

APLICAÇÃO DE TÉCNICAS DE *DATA ANALYTICS* PARA MELHORIA DO DESEMPENHO DE PROCESSOS*

Marcos de Oliveira Fonseca¹

Resumo

O crescente volume de dados gerados pelos sistemas computacionais tem estimulado o uso cada vez maior das técnicas de *Data Analytics*. Tais técnicas suportam a transformação dos dados em informações úteis para tomada de decisão nas diversas áreas de negócio. Na indústria em geral essa situação não é diferente, graças também ao uso cada vez maior de sistemas de automação e informática industrial. Esses sistemas coletam e disponibilizam milhares de dados do processo numa alta taxa de aquisição, o que normalmente representa um altíssimo volume de dados, muitas vezes chamado de *Big Data*. O presente trabalho mostra como essas técnicas podem ser aplicadas para maior conhecimento do comportamento e melhoria do desempenho de processos industriais, usando uma abordagem prática de inteligência de negócio que explora as características dos sistemas de automação e informação industrial.

Palavras-chave: *Data analytics*; *Big data*; Inteligência de negócio; Automação.

APPLYING DATA ANALYTICS TECHNIQUES FOR PROCESS PERFORMANCE IMPROVEMENT

Abstract

The increasing volume of data generated by computer systems has encouraged the increasing use of *Data Analytics* techniques. Those techniques support the data transformation into useful information for decision making in various business areas. In industry in general, this situation is not different, thanks to the increasing usage of automation and industrial IT. Those systems collect and provide thousands of process data in a high data acquisition rate, which usually represents a very large volume of data often called *Big Data*. This paper shows how these techniques can be applied for better understanding the behavior and improving the performance of industrial processes, using a practical approach of business intelligence that explores the characteristics of industrial IT and automation systems.

Keywords: *Data analytics*; *Big data*; *Business intelligence*; *Automation*.

¹ Engenheiro Eletricista, M.Sc, Senior Manager da Accenture Plant and Commercial Services (APCS), Belo Horizonte, MG, Brasil.

* Contribuição técnica ao 18º Seminário de Automação e TI Industrial, 23 a 26 de setembro de 2014, São Paulo, SP, Brasil.

1 INTRODUÇÃO

Sistemas de automação e TI industrial são utilizados pela maioria das empresas com o foco principal de monitoração e controle de processos já há algumas décadas, fazendo a aquisição de dados em tempo real a partir da instrumentação do processo e de outros dispositivos de aquisição de dados, sejam esses embarcados ou não. Recentemente, esses sistemas têm sido integrados entre si e com os sistemas de gestão (operação, produção, manutenção, qualidade, etc.), de forma que a tendência atual é que todos os dados coletados no chão de fábrica sejam disponibilizados de forma colaborativa para os diversos sistemas e usuários das partes operacional e corporativa das empresas. Dentre os sistemas que coletam e armazenam dados do processo podemos destacar o PIMS (*Plant Information Management Systems*), [1]. Nas modernas arquiteturas de automação, esse sistema normalmente desempenha o papel de elemento unificador, concentrador e base de dados de tempo real para integração dos dados de processo com os demais sistemas superiores, sendo muitas vezes responsável pela “versão única da verdade” quando nos referimos aos dados de processo. Se considerarmos uma aplicação típica para uma planta industrial de médio porte, onde um sistema PIMS faz a coleta e armazenamento de 15.000 tags na taxa de uma amostragem por segundo, onde cada tag com 4 bytes pode representar um dado discreto ou analógico, teríamos um volume da ordem de 20 Mbytes por dia e 7,2 GBytes por ano, somente para esse conjunto de dados. Esse número, ainda que pequeno perto das definições de *Big Data*, vide abaixo, caminha nessa direção quando começamos a integrar não somente dados de monitoração e controle de processo, mas também dados de operação, manutenção, qualidade, energia, custos, etc; além da integração de dados de múltiplas plantas dentro de uma organização e de outras fontes de dados ligadas à operação do processo e sua gestão (imagem, vídeo, textos, etc.). Um exemplo são dados de monitoração de vibração e de energia, os quais possuem taxa de aquisição na faixa de milissegundos. Além disso, a Internet das Coisas (*IoT*) está chegando no ambiente industrial e contribuirá ainda mais para o aumento do volume dados. Ou seja, nas plantas industriais, os conjuntos de dados ligados ao processo e sua gestão podem chegar facilmente a volumes muito altos, sendo coletados em altas velocidades e com uma grande variedade de tipos.

