RECEBIDO
01/07/2016
APROVADO
31/08/2016

Gerenciamento de Indicadores de Desempenho Industrial: Análise de Regressão e Simulação

Management of Industrial Performance Indicators: Regression Analysis and Simulation

Walter Roberto Hernandez Vergara[1] - Univ. Fed. da Grande Dourados - Fac. de Engenharia - Eng. de Produção
Rafael Henrique Barros da Silva[2] - Univ. Fed. da Grande Dourados - Fac. de Engenharia - Eng. de Produção
Fabio Alves Barbosa[3] - Univ. Fed. da Grande Dourados - Fac. de Engenharia - Eng. de Produção
Juliana Suemi Yamanari[4] - Univ. de São Paulo - Escola de Engenharia de São Carlos - Eng. de Produção

RESUMO Os métodos estocásticos podem ser utilizados na solução de problemas e explicação de fenômenos naturais através de procedimentos estatísticos. O artigo busca associar a análise de regressão e simulação de sistemas com o objetivo de facilitar o entendimento prático da análise de dados. Os algoritmos foram desenvolvidos no *software Microsoft Office Excel*, utilizando técnicas estatísticas como a teoria da regressão, ANOVA e fatoração de Cholesky, possibilitando a criação de modelos de sistemas simples e múltiplos com até cinco variáveis independentes. Para a análise dos referidos modelos, utilizou-se a simulação de Monte Carlo e análises de indicadores de desempenho industrial, resultando em índices numéricos que tem como objetivo melhorar o gerenciamento das metas para indicadores de conformidade, através da identificação de instabilidade, correlação e anomalias do sistema. Os modelos de análise apresentados na pesquisa representaram resultados satisfatórios com inúmeras possibilidades de aplicações empresariais e acadêmicas, além do potencial para desdobramento em novas técnicas de análise.

Palavras-chave: Análise de Regressão. Simulação de Sistemas. Indicadores de Desempenho.

ABSTRACT *Stochastic methods can be used in problem solving and explanation of natural phenomena through the application of statistical procedures. The article aims to associate the regression analysis and systems simulation, in order to facilitate the practical understanding of data analysis. The algorithms were developed in Microsoft Office Excel software, using statistical techniques such as regression theory, ANOVA and Cholesky Factorization, which made it possible to create models of single and multiple systems with up to five independent variables. For the analysis of these models, the Monte Carlo simulation and analysis of industrial performance indicators were used, resulting in numerical indices that aim to improve the goals' management for compliance indicators, by identifying systems' instability, correlation and anomalies. The analytical models presented in the survey indicated satisfactory results with numerous possibilities for industrial and academic applications, as well as the potential for deployment in new analytical techniques.*

Keywords: *Regression analysis. Systems Simulation. Performance Indicators.*

1.Rua Quintino Bocaiuva, 817, Apto. 06, Centro, Dourados, Minas Gerais, CEP - 79803-030, waltervergara@ufgd.edu.br; 2. rafaelhenrique.bs@gmail.com; 3. fabiobarbosa@ufgd.edu.br; 4. jusuemi@hotmail.com

VERGARA, W. R. H.; SILVA, R. H.B.; BARBOSA, F. A.; YAMANARI, J. S. Gerenciamento de Indicadores de Desempenho Industrial: Análise de Regressão e Simulação. **GEPROS. Gestão da Produção, Operações e Sistemas**, Bauru, Ano 12, nº 4, out-dez/2017, p. 183-203.

DOI: 10.15675/gepros.v12i3.1668

1. INTRODUÇÃO

A complexidade dos problemas da ciência e dos negócios motiva os pesquisadores a buscarem o entendimento sobre o grau da relação entre as variáveis de um sistema, seja para definir a causa de um problema ou até estimar seu comportamento futuro. Comumente, questiona-se a influência que um fator exerce sobre os outros que compõem um sistema.

Os modelos matemáticos são utilizados para representar fenômenos da natureza que consistem de uma série de variáveis inseridas em equações. A análise de regressão é um método estatístico utilizado para modelar a relação entre uma variável dependente e uma ou mais variáveis independentes. Ela é muito utilizada e difundida nos negócios modernos e sua modelagem estocástica inicia com a coleta, síntese e análise de dados para estabelecer as relações sobre o mundo real. A análise de variância (ANOVA) foi utilizada para testar se as hipóteses (nula e alternativa) das médias de duas ou mais populações são iguais ou não e, a fatoração de Cholesky como ferramenta na decomposição da matriz simétrica formada na análise da regressão em uma matriz triangular inferior e sua transposta.

O método de simulação de Monte Carlos é uma ferramenta estatística de amostragem computacional muito útil para alcançar soluções numéricas bem aproximadas para problemas de sistemas complexos e parâmetros incertos. Por fim, o objetivo da pesquisa consiste em formular um modelo estatístico baseado na análise de regressão e realizar a simulação para estimar indicadores de desempenho a fim de contribuir com a eficiência dos mesmos.

2. REFERENCIAL TEÓRICO

2.1. Indicadores de desempenho

O desempenho de uma organização pode ser entendido como a quantificação de metas concretas mensuráveis e específicas derivadas de uma missão, definido e focado na otimização de valor do consumidor, usando dados de desempenho para melhorar continuamente as operações do sistema (NASA, 2015).

