Modelagem de dados de entrada para simulação estocástica de lavra

Data modeling for stochastic mine simulation

Resumo

A simulação de lavra depende da utilização de dados coerentes e representativos da operação da mina. Esse trabalho propõe uma metodologia de modelagem de dados especificamente desenvolvida para simulação de lavra. A metodologia proposta foi aplicada a um estudo de caso de mina a céu aberto, onde dados do tempo de ciclo dos caminhões foram modelados para utilização em simulação de lavra. Os resultados da aplicação da metodologia indicaram que todas as variáveis modeladas aderiram a distribuições teóricas após o tratamento de dados, incluindo a remoção de *outliers*. O estudo identificou a necessidade do rastreamento das causas de inconsistências de dados, para o desenvolvimento de um processo de gestão de dados incorretos da operação de lavra.

Palavras-chave: Simulação de Lavra, modelagem de dados, operação de lavra, tempo de ciclo.

Abstract

Mine simulation depends on data that is both coherent and representative of the mining operation. This paper describes a methodology for modeling operational data which has been developed for mine simulation. The methodology has been applied to a case study of an open-pit mine, where the cycle times of the truck fleet have been modeled for mine simulation purposes. The results obtained have shown that once the operational data has been treated using the proposed methodology, the system variables have proven to be adherent to theoretical distributions. The research indicated the need for tracking the origin of data inconsistencies through the development of a process to manage inconsistent data from the mining operation.

Keywords: *Mine simulation, data modeling, mining operations, cycle time.*

Tatiane Marin

Mestranda da Escola Politécnica da USP, Departamento de Engenharia de Minas e de Petróleo. Pesquisadora do Laboratório de Planejamento e Otimização de Lavra (LAPOL). E-mail: tatiane.marin@gmail.com

Giorgio Francesco Cesare de Tomi

Professor Associado da Escola Politécnica da USP, Departamento de Engenharia de Minas e de Petróleo. Coordenador do Laboratório de Planejamento e Otimização de Lavra (LAPOL). Bolsista de Produtividade em Pesquisa do CNPq - Nível 2 E-mail: gdetomi@usp.br

1. Introdução

O tema desse trabalho é a modelagem dos dados de entrada para simulação estocástica da lavra. A revisão bibliográfica indicou que não existem metodologias popularizadas para a modelagem de dados da lavra. Isso ajuda a explicar porque a simulação, mesmo sendo utilizada para representar a lavra desde a década de 60 (Sturgul, 2001), ainda não se tornou uma ferramenta do dia-a-dia na tomada de decisão da lavra.

A gestão de dados de lavra é, normalmente, limitada à amostragem, ao armazenamento e à geração de relatórios, dando-se pouca ênfase à análise de consistência do banco de dados operacionais da lavra. A prática, na maior parte das empresas de mineração, é a de usar os dados do sistema de despacho para reportar dados de produção e operacionais, sem nenhum tratamento a priori.

O objetivo do trabalho é propor e aplicar um método de modelagem dos dados de lavra para se obterem as informações necessárias para um estudo de simulação estocástica. A simulação estocástica reproduz o funcionamento de um sistema que possui variáveis de entrada aleatórias. Por exemplo, na mineração, o tempo de viagem de uma frente de lavra ao britador varia aleatoriamente, desde que sejam mantidas as condições operacionais da mina.

Em modelos de simulação, o comportamento estocástico é representado por distribuições de probabilidade (empíricas ou teóricas). A modelagem de dados visa a identificar essas distribuições, de forma a representar, corretamente, a aleatoriedade do fenômeno, já que a qualidade dos resultados da simulação depende do modelo de simulação e dos dados de entrada.

A metologia proposta está descrita nas seções a seguir, assim como sua aplicação e análise dos resultados em um estudo de caso de uma mina a céu aberto de minerais industriais ao sul do Estado de São Paulo.

2. Metodologia

A metodologia proposta, nesse trabalho, foi desenvolvida com base nas pesquisas de Botter (2002), Biller e Nelson (2002), Chiwf e Medina (2006) e Vicent (1998). A modelagem de dados para simulação inclui as seguintes etapas:

- **Coleta de dados**: envolve a amostragem, armazenagem e disponibilização dos dados. É importante formalizar os procedimentos adotados na coleta de dados para que seja possível identificar quais os erros que podem ocorrer nessa etapa e, assim, atuar na consolidação dos dados.
- **Tratamento de dados**: procedimento realizado para gerar uma base de dados coerente e consistente em relação ao sistema simulado. Essa etapa permite identificar eventuais erros de coleta de dados, além de aumentar o conhecimento do sistema em estudo. Ela pode ser dividida em ordenação dos dados, avaliação descritiva, visualização dos dados, limpeza dos dados e agrupamento.
- **Inferência**: identificação da distribuição estatística dos dados processados. Essa distribuição pode ser teórica ou empírica, conforme a aderência de modelos teóricos aos dados coletados. Para fins de simulação, a situação ideal é a de que os dados possam ser associados a um modelo teórico. No entanto, se os dados não passarem pelos testes de aderência, uma distribuição empírica deve ser utilizada.

