(i,1))*0.8)/(PR(i,2)-PR(i,1))+0.1;
        %loop para fazer a normalização
    end
end
end
```

-----  
%função para desnormalizar os dados

```
function Mm=dnormal(YN,M)
[R,Q]=size(YN);
n=R;%numero de linhas
m=Q;%numero de colunas
PR=[min(M,[],2) max(M,[],2)];%matriz com valores mínimos e máximos
for i=1:n
    for j=1:m
        Mm(i,j)=((YN(i,j)-0.1)*(PR(i,2)-PR(i,1)))/0.8+PR(i,1);
        %loop para desnormalizar
    end
end
end
%save desn.dat Mm -ascii;
```

---

Anexo II – Parâmetros da rede

Camada intermediária		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
w1	b1																								
1	-0,409	3,323	-1,193	-1,494	0,204	-2,364	-0,573	-0,735	0,443	0,712	-0,582	-0,853	1,115	-0,331	0,652	0,075	-0,047	-0,021	1,392	-0,083	-0,325	-0,235	1,862	0,571	
2	2,057	-1,055	-1,004	-0,924	2,172	-1,551	1,152	-0,219	-1,128	-0,320	0,564	0,932	0,486	1,291	0,021	-1,009	0,431	-0,440	0,194	1,075	0,259	-0,705	0,537	-0,002	
3	0,884	-0,420	-2,001	1,293	0,947	-0,509	-0,108	-0,878	-2,694	-0,292	-0,863	0,653	2,597	0,058	0,451	0,927	0,607	-0,299	1,967	2,895	1,787	-0,026	-0,524	0,056	
4	0,434	-0,482	0,661	-0,613	-1,562	-0,909	-0,087	2,076	2,872	0,172	-1,004	-0,647	-2,003	0,120	1,850	-1,200	0,511	0,729	-0,754	1,028	-1,431	0,655	0,977	0,775	
5	-0,928	0,380	2,033	0,205	2,420	-1,016	0,765	1,348	1,997	-2,300	-1,196	-2,348	0,312	-1,341	-0,217	-0,336	-0,666	1,855	-0,761	-0,124	-2,275	-0,479	-0,118	0,106	
6	-1,688	0,454	1,381	1,219	-0,853	-1,013	-1,337	1,750	-0,049	0,448	2,142	-0,416	0,212	0,353	0,267	-0,237	0,201	-0,098	-1,107	0,041	-0,465	-0,288	0,713	-0,519	
7	-2,222	2,394	-0,908	1,925	0,938	-0,758	1,732	-2,004	-0,493	-0,549	0,458	2,791	-0,127	-0,806	2,874	-0,423	-2,233	0,141	1,049	0,577	1,416	0,202	0,092	-0,185	
8	-0,417	-1,125	0,715	-0,590	-0,706	1,261	1,277	1,368	1,132	-1,631	-1,168	-0,144	-1,070	-0,434	-0,061	1,001	-1,038	-0,221	0,887	-0,323	1,347	2,007	-0,039	0,828	
9	0,488	-0,970	0,356	-0,545	1,534	-0,162	0,156	2,644	-0,408	-1,103	0,263	1,872	-0,786	-0,491	-0,805	-0,216	-0,164	-0,763	1,384	0,453	1,234	0,353	-0,047	0,065	
10	-2,257	-1,673	0,038	0,691	1,116	1,409	-0,657	-0,214	0,119	-1,760	1,326	1,708	1,731	-0,078	-0,697	1,208	0,807	-0,370	0,067	0,548	0,396	0,325	-1,146	0,089	
11	-1,379	1,742	1,304	-0,634	0,272	-0,859	-0,419	0,387	-0,328	-0,505	0,995	0,767	0,754	0,859	-0,014	-0,526	0,089	-1,416	0,118	0,185	-0,687	-0,319	3,370	-0,024	
12	0,150	0,915	-0,328	1,214	-1,285	0,811	1,448	-0,798	-2,220	1,228	1,199	-1,736	0,226	-0,833	1,157	1,048	-0,306	-1,380	1,089	-1,930	-2,015	-0,046	0,260	-0,691	
13	0,013	1,529	0,097	-0,932	0,601	-0,527	-1,563	1,732	0,304	-1,880	0,722	1,137	1,895	-0,958	-1,438	0,916	-0,768	0,420	0,825	-0,677	1,859	0,541	-1,316	0,231	
14	0,898	0,089	-0,071	1,667	1,398	1,905	-1,562	-0,864	-1,751	1,030	-0,877	1,515	0,118	-0,708	-0,310	-1,840	1,747	-1,759	0,121	-0,274	0,739	-1,334	0,742	0,565	
15	0,473	-0,996	0,689	1,615	1,107	0,481	-1,493	-0,912	0,356	-0,099	-3,193	0,149	-2,051	-0,658	0,293	0,080	1,289	-1,184	-1,099	-0,657	-0,263	-0,956	0,100	0,091	
16	-0,321	-2,639	-1,260	0,199	-0,716	0,067	-1,265	-0,335	1,057	0,899	-0,644	0,080	-0,308	-0,387	1,301	0,995	1,930	-1,306	0,884	0,246	-2,011	0,722	0,335	0,017	
17	0,695	0,308	-1,457	1,878	-0,351	2,260	-0,021	-2,849	-3,148	-1,127	1,442	0,516	2,994	-0,545	0,178	1,566	0,611	-1,515	2,734	1,133	2,151	1,264	-0,922	-0,472	
18	0,966	-1,820	-1,015	1,458	1,196	-3,211	0,016	0,710	-1,086	2,125	0,995	0,573	-2,249	0,866	1,069	-0,047	0,305	-1,514	-0,462	-1,197	2,637	1,177	-1,038	-0,131	
19	0,064	0,309	-2,245	-2,836	-1,871	-0,870	0,856	-0,299	-0,485	0,619	3,351	-0,294	0,295	0,670	-0,405	0,178	-0,687	1,676	1,522	0,396	-0,683	-0,538	0,052	0,191	
20	0,618	-0,704	0,915	-0,328	-0,656	1,788	-1,730	2,130	0,931	-0,022	0,217	-0,676	-0,910	0,171	-0,790	-2,583	0,456	-0,297	-0,729	-0,673	2,761	-0,042	0,314	0,912	
21	0,838	-1,040	1,299	1,153	-0,295	-1,589	-0,701	1,894	-1,990	1,563	1,476	-1,336	-0,023	0,941	-2,106	-0,213	1,388	-1,380	-1,247	-0,637	2,378	0,724	-0,054	-0,243	
22	0,402	0,196	-1,056	-0,685	0,489	-0,005	-0,275	-0,502	0,021	-0,593	2,563	0,462	0,532	0,985	0,885	-0,830	-1,088	-0,375	-0,062	-1,143	1,331	0,224	-0,765	-0,288	
23	-2,230	-1,505	-0,029	-0,144	1,227	1,203	-1,475	1,148	0,185	0,519	1,097	-1,885	-0,239	0,684	-1,830	0,085	1,027	0,785	-0,350	-0,585	-0,401	0,444	0,291	0,032	

Camada de saída		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
w2	b2																								
1	1,303	-0,339	3,168	-3,028	-1,385	0,205	1,756	-0,298	0,397	-1,753	-0,453	0,084	0,136	1,218	-0,019	0,052	-0,130	1,566	1,404	0,383	0,815	-1,807	-2,424	0,017	
2	3,468	-0,559	-0,863	-1,420	-2,567	2,882	-0,464	-1,903	-1,581	0,050	0,130	2,314	-1,682	-0,013	0,803	-2,700	0,074	2,109	4,660	1,188	0,190	-2,901	-2,058	-2,714	
3	1,541	2,397	-1,095	-0,412	0,188	-2,016	2,755	-1,053	0,537	2,494	1,857	-2,457	1,184	-2,850	-1,245	0,791	-2,141	0,747	-1,777	-1,056	1,390	-0,552	0,348	-0,988	
4	-1,505	0,822	-1,217	1,155	0,483	1,760	-2,164	0,033	-1,256	1,667	0,407	0,983	-0,980	0,112	0,372	0,381	-0,129	-0,661	1,485	-0,210	-1,644	0,660	1,114	-0,074	
5	1,596	1,916	-0,297	-0,331	0,899	-2,235	0,675	-0,502	-2,012	2,595	3,355	-1,342	2,703	-1,936	-3,315	1,918	-2,639	0,484	-1,078	-1,567	3,473	-1,277	-1,359	-0,514	



## **APENDICE I**

### Neural Modelling as a Tool to Support Blast Furnace Ironmaking

Contribuição #134 submetida e aceita junto ao  
Advanced Control of Chemical Process (ADCHEM) 2006  
em Gramado, Brasil, 2-3 de abril 2006

Evento coordenado por IFAC – International Federation of Automatic Control

## NEURAL MODELING AS A TOOL TO SUPPORT BLAST FURNACE IRONMAKING

**F. T. P de Medeiros<sup>a</sup>, A. Pitasse da Cunha<sup>b</sup>, A. M. Frattini Fileti<sup>c</sup>**

<sup>a</sup>*Companhia Siderúrgica Nacional (CSN), Brazil*

<sup>b</sup>*MetalFlexi, Brazil*

<sup>c</sup>*Chemical Systems Engineering Department, (UNICAMP), SP, Brazil*