Um indicador de desempenho é uma ferramenta de gestão do desempenho empresarial e seu valor quantitativo possibilita medir o que está sendo executado para gerenciar os negócios de forma adequada, buscando atingir as metas organizacionais. Ele contém informações sobre as características e resultados de um produto, serviço ou processo. As informações de processos, por exemplo, permitem aos gestores traçar planos de ação para atingir metas organizacionais e até mesmo determinar se estão ou não no caminho certo. Um fator importante dos indicadores é que quando são disponibilizados e coerentes podem ser comparados com outros setores específicos da empresa ou outro segmento análogo do mercado.

Muitas vezes os processos não são mensurados e controlados e, consequentemente, mal-entendidos. Um dos desafios na construção de um indicador de desempenho é a concepção da meta que se deseja alcançar. O gestor precisa saber exatamente aonde quer chegar e o que é relevante no processo em estudo. Entender como um processo funciona significa conhecer os fatores que afetam seu desempenho, como o sistema se comporta diante de intervenções e os limites da capabilidade do mesmo. Quando gestores não recebem um *feedback* adequado, oportunidades não são reconhecidas e os processos se tornam mais vulneráveis às ameaças.

2.2. Análise de regressão

As ferramentas estatísticas frequentemente são aplicadas para quantificar e manipular os componentes de um sistema, sendo a análise de regressão uma das mais influentes e amplamente utilizada em diversos campos de estudo (RISPOLI ; SHAH, 2015). A análise de regressão consiste em relacionar uma variável dependente, também chamada de resposta, e uma ou mais variáveis independentes, também conhecidas como preditoras.

Depois de definir se o modelo utilizado na regressão é linear ou não linear, simples ou múltiplo, os parâmetros de regressão devem ser estimados. Vale ressaltar que antes de utilizar o modelo é preciso validá-lo, verificar novamente a precisão dos dados e constatar que todas as suposições feitas sobre a natureza do sistema são corretas (BASTOS; GUIMARÃES; SEVERO, 2015).

2.2.1. Regressão linear simples

Um modelo de regressão é dito linear quando seus parâmetros são lineares e eles são denotados na forma mostrada na Equação 1.

$$y = \beta_0 + \beta_1 x + \varepsilon \qquad (1)$$

A função consiste de uma parte determinística e uma parte aleatória. A parte determinística da equação, $\beta_0 + \beta_1 x$, especifica que para qualquer valor da variável independente x, a média populacional da variável dependente y é descrita por uma função reta. De acordo com as propriedades de uma função linear, o parâmetro β_0, também chamado de coeficiente linear ou intercepção, é o valor de y quando x se iguala à zero. O parâmetro β_1, conhecido como coeficiente angular ou inclinação, é a variação no eixo y associada à variação de uma unidade no eixo x. A parte aleatória ε do modelo é utilizada para explicar a variabilidade das respostas em relação à média real e, é assumido como normalmente distribuído, com média zero e desvio padrão constante. Os parâmetros de regressão β_0 e β_1 são estimados com base nos dados coletados das variáveis estudadas para encontrar a reta que faz o melhor ajuste dos pontos no gráfico (RISPOLI ; SHAH, 2015).

2.2.2. Regressão linear múltipla

Quando uma variável-resposta se relaciona linearmente com mais de uma variável preditora, tem-se um modelo linear múltiplo. A função desse modelo toma a forma da Equação 2.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_n x_n + \varepsilon \qquad (2)$$

Enquanto a equação de regressão linear simples era dada por uma reta, essa equação se dá como uma superfície e é chamada de regressão linear múltipla. Nesse tipo de regressão, o coeficiente β_0 é o valor de y para $x_1 = x_2 = x_n = 0$. Os coeficientes β_n são coeficientes parciais de regressão porque representam a contribuição de x_n para a resposta y depois de ambas variáveis terem sido linearmente ajustada para outras variáveis independentes (RISPOLIv; SHAH, 2015).

Os parâmetros de modelos de regressão múltipla são diferentes da regressão simples, uma vez que as variáveis independentes estão correlacionadas. Hillier e Lieberman (2012) destacam que a qualidade da regressão múltipla pode ser obtida calculando o coeficiente de relação múltipla que mede a força da relação entre todas as variáveis do modelo.

2.3. Método Monte Carlo

A simulação está entre as técnicas de tomada de decisão mais utilizadas atualmente. Por ser uma ferramenta flexível e intuitiva, a popularidade da simulação continua crescente (HILLIER; LIEBERMAN, 2012). Essa técnica consiste no processo de criação do modelo de um sistema real e condução de experimentos com esse modelo com o objetivo de entender o comportamento do sistema e avaliar várias estratégias para sua operação.

O método de Monte Carlo permite analisar, avaliar e compreender um processo, a partir de um conjunto de variáveis representativas do sistema, sem perturbá-lo com uma série de restrições. Ele se aplica tanto a problemas determinísticos quanto problemas com estrutura probabilística e possui a menor taxa de decrescimento do erro absoluto dos estimadores à medida que as amostras de números aleatórios aumentam (NASA, 2013; RUI-MEI, 2015; BODEA ; PURNUŞ, 2012).

