

PREDIÇÃO DA CONCENTRAÇÃO DE FÓSFORO DO AÇO EM CONVERTEDORES A OXIGÊNIO

LUAN CARLOS DE OLIVEIRA*, CASSIUS ZANETTI RESENDE*, DANIEL CRUZ CAVALIERI*, HENRIQUE SILVA FURTADO†, LEANDRO RODRIGUES RAMOS†, ROBERTO DALMASO†

**Instituto Federal do Espírito Santo - Campus Serra
Rodovia ES 010, Km 6,5, 29173-087 - Manginhos
Serra, ES, Brasil*

†*Arcelor Mittal Tubarão
Av. Brigadeiro Eduardo Gomes, 930, Jardim Limoeiro, 29163970
Serra, ES, Brasil*

Emails: luandeol@gmail.com, cassius@ifes.edu.br, daniel.cavaliere@ifes.edu.br, henrique.furtado@arcelormittal.com.br, leandro.ramos@arcelormittal.com.br, roberto.dalmaso@arcelormittal.com.br

Abstract— The objective of this work is the development and performance analysis between two prediction models, based on the SVR (Support Vector Regression) and neural network Autoencoder techniques. The models aim at predicting the steel phosphorus content in basic Oxygen Furnaces at the end of primary refining, replacing sampling usually performed for laboratory analysis, reducing process time and consequently increasing steelmaking productivity. For the modeling, a set of variables was used, consisting of inputs and refining instrumentation data. The k-fold cross validation was used to obtain generalist models and as performance metrics the RMSE (Root Mean Square Error) and hit rate were adopted. Among the techniques, the SVR with linear kernel guaranteed the best results, showing a good solution for the proposed problem.

Keywords— Machine Learning, Support Vector Regression, Autoencoder, Basic Oxygen Furnace.

Resumo— O objetivo deste trabalho é o desenvolvimento e análise de desempenho entre dois modelos de predição, baseados nas técnicas SVR (*Support Vector Regression*) e rede neural *Autoencoder*. Tais modelos visam à predição do teor de fósforo do aço em convertedores a oxigênio ao final do refino primário, em substituição à amostragem comumente realizada para análises laboratoriais, reduzindo o tempo do processo e, consequentemente, aumentando a produtividade da produção do aço. Para a modelagem utilizou-se um conjunto de variáveis formado por insumos e dados de instrumentação do refino. Utilizou-se a validação cruzada *k-fold* para obtenção de modelos generalistas e como métricas de desempenho adotou-se o RMSE (*Root Mean Square Error*) e taxa de acerto em relação à concentração final de fósforo. Dentre as técnicas, a SVR com *kernel* linear garantiu os melhores resultados, se mostrando uma boa solução para o problema proposto.

Palavras-chave— Aprendizado de Máquina, Regressão por Vetores de Suporte, *Autoencoder*, Forno Básico de Oxigênio

1 Introdução

À medida que a globalização da economia avança, as siderúrgicas enfrentam uma crescente concorrência. A qualidade do produto é muitas vezes um fator diferenciador entre os produtores, impulsionando a demanda pelo desenvolvimento de ferramentas avançadas para melhorar a qualidade durante o processo de fabricação de aço. Muitas técnicas de inteligência artificial e aprendizado de máquina foram utilizadas para o aumento da produtividade e a qualidade do processo supracitado. Laha et al. (2015) modelaram o rendimento de produção de aço utilizando técnicas como SVR, redes neurais artificiais, floresta randômica e *neuro-fuzzy* e apontaram a SVR como a de melhor desempenho para o estudo a que se propuseram. Visando o aumento da produtividade, Han and Liu (2014) utilizaram algoritmo membrana para a predição dos níveis de carbono em convertedores a Oxigênio. Com técnicas de visão computacional associadas às redes neurais, Liu et al. (2014) desenvolveram modelos para a predição do fim de sopro em convertedores a oxigênio.