Abstract: This paper describes the development of a hybrid model based on artificial neural network and its industrial application to the ironmaking at *Companhia Siderúrgica Nacional* (CSN-Volta Redonda/Brazil). The Iron Blast Furnace is highly complex process subject to oscillations in raw material characteristics. A precise model is essential to adjust © 2002 charging and blow conditions to match productivity, chemical quality and costs targets. A neural model was developed in order to estimate chemical and thermal parameters to feed a first principles model capable of evaluating alternative operation standards. As a consequence, operation efficiency is enhanced leading to higher productivity and lower costs. *Copyright © 2002 IFAC*

Keywords: modelling, neural network, ironmaking, iron blast furnace.

## 1. INTRODUCTION

The impulse from the domestic market and the abundance of quality raw materials have favored the development of the Brazilian steel industry, which is viewed as playing a fundamental part in the process of industrialization and development. CSN is one of the largest steelmaking groups in Latin America, with a production capacity of 5.8 million tons of raw steel per year.

The Iron Blast Furnace reduces iron ore, producing liquid iron (hot metal) which is converted to steel by exothermic oxidation of metalloids dissolved in the iron in the basic oxygen steelmaking process.

The Blast Furnace is a very complex processes in terms of chemistry, fluidodynamics and heat exchange. The composition of the burden material to be loaded and the blast to be blown determines productivity, quality and costs. Designing burden and blast requires a fairly accurate process model to define an appropriate operation standard from an almost infinite set. Particular characteristics, associated to both materials and equipment, are to be considered in the model requiring actual data to be analysed before applying first principle models.

Many simple models exist to analyse the blast furnace process based on heat, mass and chemistry balance and some are even ingenuous. However, chemical equilibrium mismatches and kinetics parameters need to be estimated based on materials and equipment characteristics in order to quantify performance indexes. Usually, to close that gap it is necessary to apply a comprehensive statistic model. Chemical composition analysis of every furnace stream need to be taken (raw material, blow, overhead gas and liquid metal), which introduces a dead time to the performance calculations.

One of the alternative and efficient tools, which enable one to obtain a numerical description of this kind of complex process, is the artificial neural network (ANN). Interactions and the dynamics among variables are readily captured from operating data base presentation to the network. From past operating conditions and calculated mismatch parameters, a network model allows performance indexes computation.

Neural networks are becoming an effective component of the steel manufacture automation system. There are various applications of neural networks in the steel industry. Schlang et al. (1997) describes the use of neural networks in the control of flat steel rolling mills and electric furnaces (Siemens AG). Cox et al. (2002) explore the use of neural networks to predict oxygen and coolant requirements

during the end-blow period of the Port Talbot basic oxygen steelmaking - BOS - plant (Corus Group). However, the authors report that the application of the neural model 'in situ' was to be carried out just in future work. Ping et al. (2003) describe the implementation of an intelligent model for controlling BOS end-points at WISCO's No 2 steel shop. This static control model combines neural networks and first principles. Indeed for the iron Blast Furnace process there are few papers on neural networks. Radhakrishnan and Mohamed (2000) describe a successful application of a neural network for the identification and control of blast furnace hot metal quality.

A growing literature within the field of chemical processes describing the use of artificial neural networks has evolved for a diverse range of engineering applications such as fault detection, signal processing, process modelling and control (Himmelblau, 2000). According to the author, because neural networks are nets of basis functions, they can provide good empirical models of complex nonlinear process useful for a wide variety of purposes.

Considering the difficulties outlined above, obtaining accurate mismatch parameters for first principles models in iron and steelmaking has proved to be a very hard task. Usually two kinds of models are employed to blast furnace operation: those very simple using estimated mismatch parameters that are corrected as operation goes on and complex models with to many parameters to be of practical use.

The present work is concerned with developing a hybrid model - neural network and mass and heat balances - and its application to the ironmaking blast furnace at CSN (Brazil). The main goal is to obtain a tool to design burden and blast conditions in order to match the targets of productivity, chemical quality and costs of the liquid metal.

## 1. METHODS

### 2.1 Process Description

*Companhia Siderúrgica Nacional's* steelworks entails three blast furnaces, two of them in operation and one out of service. Blast Furnace 2 produces nearly 4,000 tons of hot metal per day whereas Blast Furnace 3 produces between 9,500 and 11,000 tons per day. Iron ore sinter, pellets and lumpy hematite constitute the ferrous burden. As reducing agents, metallurgical coke and pulverised coal are used, being the latter injected through the tuyeres. Sometimes dolomite or quartzite are used as fluxes to control slag composition. The blast composition (air, oxygen and steam) and the rate of coal injection are the main and most sensitive parameters of control.

Operation aims at production rate, hot metal chemical composition and temperature and ultimately, low cost. Because the plant is self-sufficient in coke, a small proportion of it is imported bringing significant characteristic variations to the mixture.

The core of the process is a counter-flow reactor where a series of chemical and thermal exchanges are performed in several internal zones (Figure 1).