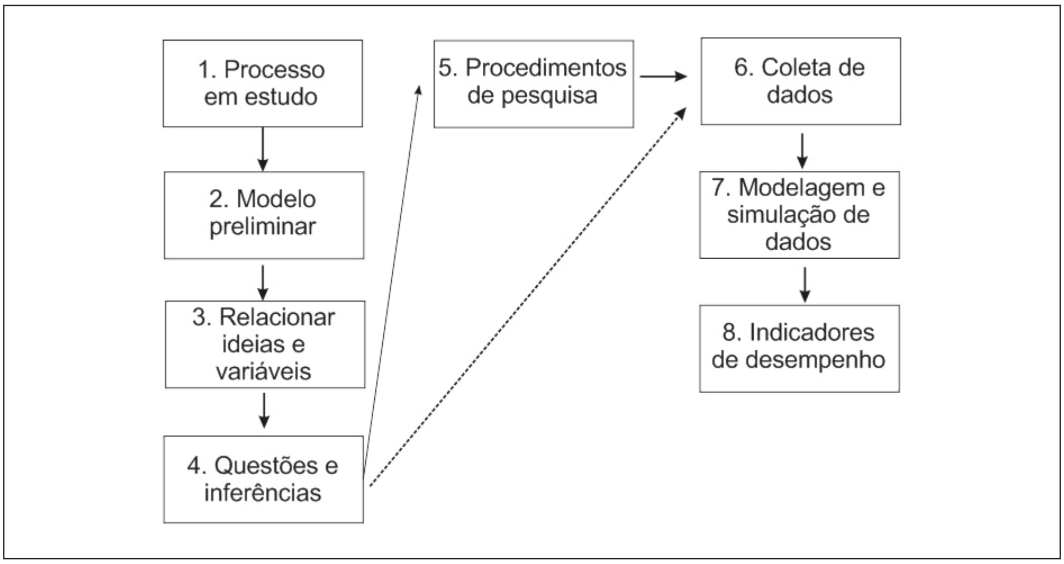
A simulação de Monte Carlo realiza a análise de risco a partir dos seguintes procedimentos: (1) identificar as variáveis de risco e configurar um modelo; (2) especificar a função de distribuição de probabilidade adequada para cada variável de risco selecionada; (3) realizar simulações, isto é, atribuir um valor aleatório para cada função de distribuição de probabilidade por um método de amostragem; (4) para cada iteração uma taxa de aproveitamento ou de benefício é calculada; e (5) repetir o processo (Passo 2-4) para uma quantidade relativamente grande de iterações. Durante a simulação, são gerados números aleatórios para cada variável de risco segundo a distribuição de probabilidade especificada (NASA, 2015).

3. MÉTODO

Do ponto de vista de sua natureza, o trabalho é classificado como como pesquisa aplicada, uma vez que busca gerar conhecimentos para a aplicação prática e solução de problemas específicos. O desenvolvimento da pesquisa é essencialmente quantitativo, visto que utiliza métodos estatísticos e matemáticos para analisar dados das variáveis estudadas (LAKATOS ; MARCONI, 2011). Os resultados numéricos das análises de regressão, análises de variâncias e simulações foram interpretados e transformados em informações para alcançar os objetivos da pesquisa.

A partir dos dados coletados e selecionados, foram desenvolvidos os modelos estatísticos no *Microsoft Office Excel*. Com o objetivo de aprimorar a pesquisa, cada indicador foi analisado em relação a outras duas variáveis através de um modelo de regressão linear múltipla. Alinhado a esse modelo, foi criado um sistema de análise de variâncias para verificar a eficiência dos resultados da regressão linear. O método da simulação de Monte Carlo foi utilizado na fatoração de Cholesky e análise de regressão linear utilizando uma grande quantidade de amostras aleatórias. Na pesquisa foram destacadas oito atividades principais que, apesar de ser apresentadas sequencialmente, não necessariamente foram realizadas nessa ordem (ver Figura 1).

Figura 1 – Metodologia de pesquisa.



Fonte: Elaborado pelos autores.

4. MODELAGEM E DESENVOLVIMENTO

4.1. Caracterização da empresa

A empresa utilizada para aplicação dos modelos é uma multinacional do segmento alimentício. Os dados foram coletados de uma unidade fabril da empresa localizada em Dourados, Mato Grosso do Sul. A unidade é um frigorífico especializado em cortes de frango pesado, produzindo uma diversa gama de produtos. A capacidade instalada de abate é de 160 mil aves por dia, funcionando em dois turnos de produção cinco dias por semana, com um quadro de funcionários de aproximadamente 1700 pessoas.

4.2. Seleção de dados

Para o desenvolvimento da análise linear simples, foram coletados resultados diários de três meses dos indicadores de rendimento e conformidade de produto. Para a análise de regressão linear múltipla, análise de variâncias e simulação, acrescentou-se ao modelo duas variáveis levantadas na fábrica como possíveis influências nos resultados de rendimento e conformidade. Essas variáveis são peso médio das aves vivas e volume total de aves abatidas por dia, das quais se realizou a mesma quantidade de observações.

