

O fenômeno da incerteza e suas implicações operacionais

Andréa Nunes Carvalho

Pesquisadora da Divisão de Avaliações e Processos Industriais do INT, atua no desenvolvimento e implantação de tecnologias de gestão da produção em diferentes segmentos da Indústria. Doutora em Engenharia de Produção, com tese premiada pela Associação Brasileira de Engenharia de Produção, é também professora convidada de cursos de pós-graduação e MBAs pela Fundação Getúlio Vargas. Atuou como pesquisadora visitante na área de otimização “data-driven” na Aalto University School of Science and Technology (Finlândia). Atualmente, coordena um grupo de pesquisa voltado ao desenvolvimento de tecnologia baseada em simulação, otimização sob incerteza e aprendizado de máquina.

Palavras-chave: Incertezas, Modelagem matemática, Otimização sob incerteza, *Machine Learning*, Sistemas de apoio à tomada de decisão.





Fenômeno comum do nosso dia a dia, a incerteza está presente em todas as áreas do conhecimento humano. Ela decorre de dados incompletos, ambíguos ou inconsistentes, mas também é originária dos próprios métodos utilizados para a coleta desses dados (1). Em muitas situações, a incerteza implica na interpretação incorreta de dados, mesmo quando são aplicados modelos estatísticos de alto nível. Na área de saúde, por exemplo, problemas de classificação de dados inconclusivos resultam em diagnósticos errados (e.g., falso positivo) e no desperdício de tempo e dinheiro além de graves consequências para os pacientes. No ambiente industrial, por sua vez, incertezas derivam de questões usuais como revisões em projetos, preços, custos e prazos, variabilidade nos tempos e indisponibilidade inesperada de recursos (2) e representam um impacto significativo na estabilidade e no desempenho das organizações, por afetarem seus níveis de serviço e o uso eficiente dos recursos produtivos (3).

Em problemas complexos que exigem inteligência – entendimento e capacidade de resolução e de adaptação ao novo –, a incerteza não deve ser ignorada. Sistemas inteligentes devem ser capazes de lidar com a incerteza, já que humanos raciocinam e decidem mesmo sem ter todas as informações necessárias (1). Na Pesquisa Operacional, modelos de programação matemática que incorporam incertezas já existem há algum tempo. Dentre as abordagens conhecidas estão a otimização estocástica e a otimização robusta. A primeira pressupõe que a distribuição de probabilidades dos fenômenos incertos é conhecida e, nesse contexto, cenários alternativos podem ser considerados simultaneamente, ponderados por essas probabilidades. A otimização robusta, por sua vez, prescinde desse conhecimento, demandando do gestor apenas o intervalo dos parâmetros incertos, ou seja, todas as possíveis realizações dos parâmetros incertos das quais se deseja proteção. Esse tipo de modelagem tem como referência a análise do pior caso, ou seja, o plano gerado é submetido à realização dos valores incertos mais desfavoráveis (um pior cenário).

Tradicionalmente, modelos de otimização sob incerteza não usam de forma extensiva dados históricos para inferir sobre as funções de distribuição de probabilidade, ou mesmo para estimar o domínio dos parâmetros incertos. Isso tem motivado o desenvolvimento de modelos de otimização orientados a dados – os *data-driven optimization models*



–, ou seja, modelos em que a definição do conjunto de dados incertos se dá em função dos dados disponíveis (4). Essa característica se torna particularmente interessante no contexto do Big Data, em que a disponibilidade de dados não é um problema. De fato, no ambiente industrial, por exemplo, grandes volumes de dados heterogêneos estão sendo gerados e armazenados e têm potencial para a produção de conhecimento e para o desenvolvimento de ferramentas analíticas para melhorar o desempenho desses sistemas de produção (5).