As the ferrous burden descends it is first dried, then reduced by the up-coming process gas containing CO and H<sub>2</sub>. This zone, called upper granulated zone, or preparation zone or even indirect reduction zone, ideally produces wustite (Fe<sub>y</sub>O) to be reduced to metallic Fe in the inferior zones. The index y, in this case, approaches 0.95. In real terms, however, the wustite will have an excess of oxygen which is quantified in terms of kg-mol of O / kg-mol of Fe. This parameter is necessary to establish a proper mass and thermal balance of the process and will be designated by ω (Rist and Maysson, 1967). The thermal balance also needs a parameter to represent real conditions. This is the heatloss that will be represented by λ.

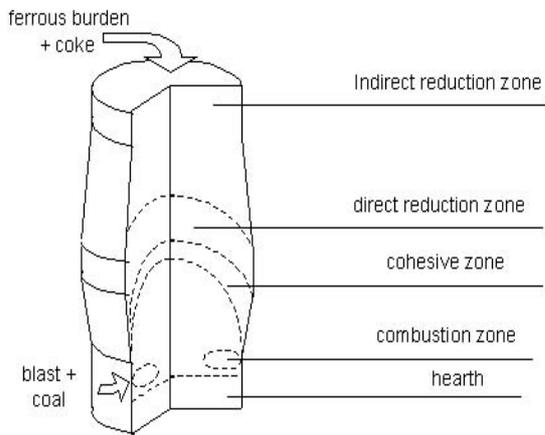
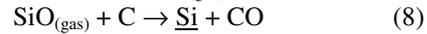
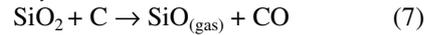
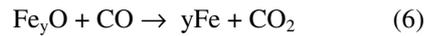
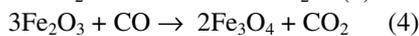
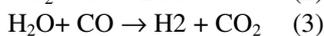
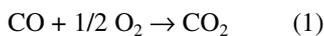


Fig. 1. The Iron Blast Furnace internal zones.

The main heat source is the combustion of coal and coke that produces mixture of CO, CO<sub>2</sub>, H<sub>2</sub>, H<sub>2</sub>O and N<sub>2</sub>. CO is regenerated in the direct reduction zone and below by the Boudouard reaction (Eq. 2). H<sub>2</sub> also plays an important role and the C-H-O will be in equilibrium in most sub-processes.

The basic chemical reactions involved are:



Silicon is partially reduced from silica into gaseous silica monoxide and incorporated to the liquid by further reduction. This process is rather complex and the metal silicon content is very hard to estimate.

The other impurities in the metal, manganese and phosphorous, do not represent a difficult estimation task depending more on the raw materials composition than on the process conditions.

## 2.2 Artificial Neural Network (ANN)

**Theory.** ANN are mathematical models constituted by several neurons, arranged in different layers (input, hidden and output), interconnected though a complex network. The multi-layer feedforward, that is the most popular of the much architectures currently available, was used. According to Equation (9), a neuron is responsible for the summation of all signals from previous layer's neurons,  $y_{i,j}$  (amplified or weakened by weight values,  $w_{i,j,k}$ ) and a value called bias,  $b_{i,j}$ .  $i$  represents the order of the layer whereas  $j$  and  $k$  indicate the order of the neuron in the layer. An activation function,  $f$  - such as hyperbolic tangent, sigmoid or linear function - is used for the activation of the neuron output,  $y_{i,k}$ .

$$y_{i,k} = f(\sum (w_{i,j,k} y_{i-1,j}) + b_{i,k}) \quad (9)$$

The data processing within the ANN structure is executed collectively and simultaneously through the dense network of neurons and their connections. Those characteristics were crucial for the this technique to be chose and not other multivariate regression ones which tend to give too much weigh to extreme values of the input variables.

**Training the ANN.** Once the network weights and biases have been initialized, the network is ready for training. The training process requires a set of examples of proper process behavior -network inputs and target outputs. During training the weights and biases of the network are iteratively adjusted to minimize the network objective function. The basic training algorithm is the backpropagation algorithm, in which the weights are moved in the direction of the negative gradient (Demuth and Baele, 2002).

The first method for improving generalization is called regularization. This involves modifying the objective function, which is normally chosen to be the sum of squares of the network errors on the training set. It is possible to improve generalization if we modify the objective function by adding a term

that consists of the mean of the sum of squares of the network weights and biases:

$$F = \beta \cdot SSE + \alpha SSW \quad (10)$$

where  $SSE$  is the sum of squared errors,  $SSW$  is the sum of squares of the network weights, and  $\beta$  and  $\alpha$  are objective function parameters (Demuth and Baele, 2002).

According to Foresse and Hagan (1997), using this objective function will cause the network to have smaller weights and biases, and this will force the network response to be smoother and less likely to overfit. One feature of this algorithm is that it provides a measure of how many network parameters, (weights and biases) are being effectively used by the network. This effective number of parameters will be called  $p$ .  $P$  is the total number of parameters in the network.

Neural network training can be made more efficiently if certain preprocessing steps are performed on the network inputs and targets. Then, before training the network the training data was normalized in the range [0.1, 0.9], as follows:

$$y_{0,j} = 0.8 ((x_j - x_{min_j}) / (x_{max_j} - x_{min_j})) + 0.1 \quad (11)$$

where  $y_{0,i}$  is the normalized value for the variable  $x_j$ , and  $x_{min_j}$  and  $x_{max_j}$  are the minimum and maximum of each variable  $x_j$ .

Modeling and data set. A neural model was developed in the present work to predict: the equilibrium mismatch parameter for the oxygen mass balance ( $\omega$ ), the thermal loss parameter for the heat balance ( $\lambda$ ), the gas flow resistance parameter ( $\rho$ ), the hot metal Silicon content ( $[Si]$ ) and the sulfur partition coefficient between slag and metal ( $K_s$ ). Feeding those parameters to a simple mass and heat balance, a precise operation pattern is defined to guide operators and technical staff for immediate and strategic decision making.

Table 1 shows the final variables selection and their meaning. Coke drum ( $x15$ ) and reactivity ( $x16$ ) indexes quantify physical strength and chemical activity, respectively, and are important both to gas flow and chemistry in the process.

Three years of Blast Furnace 3 operation were observed. Records were condensed into 23 input variables. Sets corresponding to days with missing or inconsistent data were filtered out. Records include those acquired by the furnace digital automation system, works and mines laboratories. Finally a 28 x 820 data bank was

Apêndice I

gathered, randomized and fed into a MATLAB® program. The final data-base was then split into two sets, one for training and the other for generalization tests (15% of the data). It was carefully checked the range of each variable since it should be similar to both sets.

In the search for the architecture that could yield the best possible prediction model accuracy, a study was performed to establish the number of nodes in the network hidden layer.

Table 1. Input (x) and output (y) variables used for the neural modeling.

