

ANÁLISE MULTIVARIADA DE DADOS APLICADA AO ESTUDO DE PROJEÇÃO DE ESCÓRIA DE ACIARIA *

Bruna Helena Malovini Loiola¹
Henrique Silva Furtado²
André Vicente Salazar³

Resumo

O fenômeno de projeção de escória em um BOF representa prejuízo financeiro, desperdício de matéria prima, risco de acidentes e traz consequências ambientais. Arquivos de *Big Data* da ArcelorMittal Tubarão foram estudados com o intuito de investigar as principais variáveis associadas às projeções. Realizou-se o tratamento descritivo dos dados, por meio das técnicas de análise de *cluster*, análise de fatores, e de componentes principais com a finalidade de elaborar uma equação de regressão linear multivariada que pudesse modelar o fenômeno e prever as corridas com projeções. O tratamento estatístico foi realizado com apoio dos Softwares STATISTICA e NCSS. Foram feitos acompanhamento *in loco*, na empresa, que geraram dados de pesquisa advindos do uso de botão de controle de projeção, além da categorização das variáveis sobre as corridas dos convertedores. Observou-se que o uso do botão para controle de projeção, pelo operador, é mais frequente aos 9 minutos de sopro. Além disso, o teor de Si no gusa líquido e o peso da sucata de aço adicionados no convertedor estão entre as principais causas de projeção.

Palavras-chave: Projeção de escória; BOF; Análise Multivariada; *Big Data*.

MULTIVARIATE DATA ANALYSIS APPLIED TO THE STUDY OF SLOPPING IN STEELMAKING

Abstract

The phenomenon of slopping of slag in a BOF creates some problems for steelmaking, such as financial and material loss, accident and environmental risks. Big Data files from ArcelorMittal Tubarão were studied in order to identify the main causes of slopping using multivariate analysis that are not frequent techniques used for steelmaking studies yet. STATISTIC software was used with new techniques proposals in this area as cluster analyzes, factor analysis, and principal components, and NCSS software was used for multivariate regression analysis. It was observed the steelmaking process at the company. Also, information about the use of the slopping control button was verified, as well, variables about the converter. It was contacted that the use of the slopping control button by the operator is more frequent in the middle of the blow. In addition, the Si content in the hot metal and the weight of the steel scrap added in the converter are some of the main causes of the slopping.

Keywords: Slopping; BOF; Multivariate Analysis; Big Data.

¹ Engenheira Metalúrgica pelo IFES, e Mestranda em Engenharia de Materiais. REDEMAT (UFOP/UEMG), Ouro Preto, MG, Brasil.

² Engenheiro Metalúrgico, Dr. Especialista de Pesquisa e Desenvolvimento do Centro de Pesquisa da ArcelorMittal para a América do Sul. Vitória, ES, Brasil.

³ Prof. Me. Ed. Matemática pelo IFES. Professor de Matemática do IFES, Campus Vitória. Vitória, ES, Brasil.

1 INTRODUÇÃO

Projeção de escória é um fenômeno que ocorre em convertedores de aciaria. Esse é um termo utilizado para o evento que descreve quando a escória não é contida dentro de convertedores BOF (*Basic Oxygen Furnace*), ou seja, é o transbordamento de escória durante o processo de sopro de oxigênio [1].

O objetivo do processo BOF é produzir um aço base reduzindo as impurezas da carga metálica através de oxidação. A este aço são feitas adições de ligas necessárias a sua aplicação final [2]. Como resultado, ocorre uma grande geração de gases oriundo da oxidação do carbono e a formação de uma escória líquida decorrente da reação do produto da oxidação das demais impurezas com materiais fluxantes adicionados para estabilizar tais reações e proteger o refratário do convertedor [3]. A combinação da grande evolução gasosa com o desenvolvimento da escória (volume e química), segundo Fruehan et al. [4], provoca uma espumação. Sendo assim, se por um lado a espuma acelera as reações de refino, por outro, se não controlada, pode transbordar de forma violenta para a área externa do convertedor, ocorrendo o fenômeno da projeção.

Segundo Remus *et al.* [2], o controle de projeção é importante, uma vez que o sistema de captação de gases do convertedor, em geral, não suporta o excesso de gases que repentinamente são expelidos no momento da projeção. A consequência é que acabam sendo eliminados gases em forma de poluente ao ambiente externo à aciaria. Além disso, com a projeção há um risco de ocorrência de acidentes, e também se reduz o rendimento de produção de aço porque há perdas metálicas, uma vez que cerca de 10 a 20 % da quantidade total de massa metálica é suspensa na espuma [1].

A proposta desse trabalho é utilizar métodos de análises multivariadas, ainda pouco comuns na área de siderurgia, para efetuar uma análise de *Big Data* de tal forma a detectar as variáveis de influência nas ocorrências de projeções nos convertedores da ArcelorMittal Tubarão, além da elaboração de um modelo matemático que possa prever as corridas que terão projeção. Para este estudo são utilizadas propostas de técnicas clássicas, como as análises de *cluster*, de fatores, de componentes principais, e análises de regressões múltiplas, de componentes principais (CP) e logísticas, certamente nunca antes utilizadas no assunto projeção.

2 DESENVOLVIMENTO

2.1 Materiais e Métodos

Do dia 13 ao dia 30 de março de 2017 foram acompanhadas um total de 38 corridas nos três convertedores da ArcelorMittal Tubarão, sendo que, em 9, verificou-se a ocorrência do fenômeno de projeção de escória. O intuito da visita à aciaria foi conhecer *in loco* as projeções, os problemas que podem ser gerados, e elaborar registros de fatos observados que pudessem ser relevantes para o estudo da investigação das causas do fenômeno. Os registros foram feitos por meio de anotações escritas, imagens fotográficas e vídeos de curta duração.