“Definição básica de Big Data: Coleção de conjuntos de dados tão grandes e complexos que tornam sua manipulação e processamento em tempo hábil difíceis de serem realizadas com as técnicas e ferramentas de dados tradicionais. Atualmente, volumes de dados enquadrados nessa classe geram volumes da ordem de Petabytes e Exabytes. Big Data também pode ser definido em função de três dimensões: volume, velocidade e variedade dos tipos de dados. Big Data é uma das grandes tendências de TI da atualidade”

Dentro do contexto apresentado acima, se quisermos entender o comportamento do processo e sua operação na busca de oportunidades de melhoria e maiores resultados para a empresa precisamos fazer uso de ferramentas e técnicas mais adequadas tanto para geração, coleta, armazenamento, distribuição e gestão dos dados, assim como para sua análise e geração de informação e conhecimento que possam resultar em ações efetivas para alcançar os objetivos do negócio.

Um aspecto importante nessa história toda está na Governança dos Dados, o que na prática significa basicamente quais processos e pessoas garantirão que os dados serão realmente a “versão única da verdade”. Ou seja, se não se estabelecer um modelo de governança efetivo para garantir que os dados gerados e armazenados sejam confiáveis, que análise confiável decorrente dos mesmos poderá direcionar

* Contribuição técnica ao 18º Seminário de Automação e TI Industrial, 23 a 26 de setembro de 2014, São Paulo, SP, Brasil.

decisões impactantes no negócio das empresas? Uma expressão que reflete bem isso diz, em Inglês, “*Garbage In, Garbage Out*”, ou seja, poderíamos traduzir no contexto desse trabalho como “Dado Inútil Entra, Informação Inútil Sai” como resultado da análise. Independentemente das técnicas e das ferramentas por melhores que sejam e ainda que possam ser utilizadas técnicas e abordagens para aumentar a confiança dos dados, na grande maioria das vezes um Modelo de Governança de Dados efetivo trará maiores benefícios e menores custos. Um exemplo prático consiste na gestão da calibração de um instrumento de pH como fonte de dados para uma análise. Se quisermos analisar a correlação desse dado com a recuperação metalúrgica de uma coluna de flotação em um processo mineral, que garantia teremos para essa análise tendo em vista que precisamos tanto da acurácia do dado quanto da sua precisão, dois elementos importantes controlados pela gestão da calibração do mesmo. Em linhas gerais, o modelo de Governança de Dados no âmbito industrial envolve muitas fontes de dados e equipes envolvidas. Associado ao Modelo de Governança dos Dados devemos também trabalhar a mudança cultural da organização para que as pessoas envolvidas desenvolvam a mentalidade de análise de dados dentro de um senso crítico, de forma a se somar às tecnologias e técnicas necessárias para obtenção de resultados baseada em fatos e não somente sentimento. Essa mudança de mentalidade não é um processo natural e faz-se necessária uma boa Gestão de Mudanças para que a organização possa adquirir a maturidade necessária para alcançar seus objetivos nessa área.

No que tange as tecnologias para armazenamento, processamento e visualização de dados e suas análises, podemos contar hoje com diversas soluções já disponíveis há algum tempo. O crescente aumento do poder computacional tem seguido até hoje a lei de Moore, que prevê que o número de transistores numa pastilha de silício dobra a cada dois anos. Ainda que estejamos chegando próximo ao limite do silício no que diz respeito à velocidade de processamento, outras soluções mais audaciosas estão se tornando realidade, como é o caso do processamento de alto desempenho paralelo em nuvem e também do computador quântico. Na parte de memória e unidades de armazenamento, o custo por MByte tem caído vertiginosamente, assim como aumentada a capacidade e velocidades de acesso aos dados. No campo das tecnologias para tratamento de altos volumes de dados, hoje podemos contar com o processamento em memória, que traz velocidades bem rápidas para acessar grandes volumes de dados. Além disso, existem os modelos para armazenamento e recuperação rápida de dados baseados em ROLAP, MOLAP, modelos híbridos, banco de dados orientado a colunas (colunar) e inclusive o Hadoop. Hadoop é o nome dado a uma nova infraestrutura para processamento distribuído de dados em batelada e em alta escala apoiado em vários computadores comuns ligados em rede, normalmente usando a nuvem, para permitir o tratamento de forma confiável de altíssimas quantidades de dados (múltiplos terabytes) [2]. Por ser o Hadoop uma solução de código aberto, muitas plataformas estão surgindo apoiadas nesta tecnologia, sendo liderada atualmente pelos gigantes da Internet. Existem hoje também desenvolvimentos para processamento de altos volumes de dados em memória usando GPUs (*Graphics Processing Unit*), devido ao crescente aumento do poder de processamento paralelo das placas gráficas, que utilizam milhares de núcleos e elevadas capacidades de memória de alta velocidade. As ferramentas de análise atuais permitem uma alta interatividade com a manipulação e visualização dos dados, facilitando em muito a sua utilização e a realização de análises sob diversos aspectos, o que é importante