Em certas aplicações uma concentração elevada de fósforo no aço pode comprometer seriamente a qualidade do produto final. Assim, um dos objetivos do refino é alcançar baixos teores deste elemento no aço. O refino do aço é dividido em dois estágios principais: o primário e o secundário. Este trabalho está concentrado no refino primário, que é a etapa na qual o ferro gusa, rico em carbono, é oxidado para conversão em aço. Isto é realizado no equipamento denominado convertedor, que inicialmente recebe os fundentes e o ferro gusa líquido. Na sequência, ocorre o sopro de oxigênio para a oxidação parcial das impurezas e a redução dos níveis de carbono.

Nos marcos de aproximadamente 70 % e 100 % do sopro ocorrem amostragens do aço em produção para análises laboratoriais, checagem das propriedades químicas e medição de temperatura (Unbehauen, 2009). Estas análises são realizadas por espectrômetros de emissão óptica e demandam um tempo que varia em torno de 5 minutos, o que corresponde a aproximadamente 20% do tempo de processamento. Apesar de serem necessárias no

formato convencional, as análises são limitantes para a produtividade, uma vez que seus resultados não são obtidos instantaneamente e geram um atraso na produção. Desta forma, o trabalho objetiva a predição de concentração do fósforo ao final do sopro de oxigênio, eliminando a necessidade das análises laboratoriais do processo. Propõe-se o desenvolvimento de preditores da concentração de fósforo ao fim do refino em convertedores a oxigênio (BOF, do inglês, *Basic Oxygen Furnace*), por meio das técnicas de aprendizado de máquina SVR e redes neurais *autoencoder*.

Recorrendo-se à literatura, estudos similares foram realizados com objetivo de estimar o fósforo utilizando as mais variadas técnicas de aprendizado de máquina. Han and Cao (2015) utilizaram a metodologia CBR (*Case-Based Reasoning*) aliada a um algoritmo genético e alcançaram a taxa de acerto de 42% considerando a faixa de erro de $\pm 20ppm$ (partes por milhão). Similarmente, Qiu and Dai (2014) adotaram o modelo neural RBF para predição do teor de fósforo em fornos AOD (*Argon Oxygen Decarburization*).

No mesmo segmento, Wang et al. (2012) utilizaram um método híbrido com modelo neural e *k-means* para a predição do fósforo. Por meio da clusterização de 21 variáveis independentes retiradas do processo de refino, alcançaram a taxa de acerto de 60% considerando também a tolerância de $\pm 20ppm$.

2 Referencial teórico

2.1 Support Vector Regression

A SVR é uma técnica de aprendizado não paramétrica, ou seja, não é necessário conhecer a distribuição de probabilidades do universo ou conjunto do qual os dados sob análise foram retirados. É oriunda do SVM (do inglês, *Support Vector Machine*), método de aprendizagem originalmente binário (Basak et al., 2007).

Seja um problema originalmente binário, no qual busca-se separar duas classes, tal como mostrado na Figura 1. Os dados são (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) , onde $x_i \in R^m$ é um vetor que pertence à classe $y_i \in \{-1; +1\}$. O SVM busca um hiperplano que separe as duas classes, tal como na Equação 1.

$$f(x) = wx + b \quad (1)$$

Onde w é o vetor de pesos e b é uma constante que define o desvio.

Como visto na Figura 1, se as classes são linearmente separáveis, podem existir diversos hiperplanos que resolvem o problema sem erros. Essas diversas soluções estão associadas comumente aos mínimos locais e, para eliminá-las, a SVM estabelece um hiperplano de classificação com a máxima distância de separação entre as amostras. Ele é

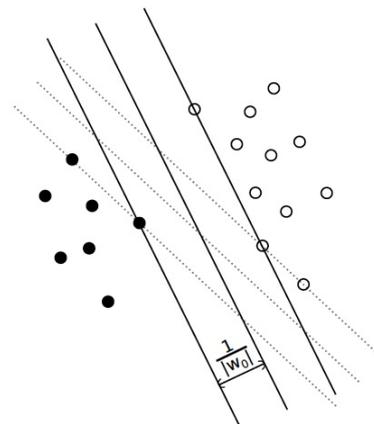


Figura 1: Hiperplanos, separando linearmente duas classes e de dois modos distintos. Fonte: Autores.