Além disso, obteve-se, junto à empresa ArcelorMittal Tubarão, cópia de arquivos de dados dos convertedores 1 e 2 referentes ao período de 1º de janeiro de 2016 a 31 de março de 2017 que descreviam, por datas, os horários de início e fim do processo de sopro (de oxigênio no convertedor), bem como data e horário de acionamento do botão de controle de projeção. Em geral, este botão é acionado

toda vez que na corrida, durante o sopro de oxigênio, há a ameaça de projeção de escória. Sendo assim, o botão tem o objetivo de tentar controlar a projeção, através de ajustes do sopro como a diminuição da vazão de oxigênio e controle da altura de lança. As ações desencadeadas pelo acionamento deste botão perduram por 90 segundos, e caso o operador necessite de mais tempo de controle, ele deve acionar o botão novamente. Fez-se uma análise dos instantes, contados a partir do início da corrida, em que o botão foi acionado (nas corridas em que houve o acionamento do botão) para verificar se havia algum padrão neste processo, tal como a identificação de um instante em que houvesse maior incidência de projeção.

Obteve-se, ainda, junto à empresa ArcelorMittal Tubarão, dados de aciaria referentes ao período de 1º de janeiro de 2016 a 07 de maio de 2017. Tais dados mostram cerca de 30 mil corridas, entre os convertedores 1, 2 e 3 da usina, com informações detalhadas de cada uma como, por exemplo, data da corrida, hora, temperatura do banho, teores dos principais elementos presentes antes e depois do sopro, volume total de oxigênio utilizado, totalizando cerca de 160 variáveis, com algumas informações recolhidas de forma automatizada e outras, adicionadas pela operação. De início, foi feita uma filtragem das 160 informações, restando, por fim, 48 variáveis de interesse. Foi feita, em seguida, uma matriz em planilha no Excel com as informações dos dados. Estes, assim, foram padronizados para poderem ser comparados sem interferências provocadas pela ordem de grandeza de suas informações (ex.: temperatura do gusa, peso do briquete, porcentagem de Mn no gusa) utilizando-se, para isso, a equação:

$$X_{a'} = \frac{X_a - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

Onde $X_{a'}$ é o valor padronizado, X_a é o valor do número a ser substituído, X_{\min} é o valor mínimo do conjunto de dados da variável que está sendo padronizada e X_{\max} é o respectivo valor máximo. Após a padronização aplicou-se as técnicas estatísticas de análise de *cluster*, análise de fatores e análise de CP, utilizando o software *STATISTICA versão 7.0 da StatSoft*. A análise de *cluster* é uma técnica da Estatística Descritiva que agrupa as variáveis com base em alguns critérios pré-definidos disponibilizados pelos softwares, gerando um dendograma que é uma espécie de diagrama de árvore, o qual possibilita a realização de uma análise de afinidade entre as variáveis independentes e destas com a variável dependente (conhecida como variável resposta). A variável resposta já veio na planilha de dados obtida junto à ArcelorMittal com seu resultado expresso numa forma semelhante a um “Item Likert” [5], que é um tipo de resposta psicométrica muito comum em questionários de pesquisas de opinião. Ao atribuir valores a essa variável, o operador (funcionário da empresa que monitora o processo na aciaria) inseria, com base em sua observação, um valor numérico representativo da intensidade da projeção avaliada, a saber: 0 = sem projeção; 1 = projeção média; 2 = projeção forte; 3 = projeção muito forte. A planilha de dados obtida possui duas de suas colunas contendo dados da variável resposta: A primeira (denominada variável 31) avalia se a projeção ocorreu até a primeira metade do sopro e a segunda (denominada variável 32) avalia da metade até o final do processo. Sendo assim, essas duas determinam qual agrupamento deve ser selecionado para as próximas etapas de análises, isto é, onde elas estão agrupadas determina o foco da investigação.

A técnica estatística da análise de fatores foi aplicada com a finalidade de verificar se haveria possibilidade de se reduzir o número de variáveis independentes do conjunto de dados a partir da criação de novas variáveis, chamadas de fatores; e

assim, determinar quantas das novas variáveis obtidas deveriam ser analisadas, bem como determinar qual seria a correlação entre as variáveis originais e os fatores, e quais os índices de contribuição das variáveis originais nos fatores. Por sua vez, a uso da técnica estatística da análise de CP foi aplicada por substituir as variáveis originais por outra categoria de variáveis, conhecidas como variáveis sintéticas, as quais possibilitaram comprovar qual era o número ideal de fatores a serem analisados (fatores obtidas na técnica anteriormente citada), uma vez que os fatores (as novas variáveis) são as CPs.