Após coletados, os dados foram analisados com o objetivo de transformar as observações em amostras mais confiáveis. A primeira seleção foi a limitação da amostra a dados correspondentes aos dias entre segunda-feira a sexta-feira, desconsiderando os resultados de sábados. Essa decisão foi tomada porque abates realizados ao sábado não são ordinários e ocorrem em condições diferentes e circunstâncias especiais, o que poderia prejudicar a confiabilidade dos resultados das análises. Posteriormente, os dados foram submetidos a um gráfico de controle para visualização dos *outliers*. Os pontos que se distanciaram da média amostral mais de três vezes o desvio padrão foram considerados pontos anômalos e eliminados da amostra. Após seleção, as amostras apresentavam 45 observações.

4.3. Indicador de rendimento

O rendimento de família (R_f) avalia a taxa de aproveitamento de uma família de produto, calculando a proporção de volume real de produto acabado (V_r) sobre o volume ideal de produto acabado (V_i). Para cada parte do frango, há um rendimento médio esperado (R_p) previamente estabelecido por veterinários baseado em variações no peso, linhagem e outras características da ave. Esse rendimento médio das partes representa a proporção do peso de cada parte, por exemplo, sobre o peso total do frango vivo. Logo, obtém-se o volume ideal de produto acabado multiplicando o volume aproveitado de matéria prima pelo rendimento médio da parte correspondente a família. O volume aproveitado de matéria prima é determinado pela quantidade total recebida (V_i) descontado do volume de aves mortas (V_{mort}) e volume de aves condenadas (V_{cond}). É possível sintetizar o cálculo do rendimento de família conforme apresentado na Equação 3.

$$R_f = \frac{V_r}{V_i} = \frac{V_r}{(V_t-(V_{mort}+V_{cond}))*R_p} \tag{3}$$

Ilusoriamente simples, o rendimento geral calculado pela empresa não é apenas a expansão dos indicadores de rendimento de família para o volume total de produção. No cálculo do rendimento geral, é atribuído um peso diferente para cada *stock keeping unit* (SKU). Esses pesos são chamados de índices unitários e são atribuídos a cada item com base na composição do produto, rendimento médio das partes, e até mesmo, valor de mercado. O volume de produção individual de cada produto (V_j) multiplicado pelo seu respectivo índice unitário (I_u) retornará a quantidade total de índices. O resultado de rendimento geral (R_g) se dá pela proporção do total de índices em relação ao volume total recebido de matéria prima (V_t). Vale ressaltar que nesse indicador, não se utiliza o volume aproveitado de matéria prima como no rendimento de família. Assim, as taxas de mortalidade de aves durante o transporte e condenação de aves pelo serviço de inspeção afetam diretamente no valor do rendimento geral. O cálculo de rendimento geral se dá conforme estabelecido na Equação 4.

$$R_g = \frac{\Sigma\left(V_j * I_u\right)}{V_t} \tag{4}$$

O rendimento é calculado diariamente referente à produção do dia anterior e referente à produção acumulada no mês até o dia anterior.

4.4. Indicador de conformidade

O Indicador de conformidade se refere à qualidade do produto, avaliando o atendimento dos requisitos exigidos pelos clientes. O resultado de conformidade é uma nota composta de cinco notas parciais, referentes a cada um dos grupos de atributos avaliados no indicador. Para todos os produtos, os grupos são embalagem primária; ossos, cartilagens e objetos estranhos; padrão de refile; peso; e outros defeitos. Porém, os atributos que compõem cada grupo variam conforme as características de cada produto. Na composição do indicador, cada grupo de atributos tem um peso diferente, peso que também varia para o mesmo grupo de um produto diferente.

As análises são realizadas por amostragem cinco vezes ao dia. No momento da análise, um monitor retira do processo uma amostra de tamanho igual ao especificado para aquele produto. Todas as peças da amostra selecionada são avaliadas individualmente para todos os atributos previstos no indicador de conformidade. Cada atributo recebe uma nota de acordo com a quantidade de peças encontradas com seu respectivo defeito. As notas vão de 0 a 3, sendo 0 o melhor resultado e 3 o pior. A quantidade limite de peças defeituosas para o enquadramento em cada faixa é claramente estabelecida para todos os atributos e definida conforme exigência do mercado consumidor. A partir das notas dos atributos, é determinada a nota de cada grupo e a nota final do indicador de conformidade. O indicador diário é obtido pela média simples dos resultados das cinco análises.

4.5. Modelo de regressão

A partir dos dados coletados foi criado um modelo interativo de análise de regressão linear simples e múltipla. O sistema permite ao usuário calcular os coeficientes de regressão linear para qualquer conjunto de amostras de uma variável dependente e até cinco variáveis independentes.

Depois de inserir os dados são exibidos os valores dos índices de correlação (R) e determinação (R^2), coeficientes de regressão linear (β_i), e a tabela de análise de variâncias contendo soma dos quadrados, graus de liberdade e resultado do teste de distribuição de probabilidade F.

4.6. Modelo de simulação

O método de Monte Carlo foi aplicado na análise de regressão linear. O algoritmo simula um conjunto de amostras de milhares de observações que possuam as mesmas características do conjunto original. As características desejadas nas amostras simuladas são média, desvio padrão, coeficiente de correlação e coeficientes de regressão.