chamado hiperplano ótimo e faz com que as diversas soluções converjam para um mínimo global. A máxima separação entre as classes é definida como margem e as amostras que delimitam a margem são os vetores de suporte. A grande vantagem da SVM, portanto, é que o treinamento é equivalente à resolução de um problema de otimização quadrática na qual a solução é global e não há mínimo local (Laha et al., 2015).

Como na maioria dos casos práticos não é possível estabelecer uma separação por um hiperplano, foi introduzido por Cortes and Vapnik (1995) o conceito de variável de folga ξ_i , no qual permite-se um erro associado à classificação.

No problema de regressão adotam-se os mesmos conceitos da classificação, fazendo-se um pequeno ajuste: a cada amostra x_i da regressão, adiciona-se e subtrai-se um valor d do valor correspondente y_i . O hiperplano ótimo formado passará exatamente pelos valores y_i e com isto, aplica-se normalmente o algoritmo do SVM binário.

O ajuste do modelo é feito de forma que o erro ponto-a-ponto não seja maior que um dado valor ξ . Pode-se encontrar os pesos que satisfaçam a condição apresentada na Equação 2:

$$|y_i - f(x_i)| < \xi \quad (2)$$

para todo $i = 1, 2, \dots, m$ (Basak et al., 2007).

2.2 Rede Autoencoder

Os modelos *Autoencoder* são redes neurais situadas no contexto das redes de aprendizado profundo e que têm por objetivo reconstruir seus dados de entrada de forma não-supervisionada.

Matematicamente, a rede *autoencoder* inicialmente codifica os dados de entrada x por meio da função f utilizando as características mais relevantes da amostra. Em seguida, aproxima os

dados originais por \hat{x} através da função g (*decoder*):

$$y = f(w_1\hat{x} + a) \quad (3)$$

$$\hat{x} = g(w_2x + b) \quad (4)$$

onde w_1 e w_2 são as matrizes de pesos e a e b são os vetores *bias*.

A Figura 2 mostra a estrutura da rede *autoencoder*. Para que a *autoencoder* realize uma boa representação dos dados, a função f (*encoder*) deve capturar as características mais importantes das amostras. Por fim, os parâmetros do modelo obtido são otimizados de modo a minimizar o erro de reconstrução dos dados, representado pela função custo L .

Com as funções de codificação e decodificação estabelecidas, f e g , respectivamente, constrói-se o modelo de predição similar a uma rede neural convencional, treinando o modelo *autoencoder* de forma supervisionada e extraindo as principais características das entradas por meio da função *encoder* (Goodfellow et al., 2016).

Para o treinamento supervisionado, neste trabalho incorporou-se um modelo neural *softmax*, que possui como característica expressar sua saída em termos da distribuição de probabilidades.

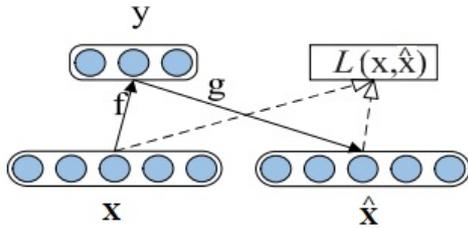


Figura 2: Estrutura simplificada da rede *autoencoder*. Fonte: Autores.

Existe o problema, no entanto, de o critério adotado pela função custo não garantir a extração das características mais relevantes. Um exemplo seria a solução trivial de simplesmente copiar a entrada para a saída. Visando eliminar situações desta natureza, introduziu-se o paradigma das *autoencoders* esparsas.