Os resultados encontrados a partir do *STATISTIC*, ou seja, as variáveis mais contribuintes para as CPs, foram aplicados no software de estatística *NCSS* versão 11 para teste. O objetivo foi obter um modelo de equação que prevê projeção a partir de análises de regressão linear múltipla, de regressão por componentes principais (RCP) e de regressão logística. As variáveis respostas do *STATISTICA* foram utilizadas nesta etapa porque demonstram a simplificação de um número grandioso de variáveis, sendo possível excluir as variáveis que não possuem efeitos significativos sobre a projeção. Antes de carregar os dados no software, a matriz com as variáveis que mais influenciam a projeção foi montada com as corridas. Depois disso, as duas variáveis de projeção foram transformadas em apenas uma, formada por 0 para não projeção e 1 para projeção, e assim, uma única variável foi considerada como dependente e todas as outras, independentes. Assim, importou se a matriz com as variáveis não padronizadas e as corridas no software. Após selecionar a variável dependente e as independentes, o programa é capaz de gerar um relatório com dados e gráficos sobre a regressão. Primeiro se fez a RCP que avalia dados de regressão múltipla que possuem multicolinearidade, que quando ocorre, as estimativas dos mínimos quadrados são imparciais, mas suas variações são grandes, e podem, por tanto, estar longe do verdadeiro valor. Assim, a RCP reduz o erro padrão, ou seja, espera-se que tenha estimativas mais confiáveis [6].

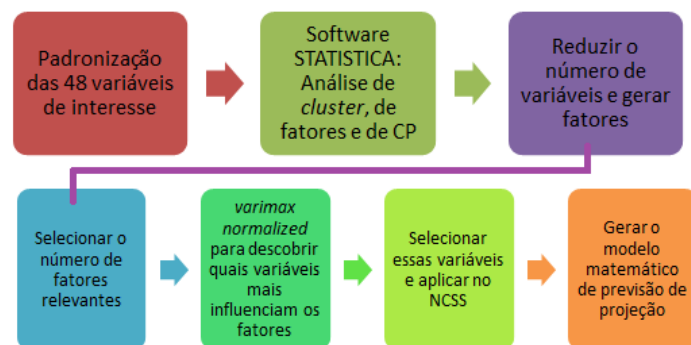


Figura 1. Fluxograma do resumo do processo realizado a partir das variáveis.

Em seguida, para que fosse possível uma comparação de resultados, fez-se a regressão múltipla, que analisa o relacionamento entre variáveis, sendo uma a chamada dependente (critério) e as outras as chamadas independentes (preditoras). Sua finalidade é utilizar as variáveis que os valores são conhecidos, as independentes, para prever os valores da dependente. O modelo de regressão é formado pelo conjunto de variáveis independentes ponderadas, que é uma combinação linear das variáveis independentes que melhor prediz a variável [7]. A regressão logística também foi realizada, uma vez que a variável dependente Y também pode ser classificada como uma variável binária, pois só assume dois valores, tal como: sim ou não, 0 ou 1, ou seja, fracasso ou sucesso [8]. Por fim, o modelo matemático para prever projeção foi feito e testado aplicando os dados de

todas as corridas do convertedor 1 na equação gerada. Um fluxograma na Figura 1 explica o resumo do processo aplicado.

2.2 Resultado e discussão

2.2.1 Análise das corridas acompanhadas na aciaria

Durante o acompanhamento das corridas na aciaria, percebeu-se que os operadores tentam ao máximo evitar projeções, já que elas também aumentam o tempo de corrida e podem atrasar a dinâmica de horários. Os padrões de sopro já são calculados automaticamente pelo sistema de acordo com a composição do gusa. Porém, durante a corrida, podem acontecer situações adversas em que o padrão de sopro não é suficiente para manter a corrida sem projeção, e então o operador precisa interferir. Existem casos em que o operador reduz a vazão de oxigênio, muda a altura de lança e o tipo de sopro, e outras vezes, utiliza-se o botão de projeção que tende a controlar o sopro.

Quando essas atitudes não são suficientes e acontece a projeção, o operador interrompe o sopro, espera estabilizar a situação no convertedor e depois volta a soprar. Pode ser comum acontecer de a corrida voltar a projetar após a continuidade do sopro após a interrupção, e então todo o processo para controle é repetido.

Algumas imagens de projeção acompanhadas podem ser vistas na Figura 2. Em certas situações, o chamado *Dog House* (portão que protege o convertedor) está fechado por questões de segurança.

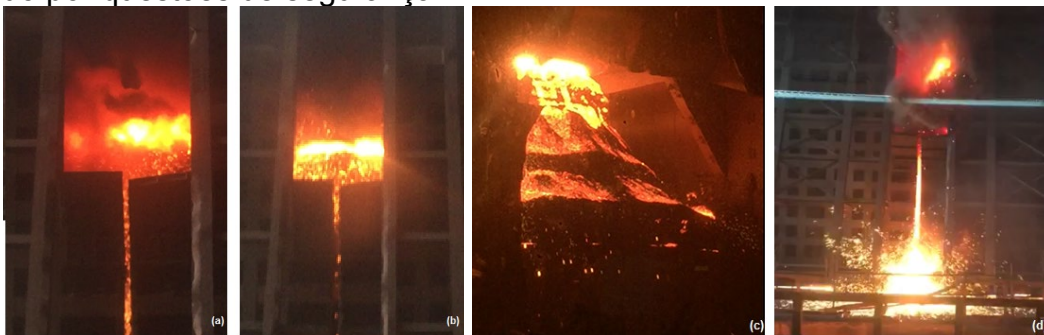


Figura 2. Projeção de escória no convertedor BOF (imagens a, b, c, e d).

2.2.2 Análise do botão de controle de projeção

De acordo com os dados cedidos, há a informação de que no convertedor 1 houve o uso do botão 1217 vezes, em 774 corridas diferentes; e no convertedor 2 o botão foi acionado 6009 em 1312 corridas distintas. A partir de então, foi montado um gráfico representado na Figura 3 que demonstra em qual momento os operadores mais utilizam o botão, o que indica que houve ameaças de projeção.