Dois modelos diferentes foram desenvolvidos para simular a regressão linear simples e múltipla. Na simulação de regressão simples, foi calculado o coeficiente a com base no coeficiente de correlação R, dado pela equação 5.

$$a = \frac{R}{\sqrt{1-R^2}} \tag{5}$$

O coeficiente a vai determinar que a correlação das amostras simuladas aleatoriamente seja equivalente à correlação real. Duas amostras x e y de dez mil números aleatórios foram criadas utilizando a fórmula inversa da distribuição normal com média igual a zero e desvio padrão igual a um. Uma terceira amostra z é criada a partir das duas amostras aleatórias e do coeficiente a, para servir de variável pivô da simulação. Finalmente, duas amostras são simuladas a partir das médias e desvios padrão reais associados às variáveis x e z.

Os resultados dos coeficientes reais e simulados e o erro residual da simulação são apresentados em uma tabela de sumário. Outro resultado importante da análise é o valor chamado no modelo de meta, que representa par de pontos com o melhor resultado mútuo para ambas variáveis. Essa meta é encontrada pelo produto máximo de cada par de observações das amostras simuladas.

Para simular um grupo de mais de três amostras, a fatoração de Cholesky fez o papel do coeficiente utilizado anteriormente. Comumente utilizada na solução matemática de problemas lineares, a fatoração de Cholesky é a decomposição de uma matriz simétrica e definida positiva no produto da matriz triangular inferior e sua transposta Hermitiana. A decomposição de Cholesky se dá na forma da Equação 6, onde é uma matriz triangular inferior com diagonal real e positiva e L^T é a transposta hermitiana de L. (KRISHNAMOORTHY; MENON, 2014).

$$A = LL^T \tag{6}$$

A matriz de Cholesky é usualmente aplicada no método de Monte Carlo para simular sistemas com múltiplas variáveis correlacionadas. A matriz L, obtida pela decomposição da matriz de correlação do sistema, é aplicada a um vetor u de amostras aleatórias não correlacionadas. O vetor produzido LU possui as propriedades de correlação e covariação do sistema sendo simulado. Executada a simulação, é possível observar os valores máximos e mínimos alcançados pelos indicadores com base nas variáveis preditoras, assim como prever o valor da variável de resposta.

5. RESULTADOS E DISCUSSÃO

5.1. Análise simples de indicadores

Os indicadores de Rendimento e Conformidade do produto "perna desossada" foram analisados a fim de testar uma tendência inversamente proporcional no seu relacionamento. Os dados foram inseridos em ordem cronológica, onde representa o indicador de rendimento de família e x1 o indicador de conformidade do produto. As 45 observações das amostras estão apresentadas na Figura 2.

Figura 2 - Dados coletados dos indicadores de rendimento e conformidade.

Y	X1	Y	X1	Y	X1	Y	X1	Y	X1
97,63	74,58	96,67	75,17	96,60	77,68	91,04	81,65	93,87	75,41
97,07	75,53	95,82	71,27	98,03	79,19	91,35	81,34	95,25	81,39
97,76	72,67	94,92	76,00	94,63	72,80	98,62	74,73	98,57	77,71
98,23	75,39	96,38	74,73	96,03	77,11	92,04	81,15	99,78	78,71
94,06	74,00	98,84	72,65	99,78	73,00	95,17	83,82	97,61	73,58
97,64	74,40	97,69	77,79	98,34	74,98	92,33	82,58	94,52	73,52
98,96	73,97	94,73	78,17	90,35	82,75	92,37	87,22	93,11	80,50
94,38	74,65	98,25	75,64	94,52	73,52	90,05	84,89	90,02	83,92
94,18	75,31	99,98	68,00	90,01	83,03	97,89	75,38	99,82	75,47

Fonte: Elaborado pelos autores.

Após a inserção das amostras, o modelo apresenta os resultados em tabelas, como pode ser visualizado na Figura 3, que mostra a interface do modelo com os resultados da análise.

Figura 3 - Resultados do modelo de análise de regressão simples.

Insira um número de variáveis dependentes entre 1 e 5:	1
Insira os valores das variáveis no quadro ao lado	

R	R^2	β0	β1
0,7872	0,6196	79,804	0,2051

Fonte	*Soma dos quadrados*	*Graus de liberdade*	*Soma dos q. média*	*F calculado*
Devido à regressão	1076,89	1	1076,89	79,82
Devido ao erro	661,12	49	13,49	
Total	1738,01	50	F tabelado	4,04

Y	X1
97,63	74,58
97,07	75,53
97,76	72,67
98,23	75,39
94,06	74,00
97,64	74,40
98,96	73,97
94,38	74,65
94,18	75,31
96,67	75,17
95,82	71,27
94,92	76,00
96,38	74,73

Fonte: Elaborado pelos autores.

O valor absoluto do coeficiente de correlação R de 0,7872 é bastante expressivo e indica uma forte dependência linear entre as variáveis. Por ser positivo, o resultado indica que quando um indicador aumenta, o outro tende a diminuir. A equação da reta de regressão linear é dada na Equação 7.

$$\text{Rendimento} = 79{,}804 - 0{,}2051 * \text{Conformidade} \qquad (7)$$

De acordo com a Equação 7, o coeficiente β_1 é negativo, indicando mais uma vez a relação inversa entre as variáveis, e baixo, indicando que ambas variáveis variam em escalas aproximadas. A eficiência do sistema em representar o sistema real é dada por R^2 ajustado de 0,6196, ou seja, 61,96% das variações nas variáveis podem ser explicadas pela equação de regressão.