A esparsidade é caracterizada pela quantidade de neurônios na camada oculta significativamente maior que a dimensão de entrada, mas cuja ativação é escassa, ou seja, quando a quantidade de neurônios inativos é significativamente maior que a quantidade de neurônios ativos. De maneira informal, a esparsidade, além de contornar o problema da solução trivial confere uma característica de representação comprimida dos dados originais (Vincent et al., 2010).

A capacidade de compressão poderia ser realizada por outras técnicas amplamente conhecidas, como a PCA (do inglês, *Principal Component*

Analysis), por exemplo (Jolliffe, 2002). Porém, a *autoencoder* possui a grande vantagem de lidar com não-linearidades, situação presente no refino do aço.

3 Metodologia

Utilizou-se duas técnicas de regressão para a predição do fósforo, a SVR (*Support Vector Regression*) e um modelo *autoencoder* conjugado a uma camada *softmax*.

A escolha das técnicas se deu, sobretudo, por suas capacidades de tratar as não-linearidades presentes no processo metalúrgico. Conforme o trabalho de Lasheras et al. (2014), a SVR pode resolver com sucesso os problemas não-lineares presentes no ambiente industrial. No mesmo sentido, Goodfellow et al. (2016) explicitam a capacidade de compressão e manipulação de não-linearidades nas redes *autoencoder*.

3.1 Conteúdo da base de dados

A base de dados utilizada nas análises é composta de dados reais obtidos de um processo siderúrgico. As variáveis independentes representam algumas entradas inerentes ao refino primário do aço em convertedores a Oxigênio. Algumas delas são: medições de temperaturas do processo; pesos de matérias-primas e fundentes utilizados; medições de teores de determinados elementos químicos, tais como carbono e enxofre; dentre outras grandezas. A variável dependente, também denominada variável explicada, é a concentração de fósforo no aço produzido.

Uma vez que o refino do aço é um processo agressivo e conduzido a temperaturas que ultrapassam $1600\text{ }^{\circ}\text{C}$, há limitações na medição das variáveis e somente 29 delas foram utilizadas no treinamento dos algoritmos.

A base é composta por 4200 conjuntos de variáveis independentes e sua variável dependente associada (teor final de fósforo). A cada um desses conjuntos, designa-se o termo *batelada*. Por fim, utiliza-se os termos ‘banho’ para menção ao aço líquido incandescente em produção e ‘sub-lança’ para referência ao marco de 70% do sopro de Oxigênio, momento em que um equipamento realiza a amostragem do aço em produção para verificação das propriedades químicas via análises laboratoriais.

3.2 Condicionamento dos dados

Uma etapa relevante na modelagem é o condicionamento das amostras utilizadas no treinamento. Em geral, o objetivo do condicionamento é eliminar ou minimizar a influência de situações como falta de dados, *outliers* e evitar duplicidade de amostras. Essas garantias conferem maior fidelidade aos modelos, reduzem sua complexidade

computacional e flexibilizam o ajuste de parâmetros, se necessário (García et al., 2016). O primeiro filtro aplicado à base foi a eliminação de amostras inconsistentes. Devido à natureza do processo sob análise, diversas variáveis independentes não poderiam assumir valores nulos, tendo sido eliminadas as bateladas em que tal situação ocorresse.

Em seguida, fez-se a verificação da existência de *outliers* via critério de 1.5IQR (*Interquartile Range*), no qual quaisquer bateladas com valores situados fora da faixa $Q_1 - 1,5(Q_3 - Q_1)$ e $Q_3 + 1,5(Q_3 - Q_1)$ foram consideradas *outliers* (Riaz, 2015). Apesar de valores fora da faixa de 1.5 IQR serem amostras autênticas do processo, a remoção de *outliers* possui grande importância sobretudo na validação dos modelos, etapa em que as análises são feitas com base nos resíduos (Hair et al., 2013).

