

PRECIFICAÇÃO DE FRETES RODOVIÁRIOS VIA MODELOS PARAMÉTRICOS

ROAD FREIGHT PRICING VIA PARAMETRIC MODELS

Luísa Maia Coppe* E-mail: luisacoppe@yahoo.com.br

Diego Augusto Diniz* E-mail: diegoaudiniz@gmail.com

Vyctor Henrique de Carvalho* E-mail: vyctor_carvalho@hotmail.com

Lívia Martins da Costa Furtado Pimentel* E-mail: liviafurtado@pucminas.br

*Pontifícia Universidade Católica de Minas Gerais (PUC Minas), Belo Horizonte

Resumo: O transporte pode representar, dentro da estrutura de custos dos produtos, até 60% do seu valor total. No Brasil o modal rodoviário é o mais utilizado, configurando 61% dos fretes realizados. A precificação do transporte rodoviário, no entanto, não é um processo trivial por envolver variáveis que vão além dos custos básicos da atividade. O objetivo deste trabalho é o desenvolvimento de um modelo de precificação de fretes com maior grau de precisão e tempo de resposta reduzido, utilizando técnicas de regressão linear múltipla. O estudo é aplicado a uma empresa de operação logística cujo modelo de negócios é a conexão entre o transportador e o demandante do frete. O modelo de precificação desenvolvido obteve um coeficiente de determinação de $R^2 = 0,9966$. Isto é, um modelo altamente explicativo, visto que explica 99,66% da variabilidade dos valores de fretes observados. As variáveis preditoras que compõem o modelo são: distância, custo do motorista, tipo de cliente, valor da nota fiscal, origem e destino. Com o modelo desenvolvido, as variáveis tiveram seu peso estatisticamente calculado, conferindo automatização ao processo e tornando mais ágil a rotina dos analistas da empresa.

Palavras-chave: Transporte rodoviário. Logística. Precificação. Regressão Linear Múltipla.

Abstract: In reference to product cost structure, transportation can represent up to 60% of its total value. In Brazil, road transportation mode is the most used kind, making up to 61% of the freight operated as a whole. However, road transportation pricing is not a trivial process because it involves variables that go beyond the basic costs of the activity. The objective of this work is the development of a freight pricing model with a higher degree of precision and reduced response time, using multiple linear regression techniques. The study is applied to a logistics operation company, which business model is the connection between the carrier and the freight demander. The pricing model developed obtained a coefficient of determination of $R^2 = 0.9966$. In other words, a highly explanatory model, once it explains 99.66% of the variability of observed freight values. The predictive variables that make up the model are distance, driver cost, type of customer, invoice amount, origin and destination. With the model developed, the variables had their weight statistically calculated, giving automation to the process and making the routine of the company's analysts more agile.

Keywords: Road transport. Logistics. Pricing. Multiple Linear Regression.

1 INTRODUÇÃO

Quando se analisa o transporte, sob a ótica de custos, conclui-se que ele representa o elemento mais importante do custo logístico na maioria das empresas.

Conforme Nazário (*apud* FLEURY; FIGUEIREDO; WANKE, 2000), o transporte representa, em média, cerca de 60 % das despesas logísticas e em alguns casos supera até o lucro operacional. Dessa forma, a terceirização desse serviço apresenta relevante importância para as empresas que buscam reduzir seus custos e reafirma a oportunidade criada para os operadores logísticos. Conforme Fleury, Figueiredo e Wanke (2000), um operador logístico é um fornecedor de serviços integrados, com a capacidade de atender de forma integrada a todas ou quase todas necessidades logísticas de seus clientes.

A logística possui origem militar, especificamente na Segunda Guerra Mundial, como garantidora de todos os meios necessários à sobrevivência das tropas no campo de batalha, como condições de movimentação, abastecimento, alojamento e transporte (MEDINA, 2002; SOUZA, 2002). No período pós-guerra, o conceito se difundiu e passou a ser utilizado por empresas e organizações, tornando-se a área responsável pelo controle e armazenamento de mercadorias e insumos (PLATT; KLAES, 2010). A logística, conforme Christopher (2007), é o conjunto de processos de gerenciamento estratégico da compra, do transporte e da armazenagem de matérias primas, partes e produtos acabados.