As variáveis reais e simuladas são apresentadas parcialmente no Figura 4. A variável de resposta é chamada de y´ e a variável preditora é chamada de x´.

Figura 4 - Valores amostrais e simulados dos indicadores de rendimento e conformidade.

Insira os dados das amostras nas colunas correspondentes		*Aqui estão as amostras simuladas com mil observações*	
X	**Y**	**X'**	**Y'**
74,58	97,63	82,34	91,86
75,53	97,07	82,00	93,38
72,67	97,76	78,90	96,74
75,39	98,23	68,76	97,63
74,00	94,06	75,72	96,88
74,40	97,64	75,54	96,13
73,97	98,96	71,49	100,03
74,65	94,38	74,40	95,31
75,31	94,18	82,94	95,33
75,17	96,67	77,73	98,94

Fonte: Elaborado pelos autores.

Os erros médios, que são apresentados no Quadro 1, são muito satisfatórios. O índice de correlação obteve um erro de 0,12%, enquanto os coeficientes linear e angular apresentaram desvios de 0,04% e 0,12%, respectivamente. Esse desempenho indica que a amostra gerada pelo modelo se aproxima muito das amostras reais.

Quadro 1 - Comparação dos coeficientes de regressão e correlação originais e simulados.

Coeficientes	***R***	***Linear (a)***	***Angular (b)***
Real	-0,68	133,66	-0,49
Simulado	-0,68	133,71	-0,49
Erro	0,12%	0,04%	0,12%

Fonte: Elaborado pelos autores.

O modelo de simulação oferece um resumo das características das variáveis, por exemplo, os valores máximos e mínimos que podem ocorrer diante das condições vigentes. Como observado no Quadro 2, os pontos máximos dos indicadores de conformidade e rendimento foram 92,92 e 104,90, indicando que mesmo sem nenhuma alteração ou melhoria no sistema, é possível obter resultados significativamente melhores que os atuais. Por outro lado, os valores mínimos obtidos na simulação foram 63,35 para conformidade e 85,94 para rendimento, mostrando que o sistema ainda é suscetível a resultados muito inferiores. Esses resultados expõem a falta de controle no desempenho de cada indicador individualmente.

Quadro 2 - Resultados da simulação e meta.

Simulação	*Conformidade*	*Rendimento*	*Paridade média*
Máximos	92,92	104,90	
Mínimos	63,35	85,94	0,54
Meta	90,87	91,04	0,82

Fonte: Elaborado pelos autores.

No Quadro 2, observou-se outro resultado importante da simulação para análise de regressão simples, o índice de paridade. Esse índice foi desenvolvido com o objetivo de encontrar o melhor resultado mútuo possível, alinhando equilíbrio e grandeza dos indicadores. O índice de paridade pode ser utilizado apenas para variáveis dentro da mesma escala numérica, o que normalmente é o caso de indicadores de desempenho. Isso porque esse valor não está ligado à correlação, mas sim à distância entre as médias de cada variável. Se o índice de paridade for baixo, podemos afirmar que os resultados dos indicadores estão ruins e desequilibrados. Quanto mais próximo de um estiver o resultado, melhores e mais próximos estarão os indicadores. O índice de paridade é dado por uma função dos indicadores e da diferença entre eles, pois não resolve ter resultados equilibrados se os mesmos forem muito ruins. Por exemplo, a paridade média de 0,54 apresentada na tabela mostra como o sistema opera atualmente, sem nenhum equilíbrio entre os indicadores.

5.2. Análise múltipla de rendimento

Uma relação inversamente proporcional entre os indicadores de rendimento e conformidade foi encontrada a partir dos modelos de análise simples. Nos modelos de análise múltipla, é possível adicionar variáveis ao sistema e aprimorar os resultados da pesquisa como um todo. As variáveis escolhidas para análise múltipla foram o peso médio das aves e a quantidade de aves abatidas no dia, a serem confrontadas com os indicadores de rendimento e conformidade em análises separadas.

Os dados do indicador de rendimento foram inseridos no modelo misto de regressão como variável de resposta, enquanto os valores de peso médio e quantidade de aves abatidas foram inseridos como variáveis preditoras.

O sistema apresentou um coeficiente de determinação R^2 muito baixo de 0,0542, resultado que indica que a relação entre as três variáveis é pouco significativa. Após avaliação desse resultado, é necessário analisar a tabela de análise de variâncias. Para um nível de confiança estabelecido para o modelo de 95%, o valor de F tabelado de um sistema com os mesmos graus de liberdade é 3,19. Para inferir que a análise de regressão confirma uma relação interdependente no sistema, espera-se que o valor de F calculado seja maior que F tabelado. Contudo, o resultado obtido de 1,35 é menor que 3,19. Logo, pode-se afirmar que a hipótese de pelo menos uma das variáveis apresentar uma variância igual é falsa.

5.3. Análise múltipla de conformidade

A variável correspondente à conformidade foi definida como variável dependente e as outras duas como variáveis independentes, sendo x1 o peso médio e x2 a quantidade de aves abatidas. Os resultados são apresentados na Figura 5.