Finalmente, visando estabelecer os mesmos graus de importância entre as variáveis independentes, normalizou-se a base de dados. Utilizou-se a metodologia por máximos e mínimos equalizados, na qual as variáveis passam a assumir valores na faixa entre 0 e 1. Esta etapa é necessária haja vista que algumas das variáveis de entrada podem assumir valores desde a ordem de partes por milhão (*ppm*) até toneladas. Diante disto, o ajuste sem a normalização dos dados pode ser tendencioso, enviesando a relevância das variáveis para aquelas que possuam os maiores valores absolutos.

3.3 Modelagem

Para a modelagem fez-se a divisão da base original em 70% dos dados para treinamento/validação e 30% para teste. Às bateadas destinadas ao treinamento e validação, aplicou-se a validação cruzada *k-fold* com 10 partições, método amplamente empregado para assegurar a capacidade de generalização dos modelos (Liang and von Davier, 2014). As dez partições de validação são disjuntas entre si e contêm bateladas tomadas randomicamente.

Tanto a SVR quanto a rede neural *Autoencoder* possuem diversos parâmetros a serem ajustados, o que acaba comprometendo seu desempenho, se não escolhidos corretamente. Os principais parâmetros do SVR são *kernel* e ξ -insensível, enquanto que no modelo neural são a quantidade de neurônios, quantidade de camadas ocultas, funções de ativação e parâmetros de esparsidade. Visando garantir o melhor desempenho dos modelos e estabelecer uma análise sob condições equiparáveis, as melhores combinações dos parâmetros foram previamente ajustadas para cada uma das técnicas de forma empírica. Ou seja, os parâmetros supracitados foram combinados em iterações encadeadas, sendo escolhidos os que resultaram modelos com menor erro de validação.

Como melhores parâmetros encontrados, para

a SVR utilizou-se ξ -insensível de $2.22e^{-4}$ e *kernel* linear. O *kernel* gaussiano apresentou por vezes resultados superiores ao linear. Porém, durante as análises de validação, os melhores resultados do *kernel* gaussiano apontaram a presença de um sobreajuste do modelo (*overtraining*), pois quando confrontados com a partição de teste, degeneravam consideravelmente a eficiência. Por este motivo, elegeu-se o *kernel* linear por manter maior capacidade de generalização. Já para a rede *autoencoder*, foram utilizados 40 neurônios na camada oculta, funções de ativação lineares no *encoder* e *decoder*, proporção de esparsidade de 0.25, 1000 épocas, peso de regularização associado à função custo de 0.004 e regularização de esparsidade em 5.

Para validação dos parâmetros dos modelos, aplicou-se os testes de normalidade, homocedasticidade e correlação serial nos resíduos de estimação.

Por fim, no conjunto de teste utilizou-se o RMSE e a taxa de acerto como métricas para avaliação dos modelos obtidos. A taxa de acerto foi analisada em relação ao valor real do fósforo com diferentes níveis tolerância, dados em *ppm*.

3.4 Análise dos resíduos

Nesta seção são apresentados os testes de validação dos modelos via análise de resíduos ϵ . Os resíduos de predição são definidos como a diferença entre o valor real e o valor predito. Tais testes têm como objetivo evidenciar a não-tendenciosidade dos modelos e foram usados os resíduos obtidos nas 10 partições de validação *k-fold*.

Sejam duas variáveis independentes X_i e X_j , onde $i \neq j$, não há correlação serial dos erros se os desvios das estimativas Y_i e Y_j em relação aos seus valores reais não apresentam erros sistemáticos de comportamento. Matematicamente:

$$cov(\epsilon_i, \epsilon_j / X_{i,j}) = 0 \quad (5)$$

Considerando que os erros são descorrelacionados e que o valor esperado dos resíduos é zero, condição para a boa eficiência dos modelos de regressão (Antonakis and Dietz, 2011), a Equação 5 torna-se:

$$E(\epsilon_i \cdot \epsilon_j | X_{i,j}) - E(\epsilon_i | X_{i,j}) \cdot E(\epsilon_j | X_{i,j}) = 0 \quad (6)$$

Em resumo, os resíduos devem estar aleatoriamente dispersos, seguindo uma distribuição normal de média zero. Para isto, fez-se a análise gráfica e o teste estatístico de *Kolmogorov-Smirnov*. De modo a restringir comentários redundantes, somente uma das partições *k-fold* será mostrada na análise gráfica de dispersão e distribuição dos resíduos. Contudo, as mesmas considerações são válidas para as demais partições. As Figuras 3 e

4 mostram os resíduos de estimação estão aleatoriamente dispersos, seguindo uma distribuição normal de centro na origem.