Nessa perspectiva, a Confederação Nacional dos Transportes – CNT, em seu *Boletim Estatístico* (CNT, 2019), descreve a participação dos cinco principais modais de transporte na movimentação de cargas em território brasileiro. De acordo com esta referência, o transporte rodoviário foi responsável por 61% das cargas, o ferroviário por 20,7%, o aquaviário por 13,6%, o dutoviário por 4,2% e o aéreo por 0,4%. Tais dados estatísticos reiteram a opção histórica de valorizar o modal rodoviário em detrimento dos outros modais existentes.

No Brasil, conforme dados constantes no *Anuário CNT do transporte* (CNT, 2018), somente no modal rodoviário de transporte de cargas existem 147.177 transportadoras regularmente inscritas, 11.482 empresas do setor de locação de veículos, 382 cooperativas e 492.408 profissionais autônomos registrados. Apenas no ano de 2017, foram realizados 51.941 novos registros de caminhões e 319.890 registros de veículos comerciais leves. Todos esses veículos circulam em uma malha rodoviária de extensão aproximada de 1.720.700,3 km. Esses dados reforçam que o mercado está cada vez mais competitivo.

Devido à complexidade operacional, ao crescente aumento de sofisticação tecnológica e a contínua busca por redução de custos, os gestores das empresas buscam cada vez mais contratar um operador logístico para realização do transporte. Conforme Fleury (2002), existem 7 critérios utilizados para selecionar os prestadores de serviços de transporte: confiabilidade; preço; flexibilidade operacional; flexibilidade comercial, saúde financeira; qualidade do pessoal operacional; e informações de desempenho. O desenvolvimento da indústria de transportes fretados levou a um aumento da competição entre essas empresas e conseqüentemente o tema da precificação do serviço ganhou mais atenção entre pesquisadores. O preço é uma variável de controle cuja alteração é uma das maneiras mais eficientes de se encorajar ou desencorajar a demanda no curto prazo (MOZAFARI; KARIMI, 2011).

Entretanto, precificar um serviço não é tão trivial quanto precificar um produto de bens de consumo. Conforme Martins (2008), a precificação do transporte é bastante complexa, pois, além dos custos da atividade, incorpora também fatores locais e conjunturais. Ortúzar e Willumsen (1994) e Caixeta Filho e Martins (2011) fizeram uma lista com 7 variáveis que influenciam no estabelecimento do preço do frete, são elas: distância percorrida; especificidade da carga transportada e do veículo utilizado; prazo de entrega da carga; custos operacionais; sazonalidade da demanda; perdas e avarias; características e aspectos geográficos das vias utilizadas; pedágios; e a possibilidade da carga de retorno para as zonas de origem.

Sendo assim, considerando a importância do transporte rodoviário no Brasil, é fundamental analisar e precificar este serviço de maneira a manter a competitividade da empresa, gerando lucro para ela e fomentando ainda mais o transporte e a geração de empregos no setor. Dado isto, este trabalho aborda a modelagem e precificação do serviço de transporte rodoviário numa empresa localizada na região metropolitana de Belo Horizonte, a Dede Transportes.

A Dede Transportes Ltda. iniciou suas operações de transporte rodoviário de carga em 2003 e, desde então, oferece um serviço diferenciado devido a seu preço acessível combinado com a qualidade oferecida, além de um relacionamento sólido e transparente com os clientes. A empresa possui seu galpão em Belo Horizonte MG

e seu escritório administrativo em Contagem – MG, e atende todo o território nacional, atuando principalmente no sudeste do Brasil.

Cabe destacar que a empresa possui um modelo de negócio diferenciado de uma transportadora comum. Apesar de ser uma transportadora, possui apenas três caminhões próprios, visto que a principal atividade da empresa é conectar seus clientes à prestadores de serviços. Devido aos custos para manter uma frota própria, foi estabelecido que a empresa teria uma rede de motoristas credenciados, cada qual com o seu veículo próprio, que realizam os fretes de acordo com a demanda da empresa.

Atualmente, a empresa possui uma sólida carteira de clientes em diversos ramos, operando com transportes de equipamentos tanto de baixo quanto de alto valor agregado, tais como: máquinas agrícolas, tratores, móveis e containers. Além disso, também transporta materiais diversos de obra, blocos de concreto, ração animal, produtos alimentícios, dentre outros. Realizando em média cerca de 400 fretes ao mês com cerca de 90 clientes cadastrados.