Normalmente, o sopro da ArcelorMittal Tubarão dura cerca de 18 minutos. Assim, no gráfico da Figura 3, representada pelas barras de azul, no meio do sopro, aos 9 minutos é onde os dados apontaram quando mais se acionou o botão de controle de projeção, com 801 vezes. Em segundo lugar, vem aos 10 minutos com 771 vezes, e depois aos 8 minutos com 746 usos. Esses tempos seguem uma tendência semelhante ao citado por Cicutti, Valdez, *et al.* [9], pois justamente pouco antes da metade do sopro que é o período propenso à máxima descarburização e a altura máxima da espuma a ser atingida, representada na Figura 3 pela linha preta, e conseqüentemente, é a etapa que há maiores riscos de projeção.

De acordo com a análise dos dados, foi feita a Tabela 1 que demonstra, nas mesmas datas, as corridas que obtiveram projeção e quais foram usados o botão de

controle de projeção. Nota-se que, o convertedor 2 possui porcentagens maiores que as do convertedor 1, tanto para o número de projeções quanto para o número de vezes do uso do botão. Vale salientar, também, que nem todas as corridas que ocorrem projeção são computadas pelo operador, uma vez que esse tipo de anotação é dependente de interpretação. Uma prova disso foi que nenhuma das 9 corridas que foram acompanhadas ocorrendo projeção estavam nos dados analisados como corridas com projeção e nem mesmo que foram utilizados o botão de controle. Isso indica que os dados analisados sobre a projeção podem não ser perfeitamente fidedignos, e que podem gerar erros de interpretação. Sendo assim, a pesar de os dados da empresa demonstrarem que ainda ocorrem muitas projeções (cerca de 12% – valor considerado alto), esse valor pode ser maior.

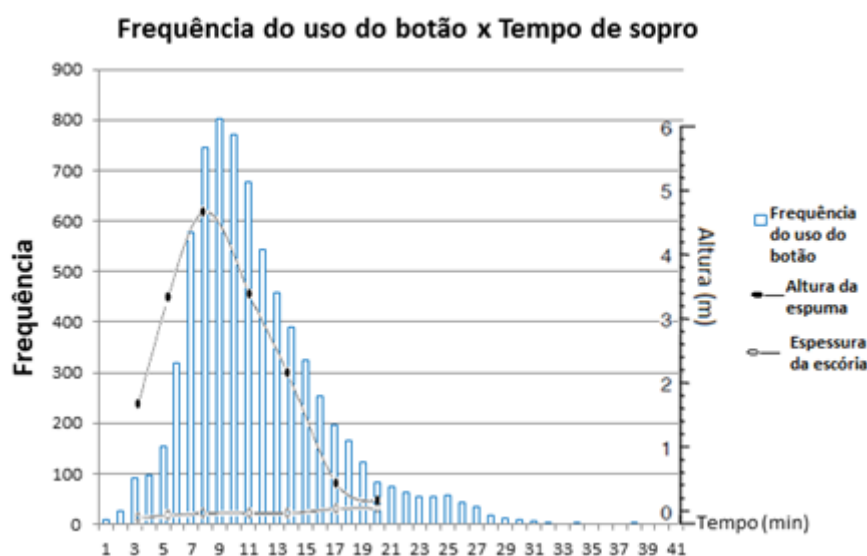


Figura 3. Gráfico do histograma de número de vezes que o botão foi apertado em relação ao tempo de sopro juntamente com a altura da espuma durante o tempo de sopro trazido da literatura.

Tabela 1 – Dados dos números de corridas, projeções, uso do botão e porcentagens

| | Convertedor 1 | Convertedor 2 |
|---------------------------------------|---------------|---------------|
| Total de corridas | 9147 | 9362 |
| Total de corridas com projeção | 1121 | 1215 |
| % de corridas com projeção | 12,26% | 12,98% |
| Total do uso do botão | 774 | 1312 |
| % do uso do botão | 8,46% | 14,01% |

A Tabela 1 ainda induz afirmar que, no convertedor 2, o uso do botão foi eficaz em algumas corridas. Assim, em determinadas situações o botão foi utilizado e surtiu efeito impedindo que algumas projeções ocorressem, ou que estas fossem controladas, uma vez que o valor do uso de botão é maior que o número de projeções ocorridas. Já no convertedor 1, ao contrário, existe uma diferença significativa entre o uso do botão e a ocorrência de projeção. O número do uso do botão foi menor do que o número de ocorrências de projeções, sugerindo que as projeções podem ter ocorrido de forma fraca e o operador optou por não o acionar. Além disso, pode sugerir também que algumas corridas com projeções ocorreram repentinamente, sem nem mesmo dar tempo de o operador usar o botão para controle.

2.2.3 Análise das causas de projeção

Foram selecionadas 48 variáveis. Fez-se uma análise mais específica só para o convertedor 1, para uma comparação de dados. A partir do STATISTICA, decidiu-se fazer um intercepto próximo a 1,8 à distância no dendograma para obter um grupo com número aceitável de variáveis para análise. Cada número abaixo (horizontalmente) na Figura 4 identifica uma variável contida nos dados, portanto obteve-se o seguinte agrupamento:

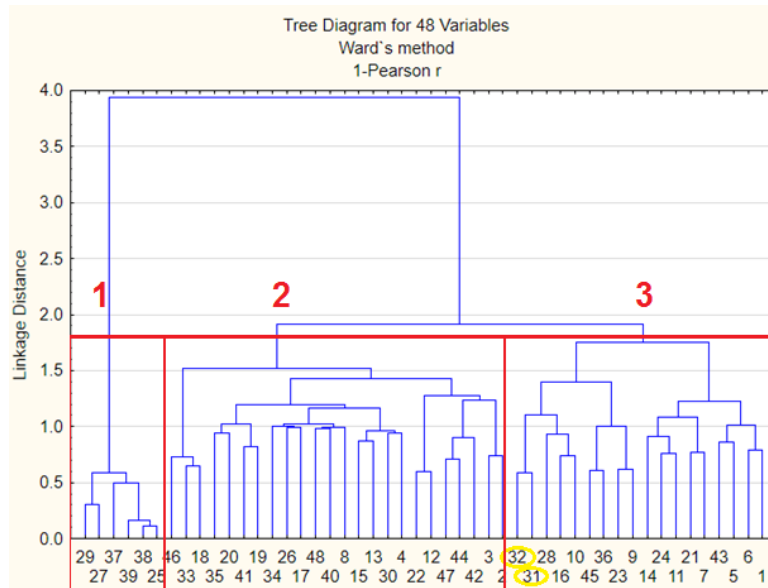


Figura 4. Dendogram formado a partir da análise de *cluster* de 48 variáveis

Todas as 18 variáveis contidas no agrupamento número 3, da Figura 4, foram para a próxima etapa da análise porque as variáveis de projeção estão melhores conectadas. Assim, com base no uso da ferramenta de análise fatorial no software, para a redução do número de variáveis, obteve-se a seguinte tabela na Figura 5:

| Eigenvalues of correlation matrix, and related statistics (CONV) | | | | |
|--|------------|------------------|-----------------------|--------------|
| Active variables only | | | | |
| Value number | Eigenvalue | % Total variance | Cumulative Eigenvalue | Cumulative % |
| 1 | 2.327693 | 12.93163 | 2.32769 | 12.9316 |
| 2 | 1.795781 | 9.97656 | 4.12347 | 22.9082 |
| 3 | 1.728470 | 9.60261 | 5.85195 | 32.5108 |
| 4 | 1.314738 | 7.30410 | 7.16668 | 39.8149 |
| 5 | 1.268747 | 7.04860 | 8.43543 | 46.8635 |
| 6 | 1.161313 | 6.45174 | 9.59674 | 53.3152 |
| 7 | 1.057765 | 5.87647 | 10.65451 | 59.1917 |
| 8 | 0.977300 | 5.42945 | 11.63181 | 64.6212 |
| 9 | 0.902424 | 5.01347 | 12.53423 | 69.6346 |
| 10 | 0.795898 | 4.42165 | 13.33013 | 74.0563 |
| 11 | 0.762235 | 4.23464 | 14.09237 | 78.2909 |
| 12 | 0.746953 | 4.14974 | 14.83932 | 82.4407 |
| 13 | 0.660022 | 3.66679 | 15.49934 | 86.1075 |
| 14 | 0.587698 | 3.26499 | 16.08704 | 89.3724 |
| 15 | 0.559114 | 3.10619 | 16.64615 | 92.4786 |
| 16 | 0.520060 | 2.88922 | 17.16621 | 95.3678 |
| 17 | 0.488003 | 2.71113 | 17.65422 | 98.0790 |
| 18 | 0.345784 | 1.92102 | 18.00000 | 100.0000 |

Figura 5. Tabela gerada no STATISTICA com o número de autovalores, variância total, autovalor acumulado e porcentagem cumulativa para cada autovalor.

Para 18 variáveis, o software gerou 18 fatores, porém, nem todas serão consideradas, pois o objetivo é reduzir o número de variáveis. Assim, para análises estatísticas, considera-se, geralmente, até o fator que possua a porcentagem acumulativa (*Cumulative %*) maior que 70, logo, até o fator 9. Na tabela, a primeira coluna (*eigenvalue*) representa os autovalores de cada fator, a segunda representa a variância total de cada um, o terceiro e o quarto representam, respectivamente, os autovalores acumulados e a porcentagem acumulativa de cada fator.

De acordo com a análise das CPs, o mesmo conteúdo da Figura 5 foi obtido. Assim, foi possível confirmar o número de CPs (fatores) a serem analisados como influências de projeção.

A etapa seguinte da análise é demonstrada na Figura 6. A imagem foi feita a partir da análise chamada *varimax normalized* que verifica quais variáveis originais mais contribuem para a formação dos fatores.