3. Estudo de caso

A metodologia proposta foi aplicada aos dados operacionais coletados como parte de um estudo de caso de simulação de lavra, com foco na obtenção das distribuições probabilísticas representativas para o estudo de simulação.

- **Coleta de dados do estudo**: Foram coletados e analisados dados equivalentes a uma semana de operação de lavra a partir do banco de dados do sistema de despacho da mina. As informações coletadas incluíram os tempos de basculamento dos caminhões, os tempos de deslocamento do britador às frentes de lavra, os tempos de carregamento dos caminhões e os tempos de deslocamento das frentes de lavra ao britador. Todos os tempos utilizados foram informados pelos operadores dos caminhões da mina através de um computador de bordo.
- **Tratamento de dados**: A metodologia de tratamento de dados para simulação foi aplicada a todos os dados coletados. Como exemplo, é apresentado a seguir o procedimento de tratamento dos dados do tempo de deslocamento da frente de lavra "FA" para o britador:

a) Ordenação dos dados: A ordenação dos dados não indicou nenhuma inconsistência, não tem valores nulos ou muito discrepantes.

b) Avaliação descritiva: Foi realizada uma avaliação descritiva completa dos dados. Os resultados obtidos foram:

Número de Registros: 299
Média: 10,21 min
Moda: 10,3833 min
Mínimo: 7,683 min
Mediana (Q_2): 10,150 min
Q_1: 9,617 min
Q_3: 10,783 min
Máximo: 15,333 min
Desvio-padrão: 0,950 min
Variância: 0,902 min²

c) Visualização dos dados: Foi criado um histograma (Figura 1a) para a visualização gráfica dos dados coletados, incluindo a distribuição de freqüências e o gráfico de caixas.

d) Limpeza dos dados: O histograma (Figura 1a) mostra que a distribuição não possui assimetria ('cauda'). A avaliação descritiva indica uma proximidade com a distribuição normal (média H", mediana H", moda), embora os testes de aderência não confirmem essa hipótese. Foram excluídos os oito valores que o gráfico de caixas da Figura 1a indicou como *outliners*, obtendo a distribuição da Figura 1b.

e) Inferência: A hipótese de que a distribuição dos dados tratados segue uma normal foi comprovada através da aplicação do gráfico de probabilidades na Figura 2. O mesmo resultado foi obtido pela aplicação do teste do Qui-Quadrado (p-value = 0,341) e do teste de Kolmogorov-Smirnov (p-value > 0.15), comprovando que os dados seguem uma distribuição normal de média 10,17 e desvio-padrão de 0,83.