O objetivo geral deste trabalho é conceber um modelo de precificação para transporte rodoviário que opere de maneira mais autônoma de modo a facilitar a atuação da equipe de vendas e que mantenha a competitividade da empresa, atendendo seus clientes, satisfazendo as necessidades dos parceiros e gerando lucro. Para isso, analisamos o modelo de precificação atual, buscando oportunidades de melhorias e foram coletados dados com informações sobre últimas precificações de fretes. Técnicas de modelos paramétricos foram utilizadas para analisar a atual precificação e desenvolver um modelo de precificação com maior grau de automação para reduzir o tempo de resposta do analista e auxiliar a tomada de decisão.

O restante do artigo é organizado da seguinte forma. O referencial teórico é apresentado na Seção 2. A Seção 3 fornece a metodologia de pesquisa. O desenvolvimento do modelo de precificação está descrito na Seção 4, assim como análise dos resultados e sua validação. A Seção 5 é dedicada à conclusão.

2 REFERENCIAL TEÓRICO

Nesta seção, apresenta-se a revisão da literatura do tema proposto com o objetivo de oferecer contextualização e consistência à precificação. São apresentados referenciais teóricos sobre a precificação e regressão linear múltipla.

2.1 Precificação

Apesar de existirem vários conceitos conflitantes, muitos autores convergem ao dizer que o preço é um ponto fundamental na definição das estratégias empresariais, posicionamento de mercado e tipo de cliente a ser atingido, influenciando diretamente os resultados das organizações (CARNEIRO *et al.*, 2018). Entretanto, de acordo com Milan *et al.* (2016), muitas empresas ainda não possuem processos definidos e baseados em pesquisas para estabelecimento dos seus preços, mas sim o fazem de acordo com a intuição ou experiências dos gestores. O referido autor cita, ainda, os constantes conflitos entre os levantamentos de custos e estratégias de marketing que interferem na assertividade da precificação.

Segundo Araújo, Bandeira e Campos (2014), a formação de preços de fretes rodoviários é complexa por levar em consideração, além de seus custos (fixos, variáveis, diretos e indiretos), também fatores locais, conjunturais, de oferta e demanda. Analisando as abordagens descritas para composição de preços de fretes rodoviários, evidencia-se uma concentração na utilização da distância como principal variável. Já as demais variáveis (custo operacional, possibilidade de obter carga de retorno, demandas, especificidades das cargas, dentre outras) possuem apenas influência complementar na precificação (CYPRIANO *et al.*, 2006). No Brasil, legislações burocráticas e fiscalizações ineficazes, corroboram para a diminuição dos preços e da qualidade do serviço prestado (PEREIRA; LESSA, 2011).

Para precificar produtos e serviços em geral, incluindo fretes rodoviários, as estratégias podem ser categorizadas em três abordagens, que também são chamadas de métodos de fixação de preços (MILAN *et al.*, 2016). A primeira delas, é a de precificação baseada nos custos, em que se utilizam os custos básicos do serviço como ponto de partida para determinação dos preços. Outra abordagem possível é a precificação baseada na concorrência, em que se faz um estudo dos

preços praticados pelo mercado e essa é a fonte primária para fixação dos preços. O último método descrito é o de precificação baseada no valor ou na demanda, que leva em consideração o valor percebido pelo cliente (ou seja, quanto ele está disposto a pagar) como base para a determinação de preços (ARAÚJO; BANDEIRA; CAMPOS, 2014).

Nos últimos 10 anos, a precificação de fretes rodoviários foi abordada em uma gama de pesquisas, incluindo: a análise do impacto comportamental da precificação de transportes via Teoria dos Jogos (HOLGUÍN-VERAS, 2010); Mozafari e Karimi (2011), no qual um modelo de jogos não cooperativos é utilizado; King, Gordon e Peters (2014) avaliam o efeito da precificação de fretes rodoviários nas atividades de transporte portuário; a diferença de preço entre fretes de transportadores brasileiros autônomos e próprio das empresas é abordada em Araújo, Bandeira e Campos (2014); e a análise de Demir *et al.* (2015) sobre o impacto no preço do transporte causado por externalidades negativas como poluição de ar, ruído, poluição de água, congestionamento e acidentes.