* Contribuição técnica ao 18º Seminário de Automação e TI Industrial, 23 a 26 de setembro de 2014, São Paulo, SP, Brasil.

para confirmar e confrontar os entendimentos de forma a tornar o resultado da análise o mais robusto possível.

Somando-se às tecnologias disponíveis, temos as diversas técnicas de *Data Analytics*, que vão desde soluções de visualização de dados mais adequadas, passando também por Métodos Estatísticos, *Data Mining*, *Text Mining*, Inteligência Computacional e algoritmos matemáticos avançados, normalmente chamados de *Advanced Analytics*, [3]. Dependendo do propósito da análise, assim como o tipo, tamanho e estruturação dos conjuntos de dados, múltiplas técnicas podem ser empregadas de forma independente ou combinadas entre si. As técnicas mais comuns normalmente figuram como técnicas de BI (*Business Intelligence*), as quais encontram-se hoje disponíveis em muitas ferramentas de mercado, como o PowerBI para MS Office, Roambi, chegando a soluções mais completas como nos pacotes Estatística, R, SAS, dentre outros. Na Figura 1 (b) são apresentadas algumas das técnicas mais utilizadas atualmente.

A utilização adequada dessas técnicas permite que sejam feitas análises sobre os dados buscando um maior entendimento sobre o comportamento do processo produtivo a partir do ocorrido no passado (*hindsight*), como reagir e tomar decisões com o observado no presente (*insight*), assim como antecipar e prever o que acontecerá no futuro (*foresight*). Através dessas análises buscam-se respostas para questões de interesse do negócio, Figura 1 (b).

A utilização de técnicas de análise de dados já é consagrada no mercado, com resultados expressivos para diversas organizações em praticamente todas as áreas de negócio. A grande aceitação por parte das organizações reforça que esse é um caminho sem volta, [4]. Pesquisas mostram que organizações que investem fortemente em *Advanced Analytics* conseguem aumentar o desempenho do índice S&P 500 em 64% em média. Sua utilização na área industrial ligada ao desempenho do processo produtivo ainda é tímida, com foco em pontos mais específicos que carecem de uma abordagem mais holística, mas diversas iniciativas já mostram seu valor.