Figura 5 - Resultados do modelo de análise de regressão múltipla.

Insira um número de variáveis dependentes entre 1 e 5:	2
Insira os valores das variáveis no quadro ao lado	

R	R^2	β0	β1	β2
0,4852	0,2354	22,7069	6,7462	0,0002

Fonte	*Soma dos quadrados*	*Graus de liberdade*	*Soma dos q. média*	*F calculado*
Devido à regressão	315,18	2,00	157,59	7,04
Devido ao erro	1051,70	47,00	22,38	
Total	1366,88	49,00	F tabelado	3,20

Y	X1	X2
75,53	2,611	159956
72,67	2,643	152458
80,79	2,722	169484
75,39	2,670	163395
74,00	2,548	162704
74,40	2,661	153078
73,97	2,589	153444
74,65	2,540	156574
75,31	2,586	170350
75,17	2,536	166559
71,27	2,407	165060
82,62	2,549	155566
76,00	2,557	159514

Fonte: Elaborado pelos autores.

Os coeficientes calculados pelo modelo de regressão apresentam erro de menos de 1%. O índice de correlação de 0,4852 indica uma relação linear razoável entre as três variáveis, enquanto o coeficiente de determinação mostra que 23,54% das variações na variável de resposta podem ser explicadas pela equação de regressão linear. Com base nos coeficientes resultantes da análise de regressão, a equação da reta é dada pela Equação 8.

$$\text{Conformidade} = 22{,}7069 + 6{,}7462 * \text{Peso Médio} + 0{,}0002 * \text{Quantidade Abatida} \tag{8}$$

Na tabela de análise de variâncias, é possível identificar que o valor de F calculado é maior que F tabelado. Assim, pode-se concluir com risco de 0,05% que existe regressão linear múltipla e o modelo pode explicar e prever a variável dependente. Estabelecida a relação linear entre as variáveis, as mesmas foram inseridas no modelo de regressão e parte das tabelas é apresentada na Figura 6.

Figura 6 - Valores amostrais e simulados de peso médio, volume de abate e indicador de conformidade.

Insira os dados das amostras nas colunas correspondentes			*Aqui estão as amostras simuladas com mil observações*		
X1	**X2**	**Y**	**X1'**	**X2'**	**Y'**
2,611	159956	75,53	2,766	167041	77,40
2,643	152458	72,67	2,795	157838	80,64
2,722	169484	80,79	2,477	152478	75,91
2,670	163395	75,39	2,678	148648	74,90
2,548	162704	74,00	2,797	149886	69,81
2,661	153078	74,40	2,558	141382	74,61
2,589	153444	73,97	2,917	168630	79,35
2,540	156574	74,65	2,592	160284	71,46
2,586	170350	75,31	2,487	150067	74,14
2,536	166559	75,17	2,697	151419	69,36

Fonte: Elaborado pelos autores.

A Figura 7, por sua vez, mostra a matriz de correlação das amostras de dados coletados no sistema real e a matriz de correlação das amostras geradas pelo modelo de simulação. O erro de cada um dos índices simulados em relação ao valor real é muito baixo, o que evidencia a eficiência da utilização da decomposição matricial de Cholesky na reprodução das amostras do sistema real.

Figura 7 - Comparação entre matrizes de correlação real e simulada.

	Matriz de correlação real		
	X1	**X2**	**Y**
X1	1,0000	0,0352	0,1844
X2	0,0352	1,0000	0,4496
Y	0,1844	0,4496	1,0000

	Matriz de correlação simulada		
	X1	**X2**	**Y**
X1	1,0000	0,0184	0,1981
X2	0,0184	1,0000	0,4352
Y	0,1981	0,4352	1,0000

Fonte: Elaborado pelos autores.

Os coeficientes de regressão constam no Quadro 3, com um erro de até 6,0% em relação aos coeficientes originais.

Quadro 3 - Comparação entre coeficientes de regressão real e simulado.

	Coeficientes de Regressão		
	β0	**β1**	**β2**
Real	22,71	6,75	0,00
Simulado	21,35	7,13	0,00
Erro	6,00%	5,70%	0,80%

Fonte: Elaborado pelos autores.

Assim como no modelo para regressão simples, esse modelo múltiplo também oferece os valores máximos e mínimos para cada variável, como demonstrado no Quadro 4. Para um indicador de conformidade dependente de ambas variáveis, esse descontrole pode levar a condições preocupantes.

Quadro 4 - Resultados da simulação.

	Resultados da simulação		
	X1	**X2**	**Y**
Máximos	3,121	186359	93,30
Mínimos	2,314	121839	60,79
Medianas	2,704	158561	77,32

Fonte: Elaborado pelos autores.

Na simulação de múltiplas variáveis, os valores ideais de cada variável independente são calculados pelas médias dos pontos nas amostras de simulação nos quais a resposta correspondente atingiu o valor desejado pelo usuário. Os resultados estão no Quadro 5 e mais uma vez, foi utilizada como exemplo a meta de indicador de conformidade buscada pela empresa, que é 90,0. Para essa meta, o cenário mais favorável é aquele com um peso médio de 2,716kg e um volume diário de abate de 165.656 aves.

