

Estimador de Fluxo em Tempo Real

Israel de Moraes Madalena¹, Francisco de Assis Boldt¹

¹Instituto Federal do Espírito Santo (IFES)
Rodovia ES-010 - Km 6,5 - Manguinhos, Serra - ES, 29173-087

israelmorais@yahoo.com.br, franciscoa@ifes.edu.br

Abstract. *Measuring instruments are used for product quality control, breakage prevention, production quantification and others. The article proposes to use a flow estimator using process variables and compare it with the information given by an integrating scale, in order to replace the instrument and be used to obtain the information.*

Resumo. *Instrumentos de medição são utilizados para o controle de qualidade do produto, prevenção de quebras, quantificação da produção e outros. O artigo propõe utilizar um estimador de fluxo utilizando variáveis do processo e compará-lo com as informações dadas por uma balança integradora, de forma a substituir o instrumento e ser utilizado para a obtenção das informações.*

1. Introdução

O minério de ferro foi até novembro de 2021 o produto mais exportado do Brasil, sendo responsável por uma receita de US\$ 42,2 bilhões, tendo como principal destino a China [FazComex 2022]. As indústrias siderúrgicas e metalúrgicas demandam ferro como matéria prima para a produção do aço, necessário para melhorias em infraestrutura e de suprimentos de bens materiais [Freitas 2012]. As balanças integradoras possuem sensor de velocidade e células de carga para medir o material em um transportador de correia em movimento. Elas são projetadas individualmente para cada transportador e a instalação mecânica é fundamental para o funcionamento. Conforme mostra a Figura 1, a vazão, em toneladas por hora (t/h), é obtida através do peso em kg/m (variável Q) fornecido pelas células de carga e a velocidade em m/s (variável V) medida pelo sensor de velocidade [brxistemas 2020]. A fórmula da vazão é dada pelo produto da carga (Q) em kg/m, da velocidade (v) em m/s e por 3,6 para a conversão em tonelada/hora ($Vazão = Q \times v \times 3,6$).

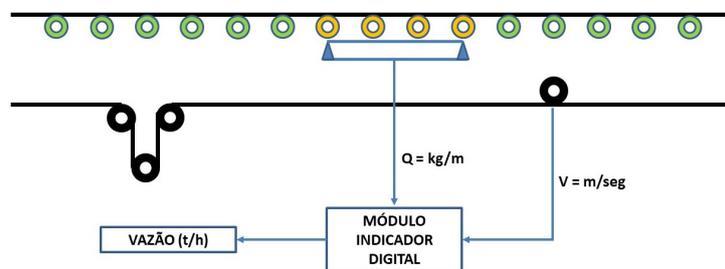


Figura 1. Esquema de uma balança integradora representada pelo sensor de velocidade e plataforma de pesagem [brxistemas 2020].

A hipótese deste trabalho é de que um medidor de fluxo pode ser obtido através da corrente elétrica dos motores e se equiparar com uma balança integradora de correia, possibilitando a obtenção de medições em tempo real.

2. Referencial Teórico

Este estudo tem o intuito de verificar através de técnicas de regressão se os modelos conseguem prever o fluxo de material em uma recuperadora de caçambas com um erro igual ou menor do que o permitido para uma balança integradora, possibilitando a utilização do modelo ao invés do instrumento. Desta forma, serão avaliados métodos de regressão que são aplicáveis a este problema e analisado o erro de cada estimador, comparando os resultados obtidos com as informações de tempo real e históricas de uma balança integradora.

As recuperadoras são utilizadas para a retirada de material das pilhas de estocagem e envio para o respectivo destino (vagões, navios, indústrias). O controle do fluxo é importante, pois estes equipamentos são projetados para uma capacidade nominal, a qual se ultrapassada, pode causar quebras e paradas de produção.

2.1. Trabalhos Correlatos

Lopes [2015] desenvolveu um modelo matemático para estimar o fluxo de recuperação, onde verificou uma alta correlação entre a corrente do motor e a vazão de recuperação, utilizando estes dados para estimar o fluxo de recuperação. Já Moreira et al. [2017] aplicaram um método de Hulthén e Evertsson [2010] que utiliza a potência consumida pelo motor do transportador para realizar a estimativa do fluxo. A partir destas afirmações, é realmente importante analisar as variáveis do processo para verificar a existência de correlações entre elas, de forma que através delas se possa fazer a previsão do fluxo real de produção. Esta solução é interessante a partir da possibilidade de eliminação da utilização de instrumentos de medição. Desta forma, a necessidade de utilizar dados históricos para treinar um modelo para que forneça os valores de fluxo em tempo real, nos leva a um problema que pode ser resolvido utilizando modelos de regressão linear. Para este trabalho utilizamos os modelos de regressão linear, descida de gradiente estocástica com reinicializações, Random Forest e o proposto por [Lopes 2015].