Blast variables	
$x1$	kg-mol of N2 in blast / ton of metal
$x2$	kg-mol de H2O in blast / ton of metal
Burden variables	
$x3$	kg of slag / ton of metal
$x4$	Primary slag B4
$x5$	Hearth slag B4
$x6$	blast temperature (°C)
$x7$	kg of small-coke / ton of metal
$x8$	kg of injected coal / ton of metal
$x9$	kg of lumpy hematite / ton of metal
$x10$	kg of pellets / ton of metal
$x11$	kg of quartzite / ton of metal
$x12$	external coke / total coke
$x13$	pulverized coal ash content
$x14$	pulverized coal oxygen content
$x15$	average coke Drum Index
$x16$	average coke Reactivity Index
$x17$	coke mean size (mm)
$x18$	hematite < 6,35 mm fraction
$x19$	hematite > 38 mm fraction
$x20$	hematite Decrepitation Index
$x21$	kg of stock sinter / ton of metal
Equipment and environmental variables	
$x22$	rain fall index (mm)
$x23$	tapping hole campaign index (1 or 0)
Output variables	
$y1$	wustite stoichiometric index ( $\omega$ )
$y2$	gas flow resistance ( $\rho$ )
$y3$	metal silicon content ( $[Si]$ )
$y4$	heat losses in MJ / ton of metal ( $\lambda$ )
$y5$	sulfur in slag / sulfur in metal ( $K_s$ )

The predicted parameters are combined with other variables in a deterministic model to estimate the overall process pattern. The parameter  $\theta$  represents the ratio between metal and gas produced.  $\phi$  represents the unity gas flow calculated from the predicted gas resistance,  $\rho$ , and the pressure imposed by the equipment, blower and reactor. The overall performance index,  $\pi$ , is the final product of the model, meaning the amount of metal produced in a unity time for each cross section area unit and results form the product  $\phi \times \theta$ . Figure 2 shows a

cause and effect diagram for the hybrid model. The four final variables:  $[Si]$ ,  $[S]$ ,  $\mu$  and  $\pi$  are efficient process performance indexes. The first two indicate metal silicon and sulfur content, respectively. The parameter  $\mu$ , as defined by Rist (Rist and Misson, 1967) quantifies the specific consumption of reducing agents ( $C + H_2$ ) and ultimate the metal cost. The index  $\pi$  quantifies the amount of metal per unity area, therefore, the overall process productivity.

### 2.3 Experimental Industrial Application

The operation process using the model as supporting toll at Blast Furnace 3 is shown on Figure 3. The application will be extended to Blast Furnace 2 after the experimental application to Blast Furnace 3.

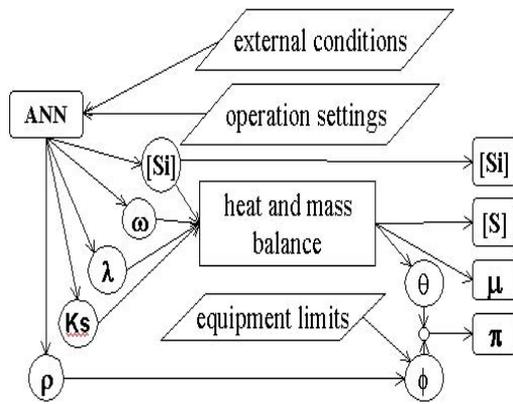


Fig 2. Data flow diagram for the hybrid model.

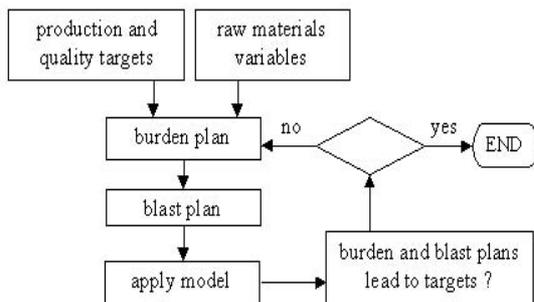


Fig 3. Industrial application procedure flow diagram.

### 3. RESULTS

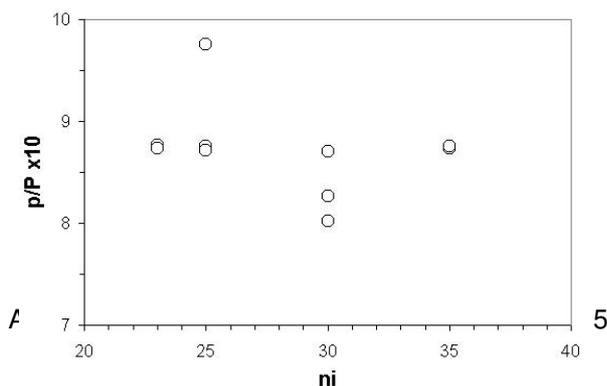


Fig 4. Relationship between effective parameter ratio ( $p/P$ ) and the number of neurons in the hidden layer ( $n_i$ ).

Figure 4 shows how the ratio between initial number ( $P$ ) of network parameters - weights and bias - and the number of effective parameters after training ( $p$ ) behaves with the increase in the number of neurons in the hidden layer ( $n_i$ ).

According to Foresse and Hagan (1997) the decreasing effective parameters ratio ( $p/P$ ) indicate that the number of neurons is excessive. Another criterion leads to the same conclusion, as illustrated by Figure 5. It is clear that the larger number of hidden layer neurons does not contribute to a smaller mean quadratic error for the generalization set although the mean error for the training set decreased. In conclusion, the best network architecture was found to be 23x23x5.

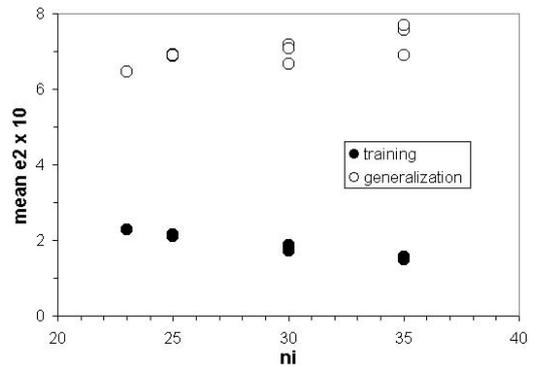


Fig 5. Relationship between mean quadratic errors ( $mean e^2$ ) and the number of neurons in the hidden layer ( $n_i$ ).

In this study and for the chosen architecture, the neuron activation function used in the hidden layer was a sigmoid one while a linear function was chosen for the output layer neurons.

Table 2 shows the mean square error for each of the 5 output variables expressed in terms of respective standard deviations. As expected, smaller mean quadratic errors are obtained for training sets. Mean errors for generalization sets were considered acceptable.

Table 2 – Square mean errors for the output variables in terms of respective standard deviations

	training	generalization
$\omega$	0.503	0.722
$\rho$	0.400	0.651
$[Si]$	0.560	0.822
$\lambda$	0.354	0.496
$Ks$	0.531	0.729

Figure 6 shows how estimated standardized values (horizontal axis) match actual ones. It could be also noted from Figure 6 the tendency of experimental seen and unseen points to follow the diagonal line, indicating the accuracy of the network approach. The estimation of low values of  $\omega$  and high vales of  $K_s$  was deficient for a few cases.

It should be pointed out that  $\phi$ ,  $\theta$  and, consequently,  $\pi$  will depend not only on the values estimated by the network but also on other process variables. Therefore there is no point in estimating them at this moment.