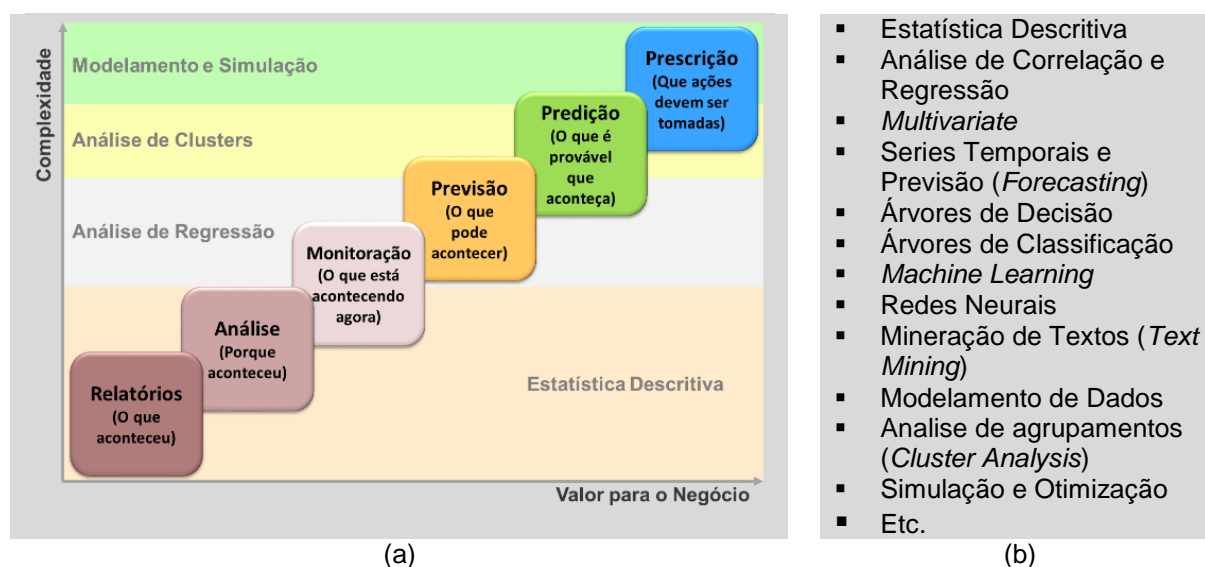


Figura 1 – (a) Questões de interesse do negócio dentro da abordagem de *Data Analytics*. (b) Exemplos das técnicas mais usadas

O presente trabalho mostra como essas técnicas podem ser aplicadas em sistemas de automação e informação industrial para melhoria do desempenho do processo produtivo e exemplifica resultados práticos obtidos.

1.1 Modelamento e Estruturação dos Dados para Aplicações Industriais

A utilização de técnicas de *Data Analytics* pode passar pelo tratamento de dados estruturados ou não-estruturados. Para o caso de processos industriais, uma abordagem bastante natural consiste na utilização de dados estruturados, uma vez que os sistemas de automação e informação fazem a geração e aquisição de dados de uma forma sistêmica a partir de fontes bem conhecidas (instrumentos, equipamentos, processos, etc.). Entretanto, tradicionalmente a maioria dos sistemas de automação foram desenvolvidos e implementados sem essa preocupação, apesar de toda a tecnologia e técnicas disponíveis, como a utilização de normas para sistemas de controle e a programação orientada a objetos [5]. A vantagem da estruturação dos dados está numa maior facilidade para habilitação de análises mais simples com ferramentas e técnicas menos sofisticadas, uma vez que o tratamento de dados não-estruturados implica normalmente em algoritmos e técnicas mais complexas.

Para a estruturação dos dados uma prática atual consiste na utilização do PIMS como principal base de dados, devido principalmente ao seu papel unificador e concentrador de dados de tempo real de diversas fontes. Neste contexto, além de promover o armazenamento de um grande volume de dados, o PIMS dispõe atualmente de tecnologias que permitem o devido modelamento dos dados de uma forma bem estruturada, ainda que a fonte envie os dados de forma não estruturada. Dessa forma é possível agregar maior contexto e significado aos dados, o que simplifica em muito a sua transformação em informação para suportar as etapas de análise [6], Figura 2.

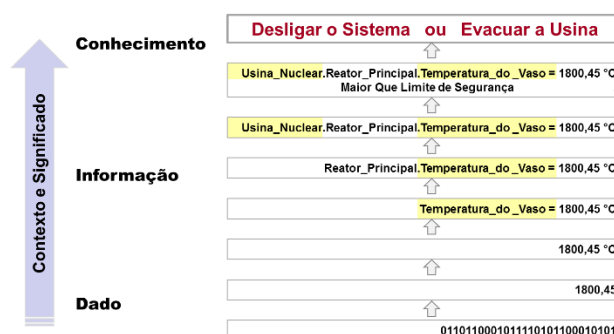


Figura 2 – Transformando Dado em Conhecimento

A partir da estruturação de dados no PIMS, podemos criar um modelo de informação que servirá tanto para organizar os dados e facilitar sua utilização no PIMS, assim como consolidar e padronizar indicadores e métricas que poderão ser utilizadas como referência para comparações e monitoração do desempenho do processo. Uma consolidação típica consiste na agregação por camadas seguindo as particularidades do processo produtivo alinhados ao contexto do negócio. Dessa forma, a análise será mais intuitiva e robusta no que se refere à estruturação e coleta de dados dentro do conceito de “fonte única da verdade” para os dados do processo produtivo. Nessa estrutura devemos associar o máximo dos conjuntos de dados (variáveis de processo, métricas, paradas, eventos, qualidade, etc.) do

* Contribuição técnica ao 18º Seminário de Automação e TI Industrial, 23 a 26 de setembro de 2014, São Paulo, SP, Brasil.

processo produtivo que possam colaborar para a qualidade, abrangência e completude da análise. A Figura 3 mostra um exemplo de como os dados podem ser estruturados usando o PIMS [7].