2.2. Estimador de fluxo com modelo linear ARX

Lopes [2012] representou matematicamente o fluxo estimado de uma recuperadora de caçambas utilizando o modelo linear ARX e o método dos mínimos quadrados estendidos para a estimação dos parâmetros [Helene 2006]. O modelo é obtido através de uma equação, onde u é a entrada do sistema (corrente dos motores), y é o fluxo estimado e k representa a variável que expressa o tempo (k é o tempo atual). A Figura 2 mostra um esquema representativo do modelo ARX, o qual temos a entrada representada por “ $u(k)$ ”, o erro representado por “ $e(k)$ ” e a saída por “ $y(k)$ ”. Com isso temos para o modelo as constantes presentes na equação $y(k) = 0,09y(k-3) - 0,76y(k-2) + 1,546y(k-1) + 12,11u(k-2) - 36,48u(k-1) + 43,238u(k)$, que foram obtidas utilizando o método dos mínimos quadrados estendidos [Lopes 2015].

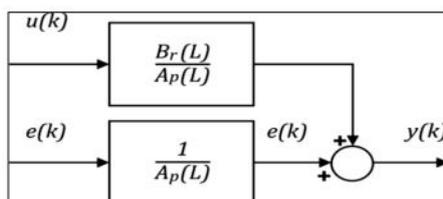


Figura 2. Representação do modelo ARX. [Lopes 2015]

2.3. Regressão Linear

A regressão linear é utilizada quando existe uma correlação linear entre uma ou mais características e a resposta [Rodrigues 2012]. A Figura 3 exibe um gráfico em função do tempo dos valores normalizados de vazão (cor azul) e a corrente dos motores 1 e 2 da roda de caçambas (cores verde e laranja), responsável por pegar o material na pilha e colocar sobre o transportador em que a balança esta localizada.

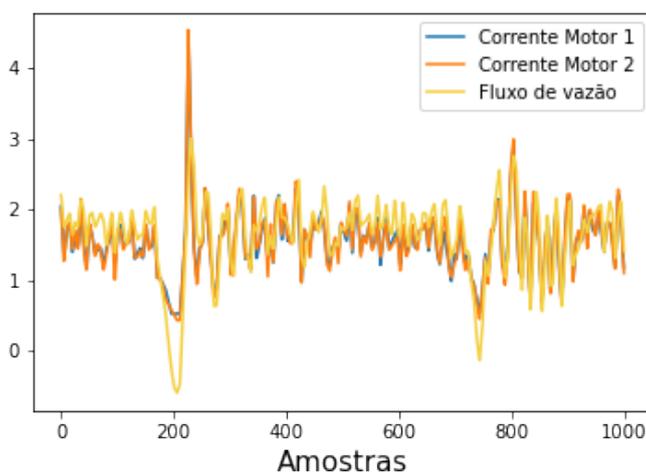


Figura 3. Valores de vazão e as corrente dos motores 1 e 2 da recuperadora RC05.

Em análise visual preliminar, observamos que uma corrente alta dos motores leva a um fluxo maior na balança e o contrário também é verdade. A partir disso, consideramos a hipótese de utilizar a regressão linear para a estimação de fluxo na comparação da variação da corrente dos motores em relação a balança [da Cunha and Carlos 2007]. Neste trabalho a regressão linear foi treinada com um algoritmo de descida de gradiente estocástica com reinicializações (SGDR). Isso faz com que a ordem das amostras de treinamento sejam randomizadas e os coeficientes atualizados após cada instância de treinamento, fazendo com que a aprendizagem seja muito mais rápida para um volume muito grande de dados [Academy 2021]. O SGDR reinicializa os pesos do algoritmo para se obter melhores resultados e evitar que fique preso em um mínimo local [Beato 2018].

2.4. Random Forest Regressor (RFR)

Random Forest é um modelo baseado em árvores de decisão, utilizando um conjunto de árvores aleatórias com o objetivo de obter o melhor desempenho, onde o resultado em cada árvore será utilizado no resultado final [El Mrabet et al. 2022]. O algoritmo busca as melhores decisões com os dados de treino e onde inseri-los dentro da estrutura, fazendo com que dada condição se siga por um ramo ou por outro no caso contrário. É um método estatístico, onde um nó é criado a partir das características de um conjunto de dados e através dos cálculos matemáticos é que as escolhas são feitas em busca do melhor resultado, alocando a característica mais adequada para o nó raiz, gerando os nós filhos e assim segue até que se tenha a quantidade de árvores desejadas [Cioccarri 2021].