Adicionalmente ao método gráfico, fez-se o teste estatístico de *Kolmogorov-Smirnov* para um intervalo de confiança de 95%. Os resultados comprovaram a análise gráfica, uma vez que resultou na aceitação da hipótese nula (H_0 : os dados seguem distribuição normal).

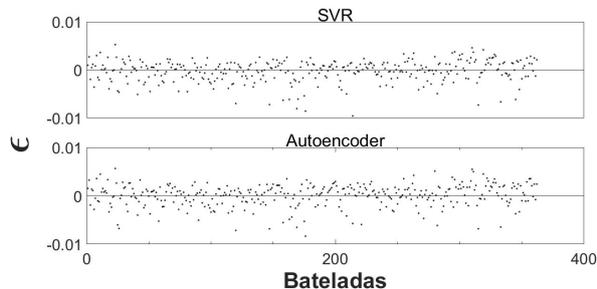


Figura 3: Dispersão dos resíduos

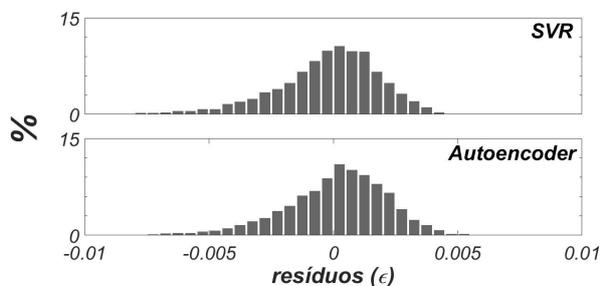


Figura 4: Histograma dos resíduos

Por fim, a Figura 5 mostra, por meio dos gráficos de caixa, a homocedasticidade dos resíduos de estimação, uma vez que as variâncias condicionais dos erros permanecem constantes ao longo das 10 partições avaliada. A homocedasticidade, assim como a normalidade é premissa para a aplicação dos testes estatísticos, dentre eles o de *Kolmogorov-Smirnov* citado anteriormente.

4 Resultados

Os resultados apresentados nesta seção referem-se ao total de 30% das bateladas destinadas ao teste dos modelos. Tais amostras não tiveram contato algum com a etapa de modelagem.

A Tabela 1 mostra os valores de RMSE obtidos pelos modelos. Nota-se uma equiparação entre as técnicas, igualdade esta que também se reflete no custo computacional. Ambas apresentaram custo de processamento similares. Entretanto, ao analisar as taxas de acerto da predição em relação ao teor de fósforo \pm tolerância, mostrada na Tabela 2, a SVR resulta nos melhores índices para uma faixa mais restrita do erro. Ou seja, a SVR resulta na melhor acurácia em torno da concentração real de fósforo.

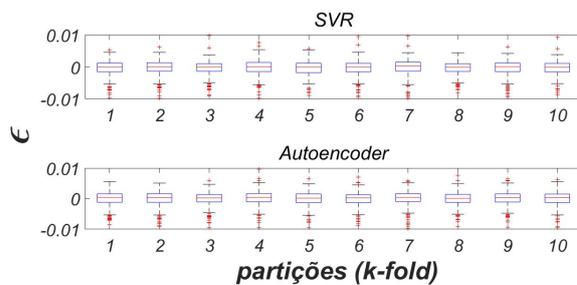


Figura 5: Análise de variâncias. Os gráficos de caixa mostram a uniformidade das variâncias (homocedasticidade), além de corroborarem os resultados das Figuras 3 e 4.