2.2 Regressão linear múltipla

Técnicas de regressão são amplamente utilizadas para auxiliar a precificação de produtos de bens e serviço. A relação entre o preço e variáveis explanatórias em economias de compartilhamento de acomodação é explorada em Wang e Nicolau (2017). Em Adrian, Crump e Moench (2013), técnicas de regressão linear são utilizadas na precificação de taxas de risco. Um modelo de regressão linear é utilizado em Hao *et al.* (2019) para determinar estratégias de precificação de tarefas em serviços de colaboração coletiva. Pereira, Garson e Araújo (2012) utilizam técnicas de regressão linear múltipla para modelar o preço de imóveis no estado de São Paulo. Similarmente, uma regressão linear é utilizada para analisar a influência da proximidade do litoral no valor de imóveis em Kim *et al.* (2020).

A análise de regressão é uma técnica estatística utilizada para a investigação e modelagem da relação entre a variável resposta e a variável explicativa, de acordo com Montgomery (2012). Quando a situação problema envolve mais de uma variável explicativa, utilizamos o modelo de regressão linear múltipla, apresentado pela seguinte equação:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (1)$$

onde Y a variável resposta; β_0 o intercepto; x_1, x_2, x_p as variáveis explicativas; β_j para $j=1, \dots, p$ representa a variação em y , causada pelo aumento de uma unidade do x_j , quando as demais variáveis são mantidas constantes e ε representa o erro.

Em notação matricial, tem-se que:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \text{ e } \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2)$$

Reescrevendo (1) em função de (2), tem-se:

$$y = X\beta + \varepsilon \quad (3)$$

onde y um vetor $n \times 1$, X uma matriz $n \times (p + 1)$, β um vetor $p \times 1$ e ε um vetor $n \times 1$.

Os valores dos coeficientes em β_s são desconhecidos e para serem estimados é necessária uma base de dados com n observações $(y_{11}, x_{11}, x_{12}, x_{1p})$, $(y_{21}, x_{21}, x_{22}, x_{2p})$, $(y_{n1}, x_{n1}, x_{n2}, x_{np})$. A partir da base é utilizado o método dos mínimos quadrados, no qual deve-se minimizar a soma da diferença entre o valor observado e o valor estimado ao quadrado. Conforme mostra a expressão:

$$\sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})]^2 = (y - X\beta)' (y - X\beta) \quad (4)$$

A fim de minimizar a expressão deve-se derivá-la simultaneamente em termos de β e igualá-la a 0. Assim, tem-se:

$$\frac{\partial y'y}{\partial \beta} - 2 \frac{\partial y'X\beta}{\partial \beta} + \frac{\partial \beta'X'X\beta}{\partial \beta} = 0 - 2(y'X)' + 2X'X\beta \quad (5)$$

onde, o vetor de dimensão $p + 1$ resulta na expressão $\hat{\beta} = (X'X)^{-1}X'y$, em que $\hat{\beta}$ é o estimador não viciado do modelo, desde que $(X'X)^{-1}$ exista.

Sendo assim, o modelo de regressão ajustado é dado por:

$$\hat{y} = X\hat{\beta} \quad (6)$$

e o vetor de resíduos é dado por:

$$\varepsilon = y - \hat{y} = y - X\hat{\beta} \quad (7)$$

De acordo com Montgomery, Peck e Vining (2012), depois da modelagem, é necessário analisar a significância da regressão, para determinar se há de fato uma relação linear entre a variável resposta e cada uma das variáveis explicativas. Para isto, é utilizado o teste de hipóteses, onde:

$$H_0: \beta_0 = \beta_1 = \dots = \beta_p = 0 \quad \text{versus} \quad H_1: \text{ao menos um } \beta_i \neq 0 \quad (8)$$

Considerando um nível de significância de 5% a hipótese nula é rejeitada se o p-valor é menor que 0,05. Caso a hipótese nula seja rejeitada, o modelo paramétrico existe, pois estatisticamente ao menos um coeficiente é diferente de 0.

Para analisar a adequação do modelo pode-se calcular o coeficiente de determinação, que é dado por:

$$R^2 = \frac{SQ_{Reg}}{SQ_T} = \frac{SQ_{Reg}}{SQ_{Reg} + SQ_E} \quad (9)$$

onde SQ_{Reg} é a soma dos quadrados devido a regressão e SQ_T é a soma total dos quadrados devido a regressão e devida ao erro (SQE). O R^2 mede qual o percentual

da variabilidade dos dados é explicado pelo modelo. Ou seja, quanto maior, mais explicativo é o modelo construído.