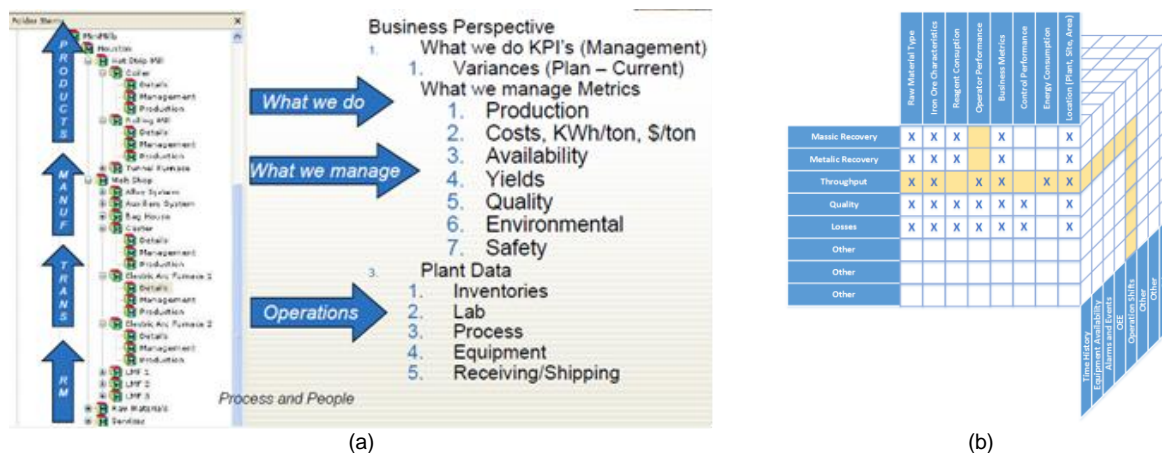


Figura 3 – (a) Exemplo de estruturação de dados no PIMS [7] (b) Relacionamento dos dados na visão de hiper-cubo

Uma vez tendo os dados devidamente estruturados no PIMS, o modelo de relacionamento de dados pode ser estabelecido no próprio PIMS ou em ferramentas externas, de forma a permitir a fácil manipulação das tabelas de fatos em relação às tabelas de dimensões. Uma forma de extrair os dados do PIMS para ferramentas de análise é mostrada na Figura 4.

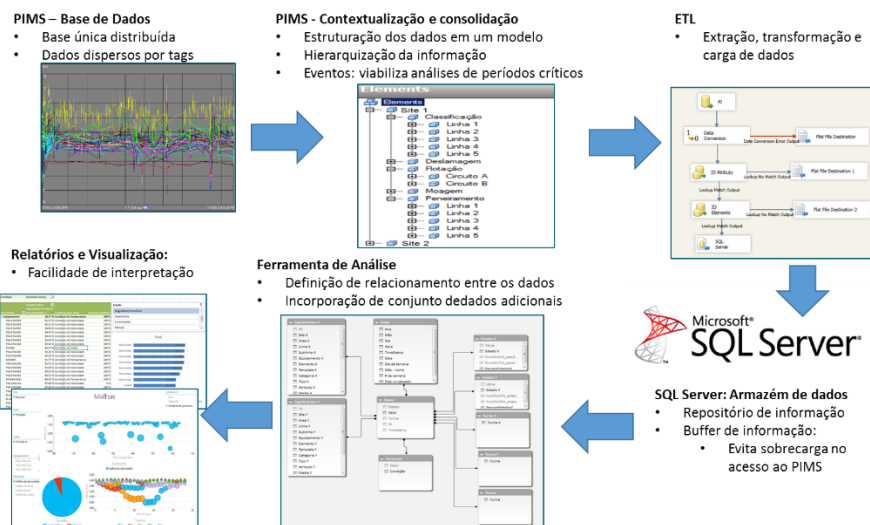


Figura 4 – Exemplo de extração de dados do PIMS para carga/acesso pelas ferramentas de análise.

Finalmente, com os dados disponíveis na ferramenta de análise é possível manipulá-los para o desenvolvimento de diversos tipos de análise usando as mais variadas técnicas. Uma abordagem interessante que pode ser aplicada em determinados casos está no conceito de autoatendimento. Nesse conceito, os usuários mais avançados conseguem, de forma relativamente mais simples, incorporar novos conjuntos de dados nos modelos, preparar esses dados, alterar os modelos de relacionamento, criar métricas e indicadores, etc. Isso elimina a necessidade de ter uma analista de TI para construção de queries elaboradas da

* Contribuição técnica ao 18º Seminário de Automação e TI Industrial, 23 a 26 de setembro de 2014, São Paulo, SP, Brasil.

forma feita tradicionalmente, agilizando o trabalho do usuário e reduzindo os custos. As manipulações clássicas de BI que a maioria das ferramentas dispõem são:

- *Slice* (fatiar): seleciona um subconjunto retangular pela escolha de um único valor para uma de suas subdimensões, criando um novo cubo sem essa dimensão.
- *Dice* (cortar em cubos): consiste em subdividir o cubo de dados em agrupamentos menores para selecionar valores específicos para múltiplas dimensões.
- *Drill Down/Up* (detalhar/resumir): permite navegar entre diferentes níveis de dados desde o mais resumido (*up*) até o mais detalhado (*down*). Exemplo, estratificar o valor de um rendimento operacional desde o nível de uma usina, passado pelos níveis de área, linha, equipamento, e assim por diante.
- *Roll Up* (sintetizar/resumir); envolve a síntese dos dados ao longo de uma dimensão.

A simples estruturação e manipulação dos dados da forma descrita anteriormente associada a uma visualização gráfica usual já permite uma análise de forma mais intuitiva e rápida. Isso proporciona ao usuário comparar comportamentos, estratificar contribuições importantes, perceber desvios relevantes, agrupar e ranquear condições de interesse relativas ao desempenho do processo. Além disso, a utilização das diversas técnicas de análise vão muito além disso, contribuindo para confirmar percepções, quebrar paradigmas e embasar decisões baseadas em fatos sem muito “sentimentalismo” sobre a real situação mostrada nos dados. Obviamente, os dados e o processo de análise têm de ser bastante confiáveis, e as conclusões devem ser interpretadas de forma adequada. Portanto, ter uma infraestrutura eficiente para coleta, tratamento e disponibilização dos dados, com seu devido modelo de informações resolve parte da equação, ficando a outra parte para os usuários que devem exercitar a mentalidade de análise e o senso crítico para tirar o máximo de resultado das técnicas disponíveis. Recomenda-se que esse modelo de trabalho seja feito de forma rotineira dentro da filosofia de melhoria contínua, principalmente porque os processos produtivos estão sujeitos a degradações e perdas de desempenho com o envelhecimento de suas partes e componentes, mudanças de estratégia e condições de trabalho, variações nas matérias-primas, etc. Além disso, as possibilidades de interações e variabilidade do processo são múltiplas, assim como a dinâmica do mercado sempre força alterações nos direcionadores de negócio que afetam as metas a serem alcançadas. Ações decorrentes das análises devem ser efetivas para contribuir para a melhoria dos resultados.

Uma vantagem da utilização do PIMS nessa abordagem é que os resultados das análises podem identificar ações que podem ser incorporadas no mesmo, permitindo automatizar como essas ações serão tratadas, desde o envio de mensagens e alertas para a operação, passando pelo suporte operacional para tomada de decisões, controle estatístico, abertura de ordens de serviços, etc.

Toda a abordagem apresentada neste trabalho pode ser implementada utilizando outros sistemas e ferramentas, estando condicionada obviamente às características de cada solução adotada.

1.2 Exemplo Prático

Para exemplificar a aplicação das técnicas, serão apresentados alguns resultados de análises da utilização das mesmas em um processo mineral.

* Contribuição técnica ao 18º Seminário de Automação e TI Industrial, 23 a 26 de setembro de 2014, São Paulo, SP, Brasil.

O processo em questão é uma planta de beneficiamento de minério para a qual foram estruturados os dados do PIMS conforme mostrado na Figura 4. A estruturação abrangeu as áreas de Moagem e Flotação. O objetivo definido para a análise consiste em entender o comportamento do processo e sua operação procurando identificar oportunidades de melhorias e otimização da variabilidade da granulometria da moagem. Todos os dados disponíveis na base do PIMS relacionados ao objetivo foram incluídos no modelo para a realização das análises. Um volume de dados referente a dois meses de operação com uma consolidação de dados a cada minuto resultou em um volume de dados de aproximadamente 1 GByte. Esses dados foram carregados na ferramenta de análise para processamento em memória, resultando em um volume de cerca de 70 milhões de linhas. Para o processamento da análise foi utilizado um computador Wintel i5 com 4 GBytes de memória RAM, que apresentou performance bastante adequada para o tipo de análise considerado. Dessa forma, as manipulações típicas de BI puderam ser realizadas com resposta em poucos segundos. Nesse tipo de análise várias formas, combinações e técnicas são possíveis, sendo importante definir bem o objetivo e foco das análises que possam ser suportadas pelos dados disponíveis.