Tabela 1: RMSE dos modelos

	SVR	Autoencoder
RMSE	0.0030	0.0029

Destaca-se que quanto menor o intervalo de erro mais próxima do valor real é a predição do modelo. Nota-se que ambos os modelos são inviáveis para a faixa de tolerância de $\pm 10ppm$, uma vez que a eficiência das predições não ultrapassa 41%. Isto reflete o fato de que as variáveis independentes utilizadas na modelagem são insuficientes para descreverem integralmente o processo.

Contudo, para tolerâncias acima de $\pm 20ppm$ os modelos entregam melhores resultados do que grande parte dos estudos correlatos, alguns deles mencionados neste trabalho. Esta é uma das contribuições deste estudo, que está em ajustar dois modelos distintos que, mesmo com poucas variáveis independentes, conseguem entregar boas taxas de acerto da variável dependente.

A coluna intervalo de erro traz, em *ppm* a faixa de tolerância em relação ao valor real e dentro da qual os valores preditos são considerados acertos. Os resultados são mostrados considerando faixas de $[-10, +10]$ a $[-80, +80]ppm$.

Tabela 2: Taxas de acerto

Intervalo de erro (<i>ppm</i>)	SVR	Autoencoder	Diferença
$[-10, 10]$	41%	38%	3%
$[-20, 20]$	71%	66%	5%
$[-30, 30]$	83%	84%	-1%
$[-40, 40]$	90%	91%	-1%
$[-50, 50]$	94%	94%	0%
$[-60, 60]$	95%	96%	-1%
$[-70, 70]$	97%	97%	0%
$[-80, 80]$	97%	97%	0%

A Figura 6 mostra o gráfico do valor real de fósforo *versus* os resíduos. As linhas tracejadas representam os limiares de $\pm 20ppm$. Devido à na-

tureza do processo metalúrgico é vista uma concentração dos níveis de fósforo na faixa que se estende de 0.008% a aproximadamente 0.017%. Este comportamento acaba por estabelecer a polarização dos modelos, embora não-tendenciosos. Uma alternativa à esta polarização é a criação de modelos mais específicos nas subfaixas do universo original. Fica, portanto, esta sugestão como um dos trabalhos futuros.

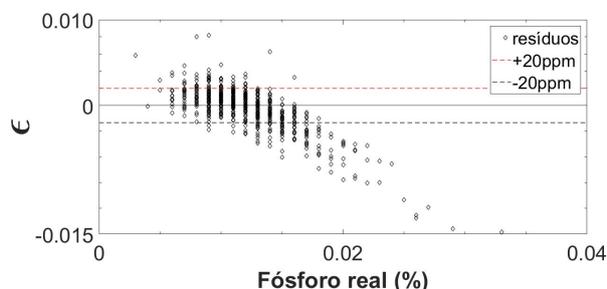


Figura 6: Resíduos *versus* fósforo real

5 Conclusões

O presente trabalho teve como objetivo o desenvolvimento de modelos para estimação da concentração de fósforo ao final do refino primário do aço em convertedores a Oxigênio. Dentre as técnicas, ambas apresentaram desempenho equiparáveis. No entanto, SVR se mostrou mais precisa no ajuste das predições em torno do valor real. Se comparado a estudos similares, dentre os quais os desenvolvidos por Han and Zhao (2011) e Shao et al. (2014), este trabalho trouxe resultados superiores.

Acredita-se, por meio das análises realizadas, que a taxa de acerto pode ser melhorada com o acréscimo de mais variáveis explicativas nos modelos, sugestão para trabalhos futuros. Dada a complexidade do processo metalúrgico e presença de não-linearidades, o uso limitado de 29 variáveis independentes é insuficiente para ter uma boa representação do processo. Por fim, dados como o perfil de sopro de Oxigênio, perfil de sopro dos gases inertes e altura relativa entre a lança e o banho metálico reúnem informações sobre a dinâmica do refino e podem ser melhor explorados para aumento da taxa de acerto e acurácia dos modelos.