2.3 Erro Quadrático Médio (MSE – *Mean Squared Error*)

O MSE de um estimador é o valor esperado do erro ao quadrado e se dá pela seguinte fórmula:

$$MSE(\hat{\theta}) = E[(\theta - \hat{\theta})^2] \quad (10)$$

O ideal é ter um MSE igual a 0, o que significaria que o estimador não contém viés. Entretanto, em situações reais, isto é extremamente raro de acontecer. Sendo assim, quanto menor o MSE, melhor é o estimador.

No contexto da regressão linear múltipla com uma amostra de n observações, o MSE pode ser calculado através da seguinte expressão:

$$MSE(\hat{y}_i) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (11)$$

3 METODOLOGIA

Nesta seção são abordados os seguintes assuntos: a classificação da pesquisa e o fluxograma que descreve passo a passo como a pesquisa será realizada.

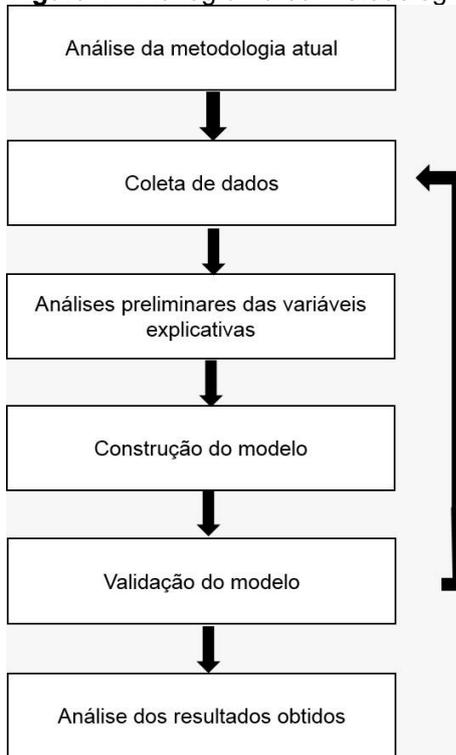
3.1 Classificação da pesquisa

O trabalho consiste na realização de uma análise do modelo atual e uma nova modelagem para precificação de frete, sendo uma pesquisa descritiva e quantitativa. Conforme Raupp e Beuren (2006), a pesquisa quantitativa preocupa-se com o comportamento geral dos acontecimentos, onde ferramentas estatísticas são utilizadas para a coleta e tratamento de dados. No caso deste trabalho, é a análise de regressão linear múltipla, conforme descrito na Seção 2.2.

3.2 Fluxograma

O fluxograma abaixo, descreve sinteticamente a metodologia utilizada no trabalho:

Figura 1 - Fluxograma da metodologia



Fonte: Os autores.

Conforme a Figura 1, o primeiro passo para a construção do modelo de precificação é analisar e discutir a metodologia atual utilizada pela empresa. Em particular, é fundamental identificar variáveis que são consideradas mais importantes. Após a identificação, será feita a coleta de dados e construção da base de dados, utilizando informações fornecidas pela própria empresa.

O terceiro passo é realizar análises preliminares dessas variáveis explicativas, utilizando o software R. Neste tópico serão feitos alguns gráficos para ilustrar a relação de cada variável com o preço do frete. Além disso, é construído um modelo de regressão simples individual para cada possível variável explicativa.

Após a análise inicial, são construídos modelos de regressões múltipla com aquelas variáveis mais explicativas. O melhor modelo feito será validado através da sua utilização para predição de 100 novos valores de frete. Os valores obtidos são comparados com os reais. Como medida de validação, são usados: (i) o MSE; e (ii)

análise dos resíduos. Caso o modelo não seja válido, a base de dados deve ser reanalisada e complementada com outras possíveis variáveis. Por fim, é analisado os resultados obtidos.

4 DESENVOLVIMENTO DO MODELO

Nesta seção são apresentadas as etapas da construção do modelo de precificação. São elas: análise da metodologia atual, coleta de dados, análises preliminares, construção do modelo múltiplo e a análise dele.

4.1 Análise da metodologia atual

A metodologia atual de precificação utilizada na empresa, é totalmente empírica. Não existe uma técnica estatística utilizada para a precificação dos serviços. Cada analista, conforme seus conhecimentos e intuições, determina o preço a ser cobrado por frete.