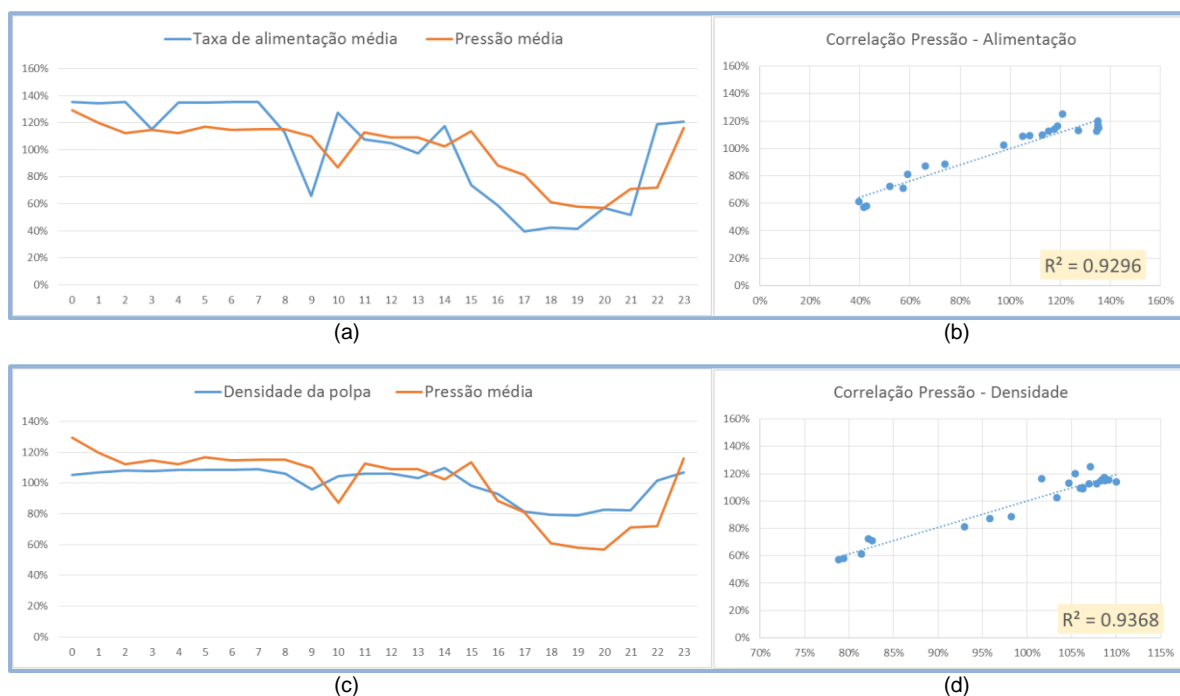


Figura 1 – Exemplo de resultados da análise do comportamento de uma planta mineral.

Conforme mostrado na Figura 5, os conjuntos de dados foram manipulados para relacionamento entre condições operacionais, estados, eventos, indicadores, variáveis de processo, etc; para identificação de oportunidades de melhoria sobre o desempenho do processo. Para direcionar a análise faz-se necessário o correto entendimento do processo produtivo e sua operação. A análise partiu de uma visão com resumo por média mensal e feito o detalhamento até a média horária, filtrando discrepâncias (*outliers*). Neste processo foi observada uma variabilidade na pressão dos ciclones da classificação. Após fatiar esses dados para escolha de um dia onde a operação aconteceu de forma normal, com pouquíssimas interrupções, foi possível analisar o comportamento temporal da pressão seguindo a taxa de alimentação do moinho (a), levantando-se à hipótese da influência entre essas variáveis. Esta

variabilidade tem efeitos negativos no ponto de corte da separação granulométrica feita pelos ciclones, de forma que a alimentação da flotação receberá uma poupa com granulometria da parte sólida fora de faixa desejada e causando degradação tanto na capacidade de recuperação mássica quanto metalúrgica, além de consumo excessivo de reagentes, perda de produto e maior impacto ambiental. A análise de correlação ajustada devido ao tempo de atraso mostrou que a alimentação tem uma contribuição de cerca de 93% na variabilidade da pressão (b).