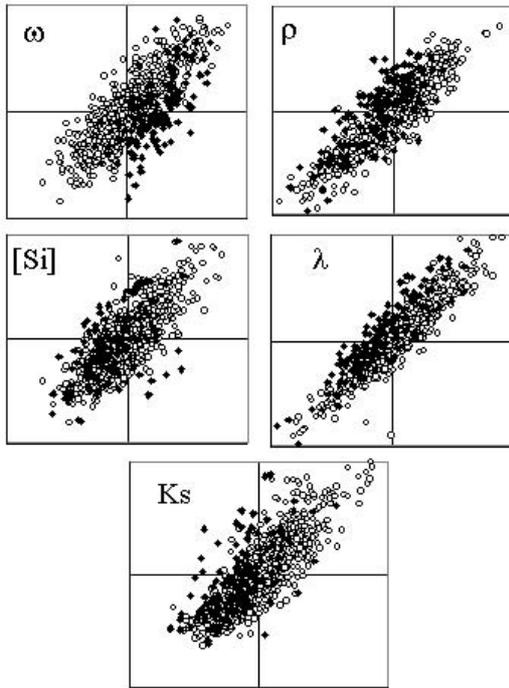


Fig 6. Dispersion plots of the network output variables (predicted values x actual values) for both training (O) and generalization (♦) sets. Axes cross at the mean value.

### 3.2 Experimental Industrial Application

Following the steps previously described (Figure 3), the experimental industrial application was carried out during a twenty-day period. During the first twenty days of September 2005, the Blast Furnace number 3 operation was guided by the model. According to Figure 2, four variables were taken to access the prediction capacity of the model: coke-rate ( $CR$ ), metal silicon content  $[Si]$ , sulfur metal content  $[S]$  and Ergun fluiddynamic resistance index ( $K$ ). Figures 7 shows the results of the industrial observations while in Table 3 results can be numerically compared.

It can be observed from table 3 that the error standard deviation was smaller than the relative

standard for the experimental set and much smaller than the standard deviation observed in actual operation. For the fluidodynamic resistance it can be pointed out that the test period did not present sufficient variation for the adequate assessment of the model on this aspect.

Table 3 – Statistical analysis of indexes observed during the test period without and with model

	error mean	error sd	set sd	population sd
CR (kg/t)	2.8	2.7	4,6	36.9
$[Si] \times 10^4$	0.8	2.9	4,4	13.5
$[S] \times 10^5$	1.5	3.8	4,4	7.9
K	19.0	11.42	8,9	24.0

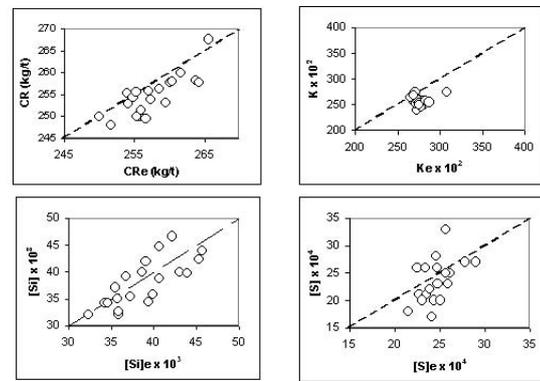


Fig 7 – Results of experimental industrial tests.

Because operational corrective actions were still too timid, fuel-rate corrections were allowed some hot metal temperature variations which contaminated sulfur control. This can be observed in Figure 8. In future, better heat control, with more confident use of the model, will also improve chemical quality, because chemical equilibrium is strongly connected to temperature.

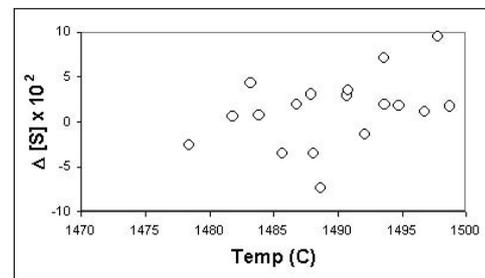


Figure 8 – Influence of hot metal temperature on the prediction error for hot metal sulfur

## 4. CONCLUSIONS

The main contribution of the present work is the development of a neural model which can increase

prediction accuracy and operation performance while reducing costs for the blast furnace process at *Companhia Siderúrgica Nacional* (CSN-Volta Redonda/ RJ/ Brazil). Obtaining liquid iron in stable conditions is a very hard task, because the Blast Furnace is a complex process, conjugating several sub-processes. Some of them are continuous, some transient, occurring in the same reactor and still subject to oscillations in raw material composition.

The developed hybrid model, based on mass and heat balance associated to an artificial neural network, aims at supporting both operational and strategic decision making.

A 23x23x5 feedforward network proved to be an efficient architecture, using sigmoid and linear activation functions for the hidden and output neurons, respectively.

Except for fluidodynamic resistance, in other words, permeability, the period in which the model was used to guide industrial furnace operation proved to be predictable and consistent. For assessment of the permeability prediction a longer period will be necessary to allow for significant variation of that parameter.

The analysis of alternative raw materials or practice standards can be held also with the support of the model as long as the variables are kept inside the operating range studied.

It could be concluded that the neural model is a relevant tool to support an iron Blast Furnace operation since some corrections and retraining are

carefully carried out by expert human operators in a systematic way. These procedures are crucial for adopting the neural model as a standard operating practice.

## REFERENCES

- Demuth, H., Beale, M., 2002. *Neural Network Toolbox User's Guide for Use with MATLAB®*. The Mathworks Inc., Version 4, Reading:
- Hagan, M., Chapter 5: *Backpropagation*.
- Foresee, F.D., Hagan, M.T., 1997. Gauss-Newton approximation to Bayesian Learning. Proc. IJCNN, 1930-1935.
- Himmelblau, D.M., 2000. Applications of artificial networks in chemical engineering. *Korean J. Chem. Eng.* **17** (4), 373-392.
- Ping, H., Liu, L., Lihong, Y., Zhigang, H., Rong, D., Jingbo, X., Wei, C., Yisheng, T., Chenghuan, Y., Fengxi, L., 2003. Combining intelligent and mathematical models for BOS control at WISCO. *Steel Times International* **27** (8), 31-32.
- Radhakrishnan, V.R. , Mohamed A.R.,2000. Neural networks for the identification and control of blast furnace hot metal quality. *Journal of Process Control* **10**, 509-524.
- Rist, A.; Meysson, N. 1967. *Journal of Metals*, April, **50**.

## **APENDICE II**

### Reviving Traditional Blast Furnace Models with New Mathematical Approach

Artigo submetido (ref IRS2047) e aceita com pendências de revisão junto ao periódico *Ironmaking and Steelmaking*, publicado por *Maney Publising* para *The Institute of Materials, Minerals and Mining*, Leeds, Inglaterra.