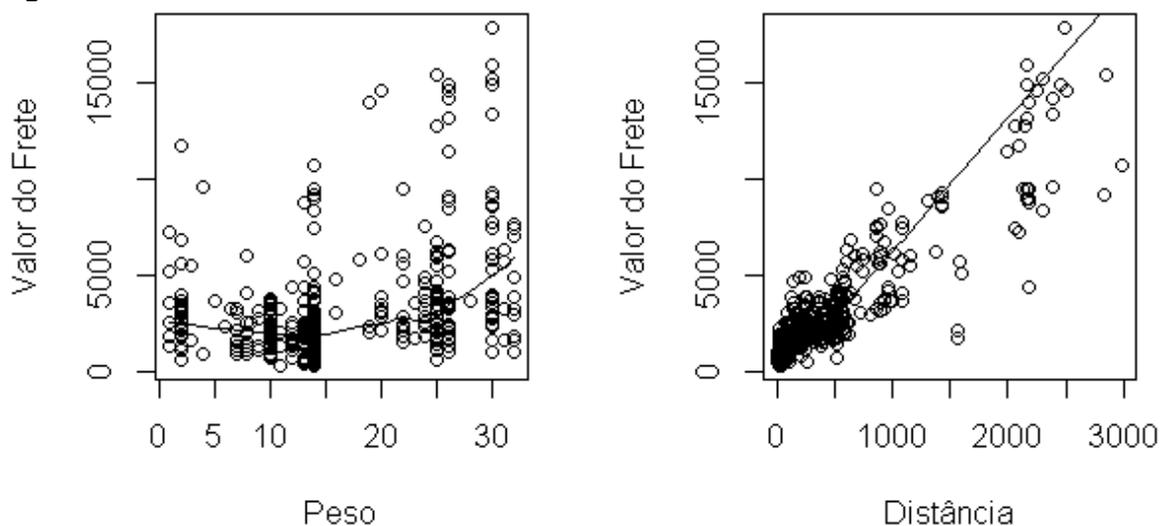
4.2 Coleta de dados

Para realizar o estudo, foi coletada, através de dados fornecidos pela empresa, uma base de dados contendo todas as informações consideradas relevantes para a atividade de precificação de um serviço de transporte. A base possui 918 linhas (ou observações), cada qual referente a um frete precificado. Cada coluna dessa base contém informações sobre: origem, destino, se possui carga de retorno ou não, distância percorrida, distância categorizada, tipo de veículo, tipos de carga, peso, valor da mercadoria, alíquota de imposto incidente, cliente, valor do frete e custo do motorista. Ou seja, cada coluna é uma variável potencialmente explicativa do valor do frete.

4.3 Análises preliminares

Com o objetivo de analisar a relação entre algumas variáveis e o preço do frete, foram feitos os seguintes gráficos:

Figura 2 - Gráfico Peso x Valor do Frete e Distância x Valor do Frete



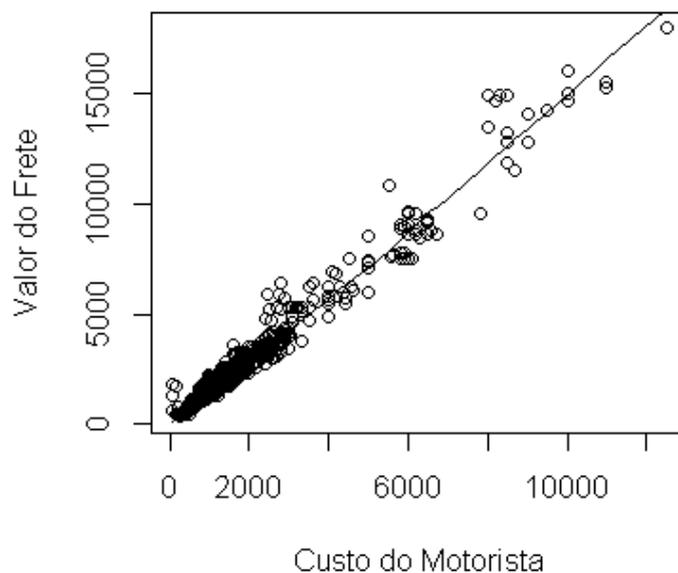
Fonte: Os autores.

Quando se analisa o valor do frete em relação ao peso, pode-se observar que há uma linha de tendência no gráfico esquerdo da Figura 2. Apesar da linha não ser exatamente uma reta, observa-se que, de modo geral, conforme esperado, quanto maior o peso, maior o valor do frete a ser cobrado. É interessante ressaltar que existem pesos com maior incidência, como 14, 26 e 32. Uma possível explicação para este fenômeno é que estes são os pesos máximos de alguns tipos de veículo utilizados pela empresa e, como a empresa trabalha geralmente com carga cheia, tende-se a utilizar o peso máximo por veículo.