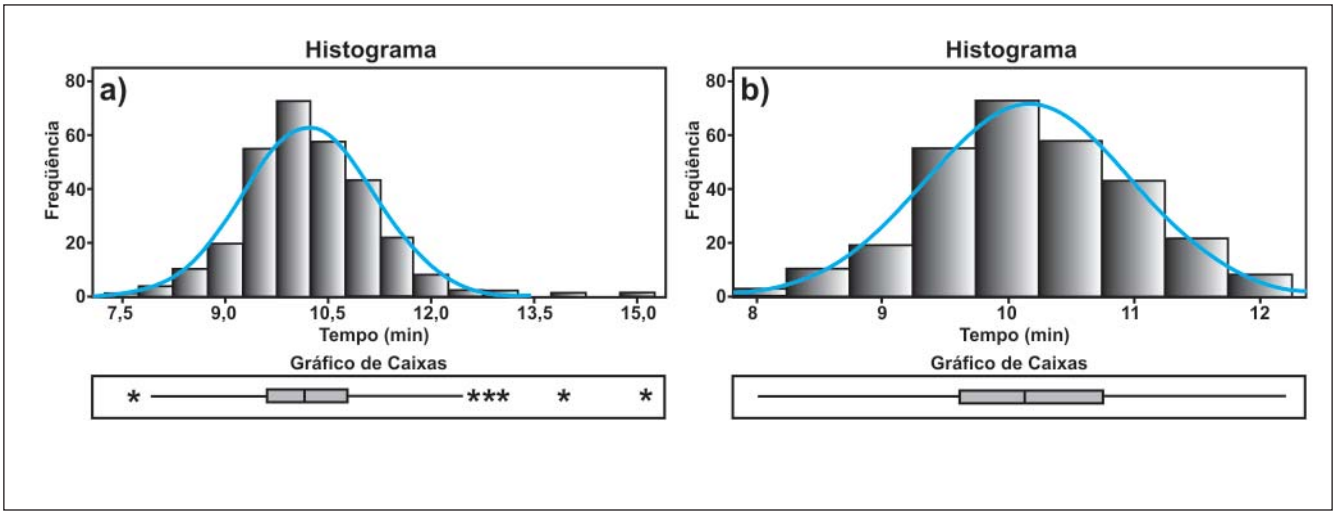

Figura 1 - **a)** Histograma dos dados originais. **b)** Histograma dos dados tratados.

4. Resultados e discussão

A metodologia de modelagem de dados aplicada aos tempos de deslocamento da frente de lavra "FA" para o britador foi também utilizada para modelar os dados das demais variáveis de interesse para o estudo de simulação, incluindo os tempos de basculamento dos caminhões, os tempos de deslocamento do britador às frentes de lavra, os tempos de carregamento dos caminhões, e os tempos de deslocamento das demais frentes de lavra ao britador.

As distribuições resultantes são as seguintes:

- Tempo de basculamento (minutos): 0.39 + ERLA(0.298, 3).
- Tempo da manobra para bascular (minutos): 0.06 + LOGN(0.167, 0.0749).
- Os tempos de deslocamento e de carregamento foram analisados para cada frente de lavra, conforme a Tabela 1.

Como já exemplificado no que se refere à modelagem dos tempos de deslocamento da frente de lavra "FA" para o britador, a etapa de limpeza de dados incluiu a identificação e a retirada de dados considerados como *outliers*. Nessa etapa, não existiu a preocupação de identificar a causa da inconsistência dos dados, que pode ser associada a erros de coleta ou mesmo a uma informação correta associada a um evento raro e factível, mas não necessariamente representativo para a simulação do sistema. Um exemplo de outliner associado a erro de coleta foi um registro do tempo de basculamento no banco de dados com valor de 0,017 minutos.

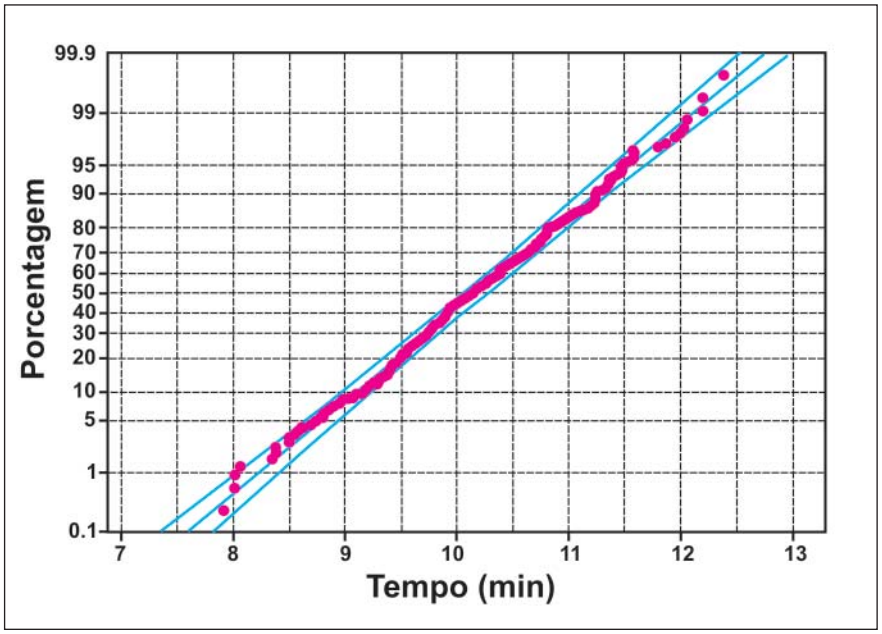

Figura 2 - Gráfico de probabilidade dos dados limpos - tempo de deslocamento da frente de lavra "FA" para o britador.