## REVIVING TRADITIONAL BLAST FURNACE MODELS WITH NEW MATHEMATICAL APPROACH

**F. T. P de Medeiros<sup>a</sup>, S. J. X. Noblat<sup>a</sup>, A. Pitasse da Cunha<sup>b</sup>, A. M. Frattini Fileti<sup>c</sup>**

<sup>a</sup>*Companhia Siderúrgica Nacional (CSN), Brazil*

<sup>b</sup>*MetalFlexi, Brazil*

<sup>c</sup>*Chemical Systems Engineering Department, (UNICAMP), SP, Brazil*

**Abstract:** This paper describes how traditional analytical blast furnace models can be revived by the inclusion of new mathematical tools. Combining some fundamental models with new mathematical algorithms can create efficient and simple to use hybrid models. A hybrid model based on artificial neural network and its industrial application to the new Blasrt Furnace number 3 at *Companhia Siderúrgica Nacional (CSN-Volta Redonda/Brazil)* was developed, test and put in use. In Blast Furnace operation, a multi variable complex process, subject to oscillations in raw material characteristics, a precise model is essential to adjust charging and blow conditions to match productivity, chemical quality and costs targets. A neural model was developed in order to estimate chemical and thermal parameters to feed a first principles model capable of evaluating alternative operation standards. As a consequence, operation efficiency is being enhanced leading to higher productivity and lower costs.

**Keywords:** modelling, neural network, ironmaking, blast furnace.

## 1. INTRODUCTION

The impulse from the domestic market and the abundance of quality raw materials have favored the development of the Brazilian steel industry, which is viewed as playing a fundamental part in the process of industrialization and development. CSN is one of the largest steelmaking groups in Latin America, with a production capacity of 5.8 million tons of crude steel per year.

Most steel produced in Brazil follows the Blast Furnace-BOF route .

The Blast Furnace is a very complex processes in terms of chemistry, fluidodynamics and heat exchange. The composition of the burden material to be loaded and the blast to be blown determines productivity, quality and costs. Designing burden and blast requires a fairly accurate process model to define an appropriate operation standard from an almost infinite set. Particular characteristics, associated to both materials and equipment, are to be considered in the model requiring actual data to be analysed before applying first principle models.

Many simple models exist to analyse the blast furnace process based on heat, mass and chemistry balance and some are even ingenuous. However, chemical equilibrium mismatches and kinetics parameters need to be estimated based on materials and equipment characteristics in order to quantify performance indexes. Usually, to close that gap, it is necessary to adopt values obtained when the models are applied as analytical tolls on a set of past operation data. As raw materials and other conditions change those values are no longer valid and need to be updated. The number of variables to be considered on this updating is considerably large.

One of the alternative and efficient tools, which enable one to obtain a numerical description of this kind of complex process, is the artificial neural network (ANN). Interactions and the dynamics among variables are readily captured from operating data base presentation to the network. From past operating conditions and calculated mismatch parameters, a network model allows performance indexes computation without time delay of chemical analysis.

Neural networks are becoming an effective component of the steel manufacture automation system. There are various applications of neural

networks in the steel industry. Schlang et al. (1997) describes the use of neural networks in the control of flat steel rolling mills and electric furnaces (Siemens AG). Cox et al. (2002) explore the use of neural networks to predict oxygen and coolant requirements during the end-blow period of the Port Talbot basic oxygen steelmaking - BOS - plant (Corus Group). However, the authors report that the application of the neural model 'in situ' was to be carried out just in future work. Ping et al. (2003) describe the implementation of an intelligent model for controlling BOS end-points at WISCO's No 2 steel shop. This static control model combines neural networks and first principles. Indeed for the iron Blast Furnace process there are few papers on neural networks. Radhakrishnan and Mohamed (2000) describe a successful application of a neural network for the identification and control of blast furnace hot metal quality.

A growing literature within the field of chemical processes describing the use of artificial neural networks has evolved for a diverse range of engineering applications such as fault detection, signal processing, process modelling and control (Himmelblau, 2000). According to the author, because neural networks are nets of basis functions, they can provide good empirical models of complex nonlinear process useful for a wide variety of purposes.

Considering the difficulties outlined above, obtaining accurate mismatch parameters for first principles models in iron and steelmaking has proved to be a very hard task. Usually two kinds of models are employed to blast furnace operation: those very simple using estimated mismatch parameters that are corrected as operation goes on and complex models with to many parameters to be of practical use.

The present work is concerned with developing a hybrid model - neural network and mass and heat balances – and its application to the ironmaking blast furnace at CSN (Brazil). The main goal is to obtain a a tool to design burden and blast conditions in order to match the targets of productivity, chemical quality and costs of the liquid metal.

## 2. DEVELOPMENT OF THE MODEL

*The blast furnace process*

*Companhia Siderúrgica Nacional's* steelworks entails three blast furnaces, two of them in operation and one out of service. Blast Furnace 2 produces nearly 4,000 tons of hot metal per day whereas Blast Furnace 3 produces between 9,500 and 11,000 tons per day. Iron ore sinter, pellets and lumpy hematite constitute the ferrous burden. As reducing agents, metallurgical coke and pulverised coal are used, being the latter injected through the tuyeres. Sometimes dolomite or quartzite are used as fluxes to control slag composition. The blast composition (air, oxygen and steam) and the rate of coal injection are the main and most sensitive parameters of control. Operation aims at production rate, hot metal chemical composition and temperature and ultimately, low cost. Because the plant is self-sufficient in coke, a small proportion of it is imported bringing significant characteristic variations to the mixture.

The core of the fundamental model is Rist-Michard-Meysson concept of oxygen exchange. Figure 1. (Rist and Maysson, 1967)

As the ferrous burden descends it is first dried, then reduced by the up-coming process gas containing CO and H<sub>2</sub>. This zone, called upper granulated zone, or preparation zone or even indirect reduction zone, ideally produces wustite (Fe<sub>y</sub>O) to be reduced to metallic Fe in the inferior zones. The index y, in this case, approaches 0.95. In real terms, however, the wustite will have an excess of oxygen which is quantified in terms of kg-mol of O / kg-mol of Fe. This parameter is necessary to establish a proper mass and thermal balance of the process and will be designated by ω. The thermal balance also needs a parameter to represent real conditions. This is the heat loss which will be represented by λ.

The main heat source is the combustion of coal and coke that produces mixture of CO, CO<sub>2</sub>, H<sub>2</sub>, H<sub>2</sub>O and N<sub>2</sub>. CO is regenerated in the direct reduction zone and below by the Boudouard reaction (Eq. 2). H<sub>2</sub> also plays an important role and the C-H-O will be in equilibrium in most sub-processes.

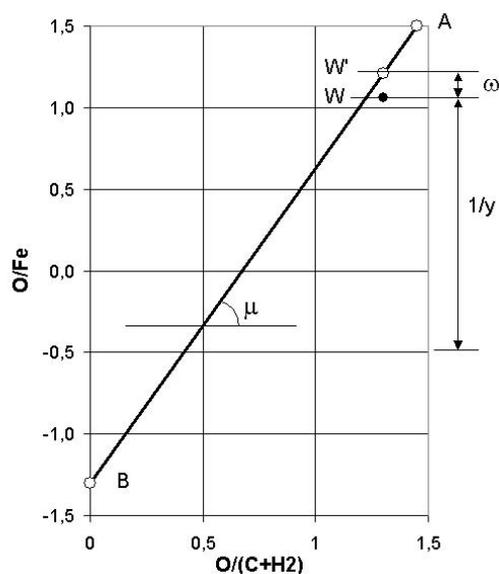
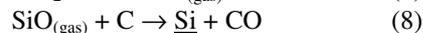
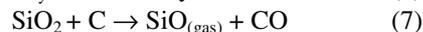
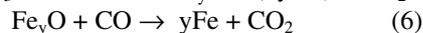
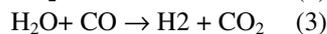
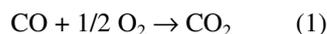


Figure. 1 – Rist-Maysson-Michard diagram.