3. MATERIAIS E MÉTODOS

O trabalho foi elaborado mediante a análise dos dados coletados da recuperadora de caçambas 05 (RC05) do Terminal de Minério, da empresa Vale no Porto de Tubarão - Vitória/ES. Os dados de corrente dos motores da recuperadora e da vazão da balança são historizados em uma base histórica da companhia. A Figura 3 mostra o gráfico da corrente dos motores da roda de caçambas da recuperadora. Os valores de corrente e vazão foram tratados de forma que as variáveis fiquem com a mesma ordem de grandeza, podendo assim serem comparados. Esta padronização foi feita subtraindo o valor com a média e dividindo o resultado pelo desvio padrão.

3.1. EXPERIMENTOS

A Tabela 1 mostra como os dados estão dispostos: data e horário da coleta, vazão em toneladas/hora e a corrente dos motores 1 e 2 da roda de caçamba. Com os dados coletados, há a necessidade de padronizá-los para que seja uma comparação justa e efetuar a análise de correlação. A Figura 3 mostra as informações coletadas e que foram padronizadas. Desta forma, conclui-se que os dados de vazão possuem uma correlação com as informações de corrente dos motores, podendo caracterizar uma boa referência para a estimação da vazão de acordo com a corrente dos motores, conforme sugere [Lopes 2015].

Tabela 1. Exemplos de dados de vazão e corrente nos motores da RC05.

Data	Vazão da Balança	Corrente no Motor 1	Corrente no Motor 2
2020-10-03 00:00:00	7262,27	23,23	28,47
2020-10-03 00:01:00	7088,69	26,28	34,46
2020-10-03 00:02:00	6988,88	20,18	26,37

A base de dados utilizada compreende o período entre 1º de Janeiro de 2019 e 30 de Abril de 2021. A taxa de amostragem de cada dado é de 1 segundo, gerando um total de 86.400 amostras por dia e mais de 73 milhões em todo o período. Contudo, as amostras foram utilizadas considerando a data e hora, não comparando uma amostra mais atual com uma antiga. Além disso, os dados foram separados em treino e teste, para não termos casos em que o algoritmo aprenda a prever os resultados.

3.2. ARX, SGDR e RFR

Após implementação dos algoritmos, foi realizada uma comparação da performance deles. Para isso foi utilizado uma base histórica real com as informações de corrente dos motores e da medição da vazão (fluxo) da balança. O tamanho da base de dados utilizada foi de 820 dias, contando os dados de teste e treinamento. Os dados estão em ordem cronológica e assim são executados os treinamentos e os testes. Sendo assim, um teste não é executado com dados de um período anterior as informações do treinamento. Os algoritmos SGDR e RFR são treinados com 70% da base e testados com os outros 30%. Já o algoritmo ARX não é treinado, portanto não utiliza os 70% da base que são usados para treinamento. Para os dados de treinamento dos algoritmos de aprendizado de máquina foram utilizados 595 dias, sendo o intervalo entre amostras de 1 segundo. O período dos dados de treinamento foi entre 1º de janeiro de 2019 e 18 de agosto de 2020. Já para o teste foi considerado 255 dias, de 19 de agosto de 2020 a 30 de abril de 2021.

Os coeficientes do ARX são obtidos analiticamente, como pode-se observar na equação da subseção 2.2. Com isso, não se trata de um algoritmo de aprendizado de máquina, mas é extremamente útil na aplicação prática, visto que é responsável por informar os valores de vazão em tempo real, possibilitando o ajuste da velocidade e carga da roda de caçamba, evitando sobrecargas e quebras no equipamento. O ARX não utiliza base de treinamento mas usa o histórico, as 8 primeiras amostras foram desconsideradas do cálculo geral, de forma que, a partir da 9ª amostra o resultado já considera o histórico das últimas 8 e não há nenhum coeficiente com o valor zerado. Para a execução dos testes do SGDR e do RFR também não consideramos estas 8 amostras iniciais. A Tabela 2 exibe os resultados após a execução do treinamento e teste. Para uma comparação justa, todos os algoritmos foram testados com a mesma base (30% dos dados). Percebemos que o código que executou em menor tempo foi o SGDR, seguido pelo RFR e por último o ARX. Quanto ao erro quadrático médio, o RFR obteve um melhor resultado.