Em relação a análise do gráfico da direita na Figura 2, é interessante observar 3 pontos: (i) existe claramente uma relação positiva entre distância e valor do frete, o que já era esperado conforme a literatura (ii) as distâncias com maiores frequências na base de dados se encontram entre 0 e 500 Km, o que reflete a realidade da empresa estudada e (iii) a partir da distância 1500 Km, os dados não seguem tão bem a linha de tendência. Uma hipótese que pode justificar este fenômeno é o fato de não só a distância ser responsável pelo valor do frete. Provavelmente, estas cargas com alta distância e valor do frete não tão alto, possuem baixo valor na mercadoria transportada. Outra justificativa seria que, apesar da distância, não há incidência de impostos, ou são cargas para capitais e cidades que possibilitam o frete de retorno.

Outra análise importante de ser feita é a comparação entre o preço do frete e o custo com o motorista, ilustrada na Figura 3.

Figura 3 - Gráfico Custo do Motorista x Valor do Frete



Fonte: Os autores.

Pode-se observar no gráfico da Figura 3 que existe uma relação positiva entre o custo do motorista e o valor do frete. Esse comportamento é esperado pois, conforme a literatura, uma das possíveis abordagens para precificação é a baseada no custo. Como a empresa não possui frota própria, seu maior custo é o valor pago aos motoristas, que incorporam valor de pedágio, preço do diesel, manutenção do caminhão entre outros no valor que cobram da empresa. Vale ressaltar que esta informação do custo do motorista é obtida antes da precificação. Isto é, ela é uma variável de entrada no modelo da precificação e não uma consequência.

Após as análises preliminares e antes de ajustar modelos de regressão linear múltipla, são ajustados modelos individuais para cada variável explicativa. Os p-valores e os valores de R^2 para cada variável se encontram na Tabela 1.

Tabela 1 - Análise de Regressão Simples

Variável Explicativa	p- valor	R2
Custo do motorista	2.2e-16	0,9825
Distância categorizada	2.2e-16	0,8976
Distância	2.2e-16	0,8707
Destino	2.2e-16	0,7890
Destino UF	2.2e-16	0,6397
Origem	2.2e-16	0,5124
Imposto	2.2e-16	0,4004
Tipo de Veículo	2.2e-16	0,3805
Tipo de Carga	2.2e-16	0,2507
Peso	2.2e-16	0,2082
Valor da mercadoria	2.2e-16	0,1662
Carga Retorno	2.2e-16	0,0839
Origem UF	2.2e-12	0,0802
Tipo de Cliente	3.5e-06	0,0303

Fonte: Os autores.

Todas as variáveis estudadas possuem p-valores abaixo de 0,05. Ou seja, considerando um nível de significância de 5%, rejeita-se a hipótese nula na qual o coeficiente β_j da variável x_j é igual a 0, e aceita-se a hipótese alternativa, na qual o $\beta_j \neq 0$. Portanto, as variáveis preditoras x_j são de fato explicativas do comportamento da variável resposta y .

Apesar do p-valor indicar que as alterações no valor de algumas variáveis preditoras estão relacionadas a alterações na variável resposta, não são todas elas que possuem um alto grau de explicação, R^2 . O “tipo de cliente”, por exemplo, em um modelo individual, explicaria apenas 3,03% da variabilidade no valor do frete. Por isso, faz-se necessário a criação de um modelo múltiplo, que irá considerar todas as variáveis que têm efeito no valor do frete e que, desconsiderando a correlação entre as variáveis, terá um maior grau de explicação.

Além disso, observa-se que as variáveis isoladas com maior poder explicativo individual do valor do frete são: o “custo do motorista” e a “distância”, tanto categorizada quando numérica. Isto era esperado pois, quanto maior a distância, maior o custo que o motorista irá ter e cobrar. Conseqüentemente, maior o valor que a empresa precisa cobrar dos seus clientes.

Analisando a variável “UF Destino”, era esperado que ela fosse significativa pois a alíquota de imposto recolhido depende do estado de destino. Por exemplo, no caso do ICMS, esse será de 7 % se o destino for Espírito Santo e 12% se for Rio Grande do Sul. Porém, como já existe na base a variável “imposto”, e ela não é tão significativa quanto o estado destino. Conclui-se que o estado destino tem um maior poder explicativo do que o esperado, e que seu poder explicativo se dá além da alíquota de imposto.