The basic chemical reactions involved are:



Silicon is partially reduced from silica into gaseous silica monoxide and incorporated to the liquid by further reduction. This process is rather complex and the metal silicon content is very hard to estimate. No equilibrium is considered here but the final value of the hot metal silicon content, [Si].

Sulphur content on the metal depends strongly on the process, but one parameter permits its estimation. If K<sub>s</sub>, the sulphur partition coefficient can be statistically determined, sulphur load and slag-rate will complete the data needed.

The other impurities in the metal, manganese and phosphorous, do not represent a difficult estimation task depending more on the raw materials composition than on the process conditions.

### **Artificial Neural Network (ANN)**

ANN are mathematical models constituted by several neurons, arranged in different layers (input, hidden and output), interconnected through a complex network. The multi-layer feedforward, that is the most popular of the much architectures currently available, was used. According to Equation (9), a neuron is responsible for the summation of all signals from previous layer's neurons,  $y_{i,j}$  (amplified or weakened by weight values,  $w_{i,j,k}$ ) and a value called bias,  $b_{i,k}$ .  $i$  represents the order of the layer whereas  $j$  and  $k$  indicate the order of the neuron in the layer. An activation function,  $f$  - such as hyperbolic tangent, sigmoid or linear function - is used for the activation of the neuron output,  $y_{i,k}$ .

$$y_{i,k} = f(\sum (w_{i,j,k} y_{i-1,j}) + b_{i,k}) \quad (9)$$

The data processing within the ANN structure is executed collectively and simultaneously through the dense network of neurons and their connections.

Once the network weights and biases have been initialized, the network is ready for training. The training process requires a set of examples of proper process behavior -network inputs and target outputs. During training the weights and biases of the network are iteratively adjusted to minimize the network objective function. The basic training algorithm is the backpropagation algorithm, in which the weights are moved in the direction of the negative gradient (Demuth and Baele, 2002).

The first method for improving generalization is called regularization. This involves modifying the objective function, which is normally chosen to be the sum of squares of the network errors on the training set. It is possible to improve generalization if we modify the objective function by adding a term that consists of the mean of the sum of squares of the network weights and biases:

$$F = \beta \cdot SSE + \alpha SSW \quad (10)$$

where  $SSE$  is the sum of squared errors,  $SSW$  is the sum of squares of the network weights, and  $\beta$  and  $\alpha$  are objective function parameters (Demuth and Baele, 2002).

According to Foresee and Hagan (1997), using this objective function will cause the network to have smaller weights and biases, and this will force the network response to be smoother and less likely to overfit. One feature of this algorithm

is that it provides a measure of how many network parameters (weights and biases) are being effectively used by the network. This effective number of parameters will be called  $p$ .  $P$  is the total number of parameters in the network.

Neural network training can be made more efficiently if certain preprocessing steps are performed on the network inputs and targets. Then, before training the network the training data was normalized in the range [0.1, 0.9], as follows:

$$y_{0,j} = 0.8 ((x_j - x_{min_j}) / (x_{max_j} - x_{min_j})) + 0.1 \quad (11)$$

where  $y_{0,i}$  is the normalized value for the variable  $x_j$ , and  $x_{min_j}$  and  $x_{max_j}$  are the minimum and maximum of each variable  $x_j$ .

A neural model was developed in the present work to predict: the equilibrium mismatch parameter for the oxygen mass balance ( $\omega$ ), the thermal loss parameter for the heat balance ( $\lambda$ ), the gas flow resistance parameter ( $\rho$ ), the hot metal Silicon content ( $[Si]$ ) and the sulfur partition coefficient between slag and metal ( $K_s$ ). Feeding those parameters to a simple mass and heat balance, a precise operation pattern is defined to guide operators and technical staff for immediate and strategic decision making.

Table 1 shows the final variables selection and their meaning. Coke drum ( $x_{15}$ ) and reactivity ( $x_{16}$ ) indexes quantify physical strength and chemical activity, respectively, and are important both to gas flow and chemistry in the process.

Three years of Blast Furnace 3 operation were observed. Records were condensed into 23 input variables. Sets corresponding to days with missing or inconsistent data were filtered out. Records include those acquired by the furnace digital automation system, works and mines laboratories. Finally a 28 x 820 data bank was gathered, randomized and fed into a MATLAB® program. The final data-base was then split into two sets, one for training and the other for generalization tests (15% of the data). It was carefully checked the range of each variable since it should be similar to both sets.

In the search for the architecture that could yield the best possible prediction model accuracy, a

study was performed to establish the number of nodes in the network hidden layer.

For the chosen architecture, the neuron activation function used in the hidden layer was a sigmoid one while a linear function was chosen for the output layer neurons.

**Table 1. Input (x) and output (y) variables used for the neural modeling.**

Blast variables	
X1	kg-mol of N2 in blast / ton of metal
X2	kg-mol de H2O in blast / ton of metal
Burden variables	
X3	kg of slag / ton of metal
x4	Primary slag B4
x5	Hearth slag B4
x6	blast temperature (°C)
x7	kg of small-coke / ton of metal
x8	kg of injected coal / ton of metal
x9	kg of lumpy hematite / ton of metal
x10	kg of pellets / ton of metal
x11	kg of quartzite / ton of metal
x12	external coke / total coke
x13	pulverized coal ash content
x14	pulverized coal oxygen content
x15	average coke Drum Index
x16	average coke Reactivity Index
x17	coke mean size (mm)
x18	hematite < 6,35 mm fraction
x19	hematite > 38 mm fraction
x20	hematite Decrepitation Index
x21	kg of stock sinter / ton of metal
Equipment and environmental variables	
x22	rain fall index (mm)
x23	tapping hole campaign index (1 or 0)
Output variables	
y1	wustite stoichiometric index ( $\omega$ )
y2	gas flow resistance ( $\rho$ )
y3	metal silicon content ( $[Si]$ )
y4	heat losses in MJ / ton of metal ( $\lambda$ )
y5	sulfur in slag / sulfur in metal ( $Ks$ )

The predicted parameters are combined with other variables in a deterministic model to estimate the overall process pattern. The parameter  $\theta$  represents the ratio between metal and gas produced.  $\phi$  represents the unity gas flow calculated from the predicted gas resistance,  $\rho$ , and the pressure imposed by the equipment, blower and reactor. The overall performance index,  $\pi$ , is the final product of the model, meaning the amount of metal produced in a unity

time for each cross section area unit and results form the product  $\phi \times \theta$ . Figure 2 shows a cause and effect diagram for the hybrid model. The four final variables:  $[Si]$ ,  $[S]$ ,  $\mu$  and  $\pi$  are efficient process performance indexes. The first two indicate metal silicon and sulfur content, respectively. The parameter  $\mu$ , as defined by Rist (Rist and Misson, 1967) quantifies the specific consumption of reducing agents (C + H<sub>2</sub>) and ultimate the metal cost. The index  $\pi$  quantifies the amount of metal per unity area, therefore, the overall process productivity.