De mais a mais, técnicas de *Machine Learning* (Aprendizado de Máquina) têm sido aplicadas para analisar e extrair informações a partir de dados para apoiar a tomada de decisão (6). Mais especificamente, técnicas de aprendizado não supervisionado – para problemas em que se tem variáveis de entrada, mas não de saída – podem ser utilizadas para a condução de análises exploratórias de dados, úteis para a formulação de hipóteses sobre as causas e os fenômenos observados. Por outro lado, técnicas de aprendizado supervisionado – para problemas em que se tem tanto variáveis de entrada quanto de saída – podem ser adotadas para se estimar ou prever variáveis de interesse e precisar a variabilidade de parâmetros incertos a partir de uma base de dados histórica. A intenção dessas aplicações é delimitar a incerteza, descrevendo-a com mais precisão.

Dentro dessa perspectiva, a associação da otimização sob incerteza com as técnicas do *Machine Learning* possibilita o desenvolvimento de aplicações *data-driven*, capazes de gerar soluções mais robustas que suas formulações tradicionais (as não *data-driven*), porém menos conservadoras. A ideia de robustez, no caso, é uma combinação entre resiliência, que indica estabilidade, e flexibilidade, decorrente da capacidade de adaptação a eventos inesperados. No contexto de planejamento da produção, por exemplo, um plano é dito robusto quando é capaz de manter o passo ou cadência na exe-



cução das atividades planejadas, absorvendo as interrupções, sem perder a consistência (7). Por outro lado, o conservadorismo se refere à proteção mediante os fenômenos incertos (i.e., quanto mais conservadora, mais protegida e inevitavelmente mais cara a solução). Em outras palavras, com as incertezas descritas de forma mais precisa, através do *Machine Learning*, um modelo de otimização é capaz de gerar soluções mais aderentes e que representem ganhos para as organizações.

Em resumo, incertezas estão presentes no nosso cotidiano e fazem parte do processo de tomada de decisão em inúmeras áreas do conhecimento humano. É preciso incorporá-las aos modelos ou sistemas que utilizamos para decidir. A combinação de técnicas advindas da Pesquisa Operacional (e.g., otimização estocástica e otimização robusta) com o ferramental

estatístico inerente ao *Machine Learning* é um caminho promissor para a gestão das incertezas na tomada de decisão, sobretudo num momento em que, cada vez mais, há disponibilidade de dados em nossos sistemas. Pensando nisso, iniciamos no INT um projeto de pesquisa intitulado "*Planejamento da produção via otimização data-driven: combinando a programação matemática e técnicas de aprendizado de máquina*", que vem sendo desenvolvido por mim, Andréa Carvalho, e minha equipe de pesquisa, em parceria com os pesquisadores Fabricio Oliveira (Aalto University – Finlândia) e Fernando Cyrino (PUC-Rio). Recém-aprovado pela FAPERJ, esse trabalho tem como objetivo gerar uma metodologia para o desenvolvimento de modelos de otimização sob incerteza *data-driven*, que deverão apoiar a tomada de decisão no planejamento da produção. ●

Referências

1. Ogheneovo EE, Nlerum PA. Managing Uncertainty in Artificial Intelligence and Expert Systems Using Bayesian Theory and Probabilistic Reasoning. *Am J Eng Res (AJER)*. 2020;9(March):53–9.
2. Artigues, C., Leus, R., Talla Nobibon F. Robust optimization for resource-constrained project scheduling with uncertain activity durations. *Flex Serv Manuf J*. 2013;(25):175–205.
3. Tolio T, Urgo M. A Rolling Horizon Approach to Plan Outsourcing in Manufacturing-to-Order Environments Affected by Uncertainty. *CIRP Ann - Manuf Technol*. 2007;56(1):487–90.
4. Bertsimas D, Gupta V, Kallus N. Data-driven robust optimization. Vol. 167, *Mathematical Programming*. Springer Berlin Heidelberg; 2018. 235–292 p.
5. Kozjek D, Vrabič R, Rihtaršič B, Butala P. Big data analytics for operations management in engineer-to-order manufacturing. *Procedia CIRP*. 2018;72:209–14.
6. Ning C, You F. Data-driven adaptive robust unit commitment under wind power uncertainty: A Bayesian nonparametric approach. *IEEE Trans Power Syst*. 2019;34(3):2409–18.
7. Policella N, Smith SF. Generating Robust Schedules through Temporal Flexibility Nicola. In: *Proceedings of the Fourteenth International Conference on Automated Planning and Scheduling*. 2004. p. 1–9.