Tabela 2. Comparação entre os algoritmos ARX, SGDR e RFR

Descrição	ARX	SGDR	RFR
Erro quadrático médio (em T/h)	833,84	1148,96	310,84
Tempo de execução (s)	235,49	17,43	28,52

4. Conclusão

Comparando os modelos Random Forest Regressor (RFR), Stochastic Gradient Regressor (SGDR) e o modelo ARX [Lopes 2015], verificamos que as previsões realizadas pelo ARX foram piores do que o RFR. Isso mostra que o RFR pode ser aplicado no lugar do ARX, obtendo uma melhor performance no processo. Além de obter um melhor resultado do que o ARX, o RFR é um algoritmo de aprendizado de máquina, que pode utilizar sempre das informações históricas para calibrar o modelo. Já o ARX não tem este diferencial e pode fazer com que o erro aumente com o tempo, dado que as condições do equipamento podem modificar. A Figura 4 mostra o gráfico comparativo dos 3 algoritmos durante os testes, sendo que o eixo x é o sequencial da amostra e o eixo y é o

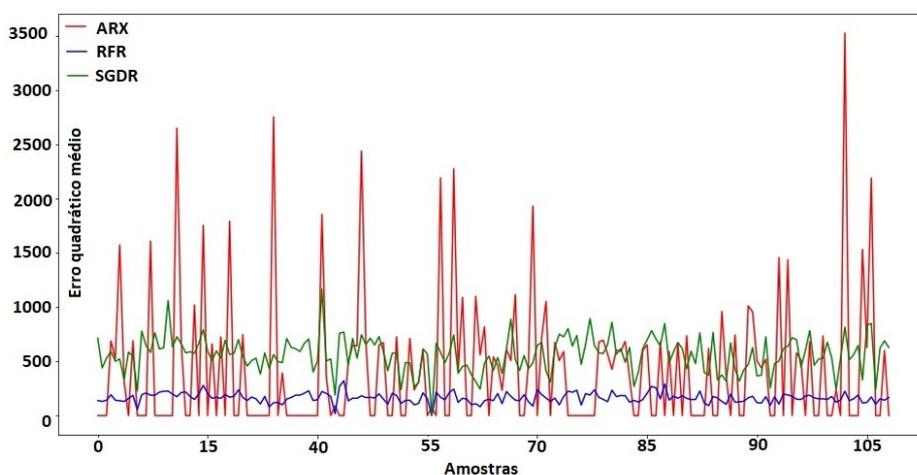


Figura 4. Gráfico do erro dos algoritmos X amostras de teste.

erro quadrático. Podemos ver que o RFR (linha azul) possui uma maior linearidade do erro durante todo o teste. O SGRD sempre com valores superiores e o ARX com valores bem variados e com vários picos durante todo o processo. A base de teste conta com mais de 22 milhões de amostras, mas a Figura 4 mostra apenas 105 para facilitar a visualização gráfica. Dessa forma, o modelo RFR pode ser utilizado no lugar do ARX, trazendo uma estabilidade maior ao processo e com possibilidade de substituir a balança integradora, fornecendo as informações de vazão do equipamento, utilizada para mensurar a produção, evitar sobrecarga, quebra e aumento da vida útil do ativo. Outra melhoria é que ao retirar a balança, reduz-se os custos com manutenção, substituição de peças e calibração.

5. Agradecimentos

Os autores agradecem ao Ifes, apoio da FAPES e CAPES (processo 2021-2S6CD, nº FAPES 132/2021) por meio do PDPG (Programa de Desenvolvimento da Pós-Graduação, Parcerias Estratégicas nos Estados)

Referências

- Academy, D. S. (2021). Deep Learning Book. <https://www.deeplearningbook.com.br/>. [Online; acessado 10-junho-2021].
- Beato, R. F. (2018). Obtenção de modelos de deep learning para a classificação automática de leucócitos. *Instituto Politécnico de Bragança*.
- brxsistemas (2020). Balança Integradora de Correia. <http://brxsistemas.com.br/produtos.php?id=13>. [Online; acessado 22-julho-2020].
- Cioccari, G. B. (2021). Integração de um algoritmo de previsão de demanda a uma linha de manufatura simulada utilizando tecnologias da indústria 4.0. *Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina*.
- da Cunha, L. S. and Carlos, A. (2007). Regressão linear múltipla. *Análise multivariada: para os cursos de administração, ciências contábeis e economia*. São Paulo: Atlas.
- El Mrabet, Z., Sugunaraj, N., Ranganathan, P., and Abhyankar, S. (2022). Random forest regressor-based approach for detecting fault location and duration in power systems. *Sensors*, 22(2):458.
- FazComex (2022). Exportação no Brasil: Quais os principais produtos exportados. <https://www.fazcomex.com.br/blog/quais-principais-produtos-exportados-brasil/>. [Online; acessado 19-maio-2022].
- Freitas, G. S. B. d. (2012). A demanda por minério de ferro no contexto das relações entre brasil e china. *Monografia (Especialização em Relações Internacionais)—Universidade de Brasília*.
- Helene, O. (2006). *Metodos dos Minimos Quadrados*. Editora Livraria da Física.
- Lopes, B. E. (2015). Bulk material reclaim control system. US Patent 9,073,701.
- Rodrigues, S. C. A. (2012). *Modelo de regressão linear e suas aplicações*. PhD thesis, Universidade da Beira Interior.