| Factor Loadings (Varimax normalized) (CONV. 1, 2, 3 in 1 - grupo) | | | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Extraction: Principal components | | | | | | | | | |
| (Marked loadings are > .700000) | | | | | | | | | |
| Variable | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 | Factor 6 | Factor 7 | Factor 8 | Factor 9 |
| 1 | 0.002412 | 0.018820 | 0.984620 | 0.007564 | -0.034268 | 0.029991 | 0.018509 | -0.041212 | -0.010515 |
| 5 | 0.031031 | 0.092302 | 0.018402 | -0.012371 | -0.055886 | -0.002957 | 0.986378 | 0.022320 | -0.007852 |
| 6 | 0.021279 | 0.003656 | 0.098307 | 0.014286 | -0.046230 | -0.002984 | -0.014593 | 0.031172 | -0.018510 |
| 7 | 0.010629 | -0.975446 | -0.019281 | -0.017011 | 0.012766 | 0.057017 | -0.095492 | -0.078983 | 0.005053 |
| 9 | 0.102810 | 0.039184 | -0.025355 | 0.027584 | 0.039428 | 0.049019 | -0.058375 | 0.076663 | 0.191187 |
| 10 | 0.142271 | 0.042309 | 0.012036 | 0.072931 | -0.077115 | -0.142466 | 0.034181 | 0.025717 | 0.018327 |
| 11 | 0.125507 | -0.090330 | -0.037369 | 0.013909 | -0.076027 | 0.114109 | 0.042996 | -0.011758 | 0.018569 |
| 14 | -0.082438 | 0.012505 | 0.034533 | -0.016996 | -0.982776 | 0.052758 | 0.056548 | -0.023706 | -0.002989 |
| 16 | -0.125442 | 0.059518 | -0.032334 | 0.070229 | 0.056721 | -0.960960 | 0.003304 | 0.076517 | -0.033795 |
| 21 | 0.060741 | -0.106014 | -0.015133 | -0.021321 | 0.029200 | 0.034995 | -0.000387 | -0.005297 | -0.001700 |
| 23 | 0.063879 | -0.005356 | -0.010675 | 0.001152 | 0.002799 | 0.032077 | -0.007733 | -0.007359 | 0.975973 |
| 24 | 0.042049 | -0.046082 | 0.090046 | 0.016503 | -0.052034 | -0.045745 | 0.052003 | 0.102688 | -0.028771 |
| 28 | 0.037329 | 0.077505 | -0.041547 | 0.032515 | 0.023818 | -0.071593 | 0.022623 | 0.981889 | -0.006672 |
| 31 | 0.076318 | 0.008179 | 0.013174 | 0.206160 | 0.019844 | -0.037406 | -0.010567 | 0.043837 | 0.013915 |
| 32 | 0.054093 | 0.017293 | 0.007828 | 0.968907 | 0.017562 | -0.068115 | -0.012968 | 0.033454 | 0.001282 |
| 36 | 0.163510 | 0.010330 | 0.033560 | 0.041253 | -0.011180 | 0.004745 | 0.001381 | 0.041723 | 0.087141 |
| 43 | 0.011603 | -0.064029 | 0.077776 | -0.024130 | 0.042787 | 0.024700 | 0.073245 | 0.009343 | 0.002890 |
| 45 | 0.916573 | -0.012754 | 0.002850 | 0.063931 | 0.102072 | 0.144781 | 0.038355 | 0.045409 | 0.076820 |

Figura 6. Tabela de contribuição das variáveis originais em cada fator.

As variáveis destacadas em vermelho são aquelas que possuem valores significativos (maiores que 0,7) e que mais contribuem para a formação de cada fator. Logo, para o fator 1, a variável 45 é a com maior representatividade, já para o fator 2, é a variável 7, por exemplo. Observa-se que em todos os fatores, apenas uma variável tem uma contribuição expressiva. Assim, as variáveis de maior contribuição, para cada fator, estão presentes na Tabela 2:

Tabela 2. Variáveis que mais influenciam nas projeções do convertedor 1

| Número da variável | Fator | Nome da variável |
|--------------------|-------|--|
| 45 (ou C43) | 1 | Si no gusa líquido (%) |
| 7 (ou C7) | 2 | Peso sucata de aço comprada tipo a (t) |
| 1 (ou C1) | 3 | Peso de sucata de gusa lingotado (t) |
| 14 (ou C14) | 4 | Peso de SiC adicionado durante o sopro de O ₂ (t) |
| 16 (ou C16) | 5 | Peso de dunito adicionado durante o sopro de O ₂ (t) |
| 5 (ou C5) | 6 | Peso sucata de aço tipo A (t) |
| 28 (ou C28) | 7 | Volume de argônio entre a sublança e o final do sopro (Nm ³) |
| 23 (ou C23) | 8 | Peso minério no início do sopro (t) |
| 21 (ou C21) | 9 | Peso de briquete misto adicionado durante o sopro de O ₂ (t) |

É possível observar que os resultados foram de acordo com o esperado segundo revisões da literatura. Um teste com os três convertedores juntos também foi realizado, porém surgiram variáveis menos apropriadas para a situação. Outros

testes com os convertedores 3 e 4 foram feitos, e os resultados também foram bem similares ao convertedor 1, com algumas mudanças na ordem em que as variáveis aparecem. Então, vale considerar a análise do convertedor separadamente, devido às particularidades de cada um.

De acordo com o obtido no convertedor 1, portanto, a presença do Si no gusa líquido está relacionada a projeção de acordo com sua quantidade. Quanto maior o percentual deste elemento, maior o volume de escória e maior a geração de gás. Sobre as sucatas, vale ter atenção na sua composição. Quanto maior o nível de impurezas mais a geração de óxidos formados na escória e mais propensão à projeção devido ao aumento do volume da escória.

Sobre os pesos de SiC e dunito, ou seja, o excesso de adição dos fluxantes, podem gerar o acréscimo da viscosidade e inibir a liberação dos gases das reações, o que também auxilia no crescimento da escória.

Os resultados de causas também incluem a adição de refrigerantes, (briquete e minério), o que aumenta a contribuição de oxigênio para a reação de descarburização na emulsão. Além disso, causa também o aumento da viscosidade da emulsão e seu volume devido à diminuição da temperatura. Sendo assim, é preciso ficar atento a um controle melhor das adições na tentativa de se evitar as projeções.

O volume de argônio entre a sublança e o final do sopro, apesar de estar envolvido no agrupamento, não está associado ao momento da projeção. O motivo de essa variável ter sido associada precisa ser investigado.

Em resumo, o uso da análise de *cluster*, análise de fatores e análise de CPs não indica causa-efeito das variáveis na projeção. Apenas agrupa variáveis correlacionadas à projeção, e reduz o número de variáveis existentes para análise.