Quadro 5 - Ferramenta de determinação de valores para atingimento de meta.

<table>
<tr><td rowspan="2">Insira o valor desejado da variável de resposta</td><td>Y</td><td></td></tr>
<tr><td>90,00</td><td></td></tr>
<tr><td rowspan="2">Aqui estão os valores ideais para obter o resultado desejado</td><td>X1</td><td>X2</td></tr>
<tr><td>2,716</td><td>165656</td></tr>
</table>

Fonte: Elaborado pelos autores.

A partir dos valores ideais encontrados para cada variável independente, o modelo traz as probabilidades da simulação. A probabilidade apresentada no Quadro 6 é calculada individualmente para cada variável pela função cumulativa de distribuição normal e pela proporção real encontrada nas amostras geradas pela simulação.

Quadro 6 - Ferramenta de determinação das probabilidades de atingimento da meta.

Probabilidade de obter o resultado desejado				
Distribuição	**X1**	**X2**	**X1 e X2**	**Y**
Normal	46,10%	23,60%	10,90%	0,90%
Simulação	45,70%	23,40%	11,00%	0,70%
Erro	0,40%	0,20%	0,10%	0,20%

Fonte: Elaborado pelos autores.

A análise demonstra uma instabilidade nos resultados obtidos para todas as variáveis. No entanto, o processo apresenta a melhor eficiência quando se trata de peso médio, onde se identificou uma probabilidade de aproximadamente 46% de atingimento da meta. A probabilidade da variável de quantidade abatida alcançar o valor ideal é ainda mais baixa, em torno de 23,5%. Apenas 11% dos valores gerados pela simulação foram ideais tanto para o peso médio quanto para quantidade abatida.

6. CONSIDERAÇÕES FINAIS

A simulação da análise de regressão abriu novos horizontes na interpretação dos indicadores de desempenho. Foram extraídas informações essenciais ao gerenciamento dos sistemas de medição de desempenho, entre elas, o índice de paridade. O indicador desenvolvido pela pesquisa oferece uma maneira eficaz de controle das medidas de desempenho. Ao acompanhar a evolução no índice paridade, é possível avaliar as melhorias individuais nos indicadores de conformidade e rendimento enquanto se controla o equilíbrio de resultados. O controle permite identificar quando um indicador é tratado com mais prioridade, ou ações que melhoram o resultado em um indicador enquanto prejudicam o outro.

No desdobramento dos índices em análises de regressão múltipla, verificou-se pouca ou nenhuma relação de dependência entre o indicador de rendimento e as variáveis de peso médio e quantidade abatida. O problema pode estar no próprio indicador, no que se refere ao cálculo e composição e interferências externas. Apesar do embasamento teórico e estatístico realizado, essa relação é inexistente. Ainda que relativamente baixa, a inter-relação na análise múltipla do indicador de conformidade foi encontrada e comprovada na análise de variâncias. A simulação para análise múltipla também foi bem-sucedida, apresentando resultados satisfatórios. Através de cálculos implícitos do modelo, foi possível identificar os valores ideais de peso médio e quantidade abatida que comporiam o melhor cenário para obtenção de determinado valor de conformidade.

Os modelos de análise de regressão e simulação oferecem ao usuário um diagnóstico do sistema de medição de desempenho. A interpretação dos resultados não somente identifica a saúde dos indicadores de desempenho, mas viabiliza oportunidades e define objetivos para a melhoria do processo. Cada resultado pode ser interpretado de maneiras diferentes, incitando o surgimento de novas ideias. Como sugestão para trabalhos futuros, outras avaliações poderiam ser incorporadas ao modelo como, por exemplo, as variáveis categóricas.

REFERÊNCIAS

BASTOS, E. V. P.; GUIMARÃES, J. C. F.; SEVERO, E. A. Modelo de regressão linear para análise de investimentos em uma empresa do ramo petrolífero. **Revista Produção e Desenvolvimento**, v. 1, n. 1, p. 77-88, 2015.

BODEA, C. N.; PURNUŞ, A. Project risk simulation methods: a comparative analysis. **Management & Marketing, Challenges for the Knowledge Society**, v. 7, n. 4, p. 565-580, 2012.

HILLIER, F. S.; LIEBERMAN, G. J. **Introduction to Operations Research**. 9 ed. New York: McGraw Hill, 2012.

KRISHNAMOORTHY, A.; MENON, D. **Matrix inversion using cholesky decomposition**. ST-Ericsson India Private Limited, Bangalore, India, 2014.

LAKATOS, E. M.; MARCONI, M. A. **Metodologia científica**. 6 ed., São Paulo: Atlas, 2011.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). **"NASA's Approach to Performance Management"**, 2015. Disponível em: <http://www.nasa.gov/news/budget/index.html>. Acesso em: 06 jun. 2016.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). **Analytic Method for Probabilistic Cost and Schedule Risk Analysis**, Final Report, 2013.

RISPOLI, F. J.; SHAH, V. Using simulation to test the reliability of regression models. **Energy and Environment Research**, v. 5, n. 1, p. 75-81, 2015.

RUI-MEI, L. Properties of Monte Carlo and its application to risk management. **International Journal of u- and e- Service, Science and Technology**, v. 8, n. 9, p. 381-390, 2015.