5. Conclusões

- O sucesso de um estudo de simulação depende da utilização de dados coerentes e representativos do sistema.
- A modelagem de dados tem papel essencial na simulação de lavra, mas as pesquisas em simulação de lavra não destacam essa etapa. Em muitos casos, os dados coletados são modelados diretamente através de uma distribuição teórica, sem passar por testes de aderência.

Tabela 1 - Distribuições Estatísticas por Frente de Lavra (*as fórmulas representam as distribuições probabilísticas teóricas que os dados aderiram: beta (BETA), erlang (ERLA), gama (GAMM), lognormal (LOGN), normal (NORM), triangular (TRIA) e weibull (WEIB))*.

Frente de Lavra	Deslocamento Britador - Frente de Lavra - vazio (min)	Tempo de carregamento (min)	Deslocamento Frente de Lavra - Britador - cheio (min)
FA	5 + ERLA (0.426, 6)	NORM (1.96, 0.418)	NORM (10.17, 0.83)
FB	4 + ERLA (0.429, 7)	0.64 + ERLA (0.166, 8)	6 + GAMM (0.271, 11.9)
FC	5 + GAMM (0.528, 3.33)	0.47 + 5.15 * BETA (9.27, 21)	8 + ERLA (0.346, 8)
FD	5 + GAMM (0.395, 5.21)	1 + ERLA (0.159, 6)	NORM (10.9, 0.833)
FE	4 + WEIB (2,2, 2.29)	0.39 + LOGN (1,82, 0.806)	6 + LOGN (2.6, 0.97)
FF	4 + GAMM (0.576, 5.28)	NORM (2, 0.478)	8 + GAMM (0.48, 4.44)
FG	5 + GAMM (0.48, 5.11)	0,59 + 3.41 * BETA (8.18, 11.9)	8 + GAMM (0.421, 6.49)
FH	4 + ERLA (0.509, 4)	1 + ERLA (0.215, 5)	TRIA (7, 8.32, 11.5)
FI	4 + ERLA (0.674, 4)	1 + ERLA (0.199, 5)	6.29 + ERLA (0.36, 7)
FJ	TRIA (5.27, 9.09, 12)	0.5 + 5.19 * BETA (1.46, 1.5)	TRIA (9, 10.2, 12.6)
FK	5 + WEIB (2.25, 1.6)	NORM (2.16, 0.575)	8 + WEIB (1.61, 1.75)
FL	5 + 5 * BETA (2.42, 3.43)	NORM (2.16, 0.575)	NORM (10.8, 0.919)
FM	TRIA (5.05, 6.38, 10.8)	NORM (2.33, 0.537)	8.41 + 4.41 * BETA (3.36, 3.5)
FN	4.13 + ERLA (0.354, 7)	0.76 + 2.63 * BETA (5.55, 7.81)	7 + LOGN (2.55, 1.17)
FO	5 + 6 * BETA (4.3, 7.62)	0.62 + 3.31 * BETA (7.58, 11.1)	9 + WEIB (2.39, 2.37)
FP	5.62 + 2.71 * BETA (2,18, 2.63)	1.09 + LOGN (0.824, 0.386)	NORM (9.23, 0.542)
FQ	5 + WEIB (2.95, 1.5)	NORM (2.32, 0.679)	TRIA (6, 6.6, 9)

- A metodologia proposta, nesse trabalho, formaliza os procedimentos recomendados para modelagem de dados de forma consistente e representativa para simulação de lavra.
- Essa pesquisa deve ser estendida para desenvolver procedimentos de rastreamento das causas de inconsistências de dados, permitindo a gestão de dados incorretos da operação de lavra.

6. Referências bibliográficas

BILLER, B., NELSON, B. L. Answer to the Top Ten Input Modeling Questions. In: ***Proceedings of the 2002 Winter Simulation Conference***. p. 35-40, 2002

BOTTER, R. C. ***Tratamentos de dados em modelos de simulação discreta***. São Paulo: Escola Politécnica, Universidade de São Paulo, 2002. 147p. (Tese de Livre Docência).

CHWIF, L., MEDINA, A. C. ***Modelagem e simulação de eventos discretos: teoria e aplicações***. São Paulo: Bravarte, 2006. v. 1. 255 p.

STURGUL, J.R. Modeling and simulation in mining - its time has dinally arrived, simulation. ***Transactions of The Society for Modeling and Simulation International***, 76, 286, 2001.

VICENT, S. Input Data Analysis, In: BANKS, J. (Ed.). ***Handbook of Simulation***. New York: JohnWiley and Sons, 1998. p.55-91.

Artigo recebido em 20/03/2009 e aprovado em 21/05/2009.

www.rem.com.br