### 3. USING THE MODEL

The operation process using the model as supporting toll at Blast Furnace 3 is represented on Figure 3. The application will be extended to Blast Furnace 2 after the experimental application to Blast Furnace 3.

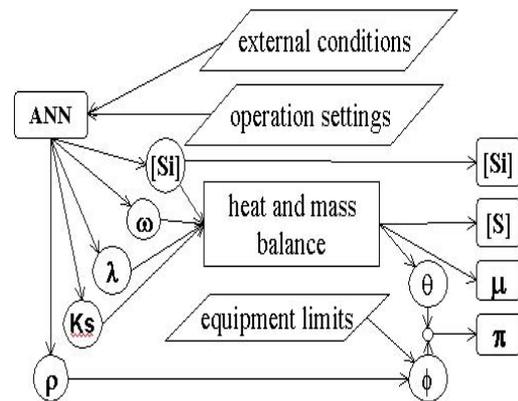


Figure 2. - Cause and effect diagram for the hybrid model.

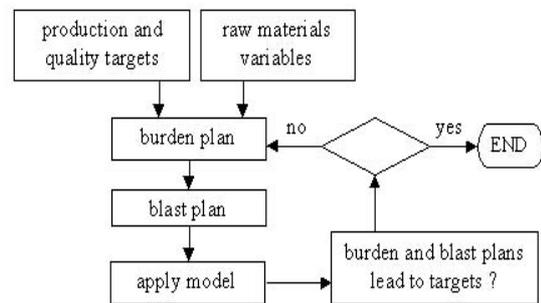


Fig 3. Industrial application procedure flow diagram.

#### 4. RESULTS

Following the steps previously described (Figure 3), the experimental industrial application was carried out during a eight-week period. During the first 4 weeks, from October, the 2<sup>nd</sup> 2005, the Blast Furnace number 3 operation data were collected and taken as reference. During the following 4 weeks, the operation was guided by the model. Six variables were chosen to assess the performance of the model as an operational guide : production rate (PR) in ton/day, coke-rate (CR) in kg/ton of hot metal, blast rate standard deviation (BRs), metal silicon content standard deviation [Si]s, sulphur metal content standard deviation [S]s and Ergun fluidodynamic resistance index (K). Figures 4 shows the results while in Table 3 results can be numerically compared.

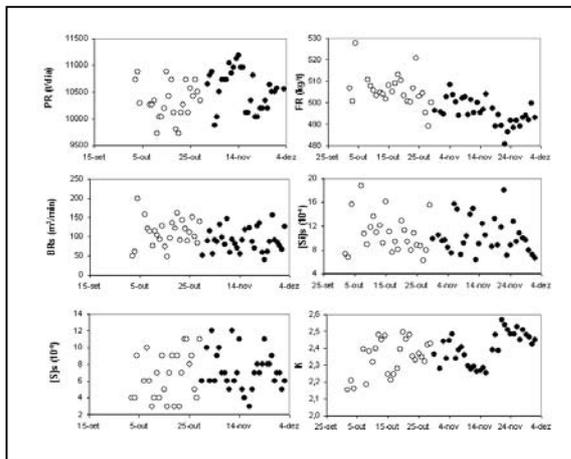


Figure 4 – Results of the industrial test.  
(o) means without model and (●)with model

Table 3 – Statistical analysis of indexes observed during the test period without and with model

	PR	CR	BRs	[Si]s $\times 10^4$	[S]s $\times 10^5$	K
with	9679	506	111	11.7	5.6	2.34
without	10554	496	99	10.3	7.3	2.40
mean sd	203	0.9				0.01
t	4.31	-11.3				4.45
F			1.49	1.07	1.24	
$P(H_0)$	$<10^{-3}$	$<10^{-3}$	0.85	0.57	0.71	$<10^{-3}$

Production rate and coke-rate are the real targets. The blast rate standard deviation shows how smoothly the furnace operated in the period whereas silicon and sulphur standard deviations

indicate chemical quality stability. K is not controlled by the model and depends on raw material quality. Here it aims at verifying external conditions for the operation.

The periods were compared calculating the probability of the null hypothesis, i.e. if the values are really significantly different. Means (PR, CR and K) were compared using the Student distribution where t is the ratio means difference and mean standard deviation. For the standard deviations, the other three indexes, the Fisher distribution was used and F represents the ration between variances.

It can be verified that productivity was higher after introducing the model as operational guide and. Fuel specific consumption was lower. There was no significant difference in chemical quality with or without the model. Raw materials conditions were different and worse during the period were the model was used.

#### 4. CONCLUSIONS

The main contribution of the present work is the development of a neural model which can increase prediction accuracy and operation performance while reducing costs for the blast furnace process at *Companhia Siderúrgica Nacional (CSN-Volta Redonda/ RJ/ Brazil)*. Obtaining liquid iron in stable conditions is a very hard task, because the Blast Furnace is a complex process, conjugating several sub-processes. Some of them are continuous, some transient, occurring in the same reactor and still subject to oscillations in raw material composition.

The developed hybrid model, based on mass and heat balance associated to an artificial neural network, aims at supporting both operational and strategic decision making.

Except for the chemical quality the period in which the model was used to guide industrial furnace operation was related to better higher productivity and less costly raw materials conditions were adverse.

The analysis of alternative raw materials or practice standards can be held also with the support of the model as long as the variables are kept inside the operating range studied.

It could be concluded that old blast furnace models can be reshaped in combination with

neural model and became efficient tools to support an iron Blast Furnace operation.

#### REFERENCES

- Demuth, H., Beale, M., 2002. *Neural Network Toolbox User's Guide for Use with MATLAB®*. The Mathworks Inc., Version 4, Reading:
- Hagan, M., Chapter 5: *Backpropagation*.
- Foresee, F.D., Hagan, M.T., 1997. Gauss-Newton approximation to Bayesian Learning. Proc. IJCNN, 1930-1935.
- Himmelblau, D.M., 2000. Applications of artificial networks in chemical engineering. *Korean J. Chem. Eng.* **17** (4), 373-392.
- Ping, H., Liu, L., Lihong, Y., Zhigang, H., Rong, D., Jingbo, X., Wei, C., Yisheng, T., Chenghuan, Y., Fengxi, L., 2003. Combining intelligent and mathematical models for BOS control at WISCO. *Steel Times International* **27** (8), 31-32.
- Radhakrishnan, V.R. , Mohamed A.R.,2000. Neural networks for the identification and control of blast furnace hot metal quality. *Journal of Process Control* **10**, 509-524.
- Rist, A.; Meysson, N. 1967. *Journal of Metals*, April, **50**..