Referências

Antonakis, J. and Dietz, J. (2011). Looking for validity or testing it? the perils of stepwise regression, extreme-scores analysis, heteroscedasticity, and measurement error, *Personality and Individual Differences* **50**(3): 409 – 415.

- Basak, D., Pal, S. and Chandra Patranabis, D. (2007). Support vector regression, **11**.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks, *Machine Learning* **20**(3): 273–297.
- García, S., Luengo, J. and Herrera, F. (2016). Tutorial on practical tips of the most influential data preprocessing algorithms in data mining, *Knowledge-Based Systems* **98**: 1 – 29.
- Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*, MIT Press. <http://www.deeplearningbook.org>.
- Hair, J., Black, W., Babin, B. and Anderson, R. (2013). *Multivariate Data Analysis*, Always learning, Pearson Education Limited.
- Han, M. and Cao, Z. (2015). An improved case-based reasoning method and its application in endpoint prediction of basic oxygen furnace, *Neurocomputing* **149**: 1245 – 1252.
- Han, M. and Liu, C. (2014). Endpoint prediction model for basic oxygen furnace steel-making based on membrane algorithm evolving extreme learning machine, *Applied Soft Computing* **19**: 430 – 437.
- Han, M. and Zhao, Y. (2011). Dynamic control model of bof steelmaking process based on anfis and robust relevance vector machine, *Expert Systems with Applications* **38**(12): 14786 – 14798.
- Jolliffe, I. (2002). *Principal Component Analysis*, Springer Series in Statistics, Springer.
- Laha, D., Ren, Y. and Suganthan, P. (2015). Modeling of steelmaking process with effective machine learning techniques, *Expert Systems with Applications* **42**(10): 4687 – 4696.
- Lasheras, F. S., Nieto, P. G., de Cos Juez, F. and Vilán, J. V. (2014). Evolutionary support vector regression algorithm applied to the prediction of the thickness of the chromium layer in a hard chromium plating process, *Applied Mathematics and Computation* **227**: 164 – 170.
- Liang, T. and von Davier, A. A. (2014). Cross-validation: An alternative bandwidth-selection method in kernel equating, *Applied Psychological Measurement* **38**(4): 281–295.
- Liu, H., Wang, B. and Xiong, X. (2014). Basic oxygen furnace steelmaking end-point prediction based on computer vision and general regression neural network, *Optik - International Journal for Light and Electron Optics* **125**(18): 5241 – 5248.

- Qiu, D. and Dai, W. J. (2014). Research on prediction model of end-point phosphorus content for aod furnace smelting ferrochrome based on rbf neural network, *Advanced Manufacturing and Information Engineering, Intelligent Instrumentation and Industry Development*, Vol. 602 of *Applied Mechanics and Materials*, Trans Tech Publications, pp. 769–772.
- Riaz, M. (2015). On enhanced interquartile range charting for process dispersion, *Quality and Reliability Engineering International* **31**(3): 389–398. QRE-13-0222.R2.
- Shao, Y., Zhou, M., Chen, Y., Zhao, Q. and Zhao, S. (2014). Bof endpoint prediction based on the flame radiation by hybrid svc and svr modeling, *Optik - International Journal for Light and Electron Optics* **125**(11): 2491 – 2496.
- Unbehauen, H. (2009). *CONTROL SYSTEMS, ROBOTICS AND AUTOMATION – Volume XIX: Industrial Applications of Control Systems-II*.
- Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y. and antoine Manzagol, P. (2010). Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion.
- Wang, H.-B., Xu, A.-J., Ai, L.-X. and Tian, N.-Y. (2012). Prediction of endpoint phosphorus content of molten steel in bof using weighted k-means and gmdh neural network, *Journal of Iron and Steel Research, International* **19**(1): 11 – 16.