Para confrontar a análise realizada e buscar uma confirmação adicional, nova análise foi realizada para a correlação entre a pressão do ciclone e a densidade da polpa. Novamente foi confirmado o comportamento temporal (c) e que a densidade tem uma contribuição de 94% na variabilidade da pressão (d). Análises complementares em outros pontos do processo confirmaram a hipótese levantada. A hipótese foi discutida com os responsáveis pelo processo e constatado efetivamente que a estratégia de controle atual e problemas nos equipamentos permitem a propagação da variabilidade da alimentação para o funcionamento do ciclone com impactos relevantes na granulometria. Para validar o potencial de melhorias será necessário a realização da coleta de amostradas para análise em laboratório da granulometria da saída da moagem e estimar o impacto nas etapas de flotação. A partir do comportamento estatístico desses impactos espera-se poder quantificar com a significância adequada o ganho esperado. Em função dessa validação será montado um plano de ação para implementação de melhorias nas estratégias de controle e operação do processo para captura dos ganhos estimados.

O exemplo apresentado sintetiza o processo de trabalho para aplicação das técnicas de *Data Analytics*. O modelamento adequado e o suporte proporcionado pelo ferramental utilizado permitiram uma alta flexibilidade e agilidade do seu uso por parte da equipe de análise, promovendo uma enorme aceitação para seu uso diário na investigação dos dados do processo produtivo, objetivando a melhoria do seu desempenho na busca de resultados para a empresa.

2 CONCLUSÕES

A utilização de técnicas de *Data Analytics* é uma realidade no mercado como um todo e sua aplicação no processo produtivo ainda tem muito para evoluir. Com o crescente aumento do volume de dados gerados nos sistemas de automação e TI industrial a demanda por soluções de *Big Data* e *Advanced Analytics* tende a aumentar para proporcionar melhores resultados para as empresas. A maioria das empresas já possui diversos conjuntos de dados de processo armazenados em sistemas historiadores com destaque para o PIMS, de forma que a aplicação de técnicas clássicas de BI permite a captura de resultados importantes de forma mais rápida.

A estruturação dos dados de processo utilizando a abordagem mostrada apresenta vantagens tanto para facilitar a aplicação das técnicas de análise quanto para democratizar o entendimento dos dados pelos diversos usuários, assim como potencializar a geração de informações e conhecimentos muito úteis para melhoria do negócio. Ferramentas mais simples de análise já estão disponíveis no mercado a custos acessíveis.

A mudança de cultura dos diversos usuários dos dados e informações do processo industrial para uma mentalidade analítica deve ser promovida pela liderança para elevar a maturidade da organização como um todo. A governança dos dados e a

* Contribuição técnica ao 18º Seminário de Automação e TI Industrial, 23 a 26 de setembro de 2014, São Paulo, SP, Brasil.

gestão dos sistemas de automação e TI industrial são fatores determinantes para o sucesso das empresas.

Os benefícios esperados pela aplicação de técnicas de *Data Analytics* para melhoria do processo produtivo são vários, dentre os quais podemos destacar:

- Tomada de decisão baseada em fatos
- Melhoria do desempenho do processo e sua operação a partir do maior conhecimento e constatação efetiva dos impactos no negócio
- Aplicabilidade em diversos tipos de análise
- Potencial de ganho sobre todos os indicadores de negócio do processo produtivo
- Aumento da maturidade da organização para pensamento crítico sobre o negócio

A escolha de uma solução para atendimento às necessidades de cada empresa passa pela definição da abrangência, volume e velocidade de dados e os tipos de análise e técnicas requeridas.

A adoção das práticas de *Data Analytics* requer o devido conhecimento das técnicas e tecnologias envolvidos, de forma a se obter o máximo de cada benefício proporcionado.

REFERÊNCIAS

- 1 Carvalho FB, Torres BS, Fonseca MO, Seixas Filho C. Sistemas PIMS - - Conceituação, Usos e Benefícios, Tecnologia em Metalurgia e Materiais, volume I, núm. 4, pág. 1-5, Abril-Junho, São Paulo, 2005.
- 2 “Hadoop 1.2.1 Documentation”, disponível em <http://hadoop.apache.org/docs/r1.2.1/index.html>, visitado em 14/05/2014.
- 3 Madsen M. Advanced Analytics – an overview, Third Nature, September-9, 2009.
- 4 “TDWI Best Practices Report - Predictive Analytics for Business Advantage”, TDWI Research, First Quarter, 2014.
- 5 Fonseca MO, Seixas Filho C, Bottura Filho JA. Aplicando a Norma IEC 61131 na Automação de Processos, ISA América do Sul, 2008, 568p.
- 6 Fonseca MO, Seixas Filho C. Abordagem Moderna para Modelamento de Informação para Automação, XV Seminário de Automação e TI Industrial, ABM, São Paulo, 2011.
- 7 Bascur OA. A top down approach for fast implementation and results, OsiSoft Users Conference, 2004.