2.2.4 Modelamento

Os resultados da regressão linear geraram um modelo matemático para a previsão de corridas com projeção. O primeiro teste feito foi a partir da RCP. Esta análise gera dados e, automaticamente, compara com o uso da técnica de mínimos quadrados de regressão linear múltipla. As variáveis respostas da Tabela 2 seguiram para serem analisadas no software NCSS e gerarem o modelo matemático. Assim, obtiveram-se os seguintes dados da Figura 7 a partir da regressão.

| Variável Independente | Coefficientes por Componentes Principais | Coefficientes Linear Múltipla |
|-----------------------|--|-------------------------------|
| Intercepto | -0.1640248 | -0.2897372 |
| C1 | 0.001348922 | 0.002465562 |
| C5 | -0.006706265 | -0.002775125 |
| C7 | -0.0009308039 | -0.0007914196 |
| C14 | 0.8312411 | 0.06920161 |
| C16 | 0.0793068 | 0.05867045 |
| C21 | 0.005540662 | -0.002192037 |
| C23 | 0.1338862 | -0.003575445 |
| C28 | 0.002590261 | 0.001236581 |
| C43 | 0.5056211 | 1.036754 |
| R ² | 0.5499 | 0.6797 |

Figura 7. Comparação dos coeficientes gerados por componentes principais e por mínimos quadrados.

Na primeira coluna se encontram o intercepto e as nove variáveis selecionadas. A segunda coluna identifica o coeficiente de cada variável por RCP e a terceira coluna mostra o coeficiente de cada variável por regressão múltipla. O R², estatisticamente, informa sobre quão bem o ajuste de um modelo está adequado, ou seja, sobre como a linha de regressão está próxima dos dados reais. Sendo assim, quanto mais o R²

se aproxima de 1, mais a linha de regressão está mais perfeitamente ajustada aos dados [10]. Vale mencionar, que outros testes com maior número de variáveis foram feitos, contudo, os resultados de R^2 estavam irrelevantes, com valores menores que 10%. Concluiu-se, assim, que com a utilização de menos variáveis, o resultado melhorou.

De acordo a Figura 7, percebe-se, então, que a linha de regressão linear múltipla está melhor ajustada aos dados quando comparada a RCP, uma vez que apresenta um valor de 67,97% de ajuste (considerado alto para dados industriais) contra 54,99%.

Durante a análise do programa NCSS, para a RCP, o próprio software realiza alguns cálculos e os valores das variáveis são padronizadas (utilizando a equação 1) e depois, em outro certo momento, são despadronizadas (voltam ao valor da variável inicial), o que pode ter gerado, ao final, alguma perda de informação, e assim, gerado um valor de R^2 menor.

Tomando a regressão linear múltipla, portanto, como a mais adequada dentre as duas, tem-se o seguinte modelo matemático gerado para previsão de corridas com projeções, dado como o índice de projeção, que indica a presença, ou não, de projeção:

$$M_{pi} = a + \sum B_i X_i \quad (2)$$

Onde M_{pi} é a variável dependente que é a de projeção (com valor de 0 para não projeção ou 1 para projeção), a é o intercepto igual a $-0,289737204931097$, e os valores dos coeficientes B e das variáveis independentes X encontram-se na Tabela 3:

Tabela 3. Tabela dos valores dos coeficientes e variáveis independentes

| Coeficiente (B_i) | Variável independente (X_i) |
|-----------------------|--|
| 0,00246556241440681 | C1 Peso de sucata de gusa lingotado (t) |
| 0,00277512443551416 | C5 Peso sucata de aço tipo A (t) |
| 0,000791419669969427 | C7 Peso sucata de aço comprada tipo a (t) |
| 0,0692016107705815 | C14 Peso de SiC adicionado durante o sopro (t) |
| 0,058670454743772 | C16 Peso de dunito adicionado durante o sopro (t) |
| 0,0021920372144974 | C21 Peso de briquete misto adicionado durante o sopro (t) |
| 0,0035754449492737 | C23 Peso minério no início do sopro (t) |
| 0,0012365805947284 | C28 Volume de argônio entre a sublança e o final do sopro (Nm^3) |
| 1,0367539367323 | C43 Si no gusa líquido (%) |

Para testar a equação obtida, utilizaram-se os dados das corridas do convertedor 1 para aplicação dos valores nas variáveis. Separaram-se, a partir dos dados cedidos, as corridas com projeção e corridas em que não houve projeção. Os resultados para M_{pi} estão presentes, em resumo, na Tabela 4:

Tabela 4. Principais valores encontrados para a variável projeção (M_{pi}) a partir dos dados cedidos das corridas

| | Corridas com projeção | Corridas sem projeção |
|---------------|-----------------------|-----------------------|
| Máximo | 0,968555696 | 0,867313 |
| Mínimo | -0,184916065 | -0,30137 |
| Média | 0,231076905 | 0,115488 |

Sabe-se que o ideal a ser encontrado seria 1 para corridas com projeção e 0 para corridas sem projeção. Porém, como se pode observar nos dados, os valores são adversos.

A imprecisão desses valores pode estar diretamente ligada à subjetividade do valor real de M_{pi} , já que essa variável, nos dados edidos, não é precisa porque depende da interpretação do operador. Assim, podem existir corridas em que houve projeção e o valor da variável foi anotado como 0. Vale considerar, ainda, que a reta de regressão se ajusta em 67,97%, ou seja, também há os valores que estão fora deste ajuste.

Além disso, a variável projeção não é contínua, e é considerada categórica. Assim, esta variável aceita apenas dois valores, 0 e 1, o que pode contribuir para o baixo desempenho dos resultados da equação. Levando esse fato em consideração, fez-se, portanto, a regressão logística no NCSS por este ser um método para variáveis categóricas.

O resultado da regressão logística, porém, deu ainda menos satisfatório, com valor de R^2 igual a 10,90%, como pode ser visto na Figura 8. Provavelmente, devido, mais uma vez, a falta de precisão dos valores da variável projeção. A automatização da anotação de tal variável na empresa, portanto, poderia gerar resultados com melhor exatidão durante a aplicação destas técnicas. Ou ainda, o acompanhamento de um número considerável de corridas, para que se tenham dados um pouco mais autênticos sobre projeção, para que se possam aplicar suas informações na regressão, para a verificação de um novo R^2 .

| Item | Valor |
|-------------------------|---------------|
| Y Variable | M_{pi} |
| Reference Value | 1 |
| Number of Y-Values | 3 |
| Frequency Variable | None |
| Numeric X Variables | 9 |
| Categorical X Variables | 0 |
| Final Log Likelihood | -3.597.5753 1 |
| Model R^2 | 0.10902 |
| Actual Convergence | 4.899459E-07 |
| Target Convergence | 1E-06 |
| Model D.F. | 20 |
| Priors | Equal |

Figura 8. Resultado do R^2 apresentado a partir da regressão logística pelo NCSS.

3 CONCLUSÃO

Por meio do presente estudo foi possível extrair, a partir de análises multivariadas, as principais variáveis de influência das causas de projeção de escória em convertedor BOF e sugerir um modelo de previsão de projeção.

De acordo com os dados cedidos pela empresa, o convertedor 1 apresentou 12,26% de corridas com projeção e 8,46% de utilizações do botão. O convertedor 2, apresentou 12,98% de corridas com projeção e 14,01% de utilizações do botão.

As evidências apuradas no acompanhamento mostram indícios de que muitas das projeções não são detectadas pelo operador, portanto, a análise de dados sobre o assunto pode estar comprometida devido à subjetividade dos dados em relação à projeção. A análise dos dados do botão de controle de projeção mostra que sua utilização é mais frequente aos 9 minutos de sopro. O que coincide com a literatura, a exemplo de Cicutti, Valdez, *et al.*, que apresentam em seus relatos sobre ser próximo ao período de máxima descarburização e máxima altura de emulsão.

Cada fator encontrado na análise de fatores foi explicado por apenas uma variável.

As principais variáveis apuradas no presente trabalho que estão associadas a projeção a partir da análise de *cluster*, de fatores e de CPs são: Si no gusa líquido (%); peso sucata de aço comprada tipo a (t); peso de sucata de gusa lingotado (t); peso de SiC adicionado durante o sopro de O₂ (t); peso de dunito adicionado durante o sopro de O₂ (t); peso sucata de aço tipo A (t); volume de argônio entre a sublança e o final do sopro (Nm³); peso minério no início do sopro (t); peso de briquete misto adicionado durante o sopro de O₂ (t).

O modelo matemático a partir de regressão múltipla proposto para a previsão de projeções durante a corrida possui intervalo de valores com incertezas. Isso se deve ao fato de que a técnica escolhida pode não ter sido a melhor, além de haver alguma imprecisão dos dados cedidos.

A regressão logística apresentou resultado ainda mais baixo com o R² de apenas 10,90%. A causa possível também pode ser devido à imprecisão e infidelidade da variável projeção, que é anotada de acordo com a interpretação do operador.

Agradecimentos

Os autores agradecem ao Centro de Pesquisa Global R&D ArcelorMittal pelo apoio à esse projeto, à planta de Tubarão por conceder os dados e o uso do software STATISTICA; a UFOP (REDEMAT) e a CAPES pelo auxílio financeiro para apresentação do trabalho neste evento.

REFERÊNCIAS

- 1 Bramming M. Avoiding Slopping in Top-Blow BOS Vessels [tese de licenciatura]. Lulea: Lulea University of Technology; 2010.
- 2 Remus, R. et al. Best Available Techniques (BAT) Reference Document for Iron and Steel Production. Seville: Joint Research Centre, 2013.
- 3 Rizzo, E. M. D. S. Processos de refino primário dos aços nos convertedores a oxigênio. São Paulo: ABM, 2006.
- 4 Fruehan, R. et al. The Making, Shaping and Treating of Steel. 10. ed. Pittsburgh: AISE Steel Foundation, 1998.
- 5 Likert, R. A Technique for the Measurement of Attitudes. Universidade de Nova York. R. S. Woodworth, Editor. Archives of Psychology, n° 140: pp. 1-55.
- 6 NCSS Statistical Software. Principal Components Analysis. [https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Principal_Components_Analysis.pdf] NCSS, 2017 [Acesso em: 10 ago. 2017]. Disponível em: <https://www.ncss.com/>.
- 7 Hair Jr, J. F. et al. Análise multivariada de dados. Porto Alegre: Bookman, 2009.
- 8 Estatcamp 4. Regressão logística. [<http://www.portalaction.com.br/analise-de-regressao/regressao-logistica>]. Portal Action, 2017. [Acesso em: 26 set. 2017]. Disponível em <http://www.portalaction.com.br>.
- 9 Cicutti C, Valdez M, Pérez T, Donayo R, Petroni J. Analysis of slag foaming during the operation of an industrial converter. Latin American Applied Research. 2001, 32:237-240.
- 10 Triola, M. F. Introdução à Estatística. 9. ed. Rio de Janeiro: LTC, 2005.