A variável “tipo de veículo” está relacionada com quantidade de carga máxima que poderá ser transportada. Era esperado que esta variável fosse significativa pois há a suposição de que quanto maior a quantidade ou o peso da carga transportada, maior deverá o valor cobrado no frete. Porém, como já existe a variável “peso”, que tem um poder explicativo menor que o “tipo de veículo” e nem sempre os veículos transportam a carga máxima, conclui-se que este não é o único fator que torna esta variável explicativa. Uma possível explicação para a significância desta variável é o fato de determinados tipos de caminhão serem mais fáceis de contratar do que outros, pois conforme a lei da oferta e da demanda, quanto maior a oferta de motoristas com determinado tipo de caminhão, menor o valor que será pago por ele.

Além disto, é válido discutir a respeito da variável “tipo de carga”, que assim como as demais variáveis discutidas também explica mais de 25% do valor do frete isoladamente. Já era esperado que ela fosse significativa pois como cada “tipo de carga” possui um grau de dificuldade para ser transportado, esta dificuldade terá reflexo no valor do frete. Por exemplo, é mais fácil e, conseqüentemente, mais barato, transportar containers do que andaimes, pois geralmente os containers vão vazios, são mais leves e mais fáceis de carregar e descarregar do que os andaimes. Além disso, cada tipo de carga tem um determinado valor e este, tem reflexo no preço, já que quanto mais valiosa e mais difícil de ser transportada, mais caro deverá ser o frete.

A principal conclusão que se tem a partir da Tabela 1 é que, de fato, diversas variáveis tem impacto no valor cobrado de cada frete e que, isoladamente, nenhuma possui R^2 que justifique um modelo de regressão simples. Ou seja, apesar de individualmente elas explicarem parte do valor do frete, é necessário a construção de um modelo múltiplo com maior grau de explicação.

4.4 Modelo de regressão linear múltiplo

Após as análises preliminares, foi construído o modelo de regressão linear múltiplo. As variáveis consideradas no modelo são: “distância”, “custo do motorista”, “tipo de cliente”, “valor da nota fiscal”, “origem” e “destino”.

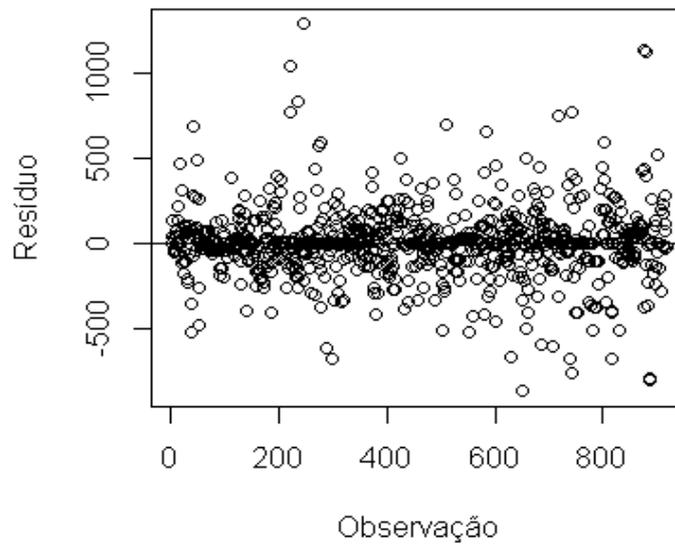
A saída fornecida pelo software estatístico para o modelo foi:

Residual standard error: 238 on 685 degrees of freedom
Multiple R-squared: 0.9966, Adjusted R-squared: 0.9955
F-statistic: 869.4 on 232 and 685 DF, p-value: < 2.2e-16

Conclui-se, portanto, que as variáveis incluídas no modelo foram extremamente significativas (p-valor do modelo menor que ordem de 10^{-16}), e que suas contribuições explicam 99,66% da variabilidade presente nos dados.

Outro aspecto importante para a validação de um ajuste de regressão linear múltipla é a análise de resíduos. A Figura 4 apresenta o gráfico de resíduos do modelo:

Figura 4 - Gráfico de Resíduos do modelo múltiplo ajustado



Fonte: Os autores.

Este gráfico indica que, conforme esperado, os resíduos possuem distribuição aleatória em torno de 0, com média de $2,12 \times 10^{-12}$ e variância de 42.315,22 validando a hipótese de que os dados coletados são independentes e que há homocedasticidade.

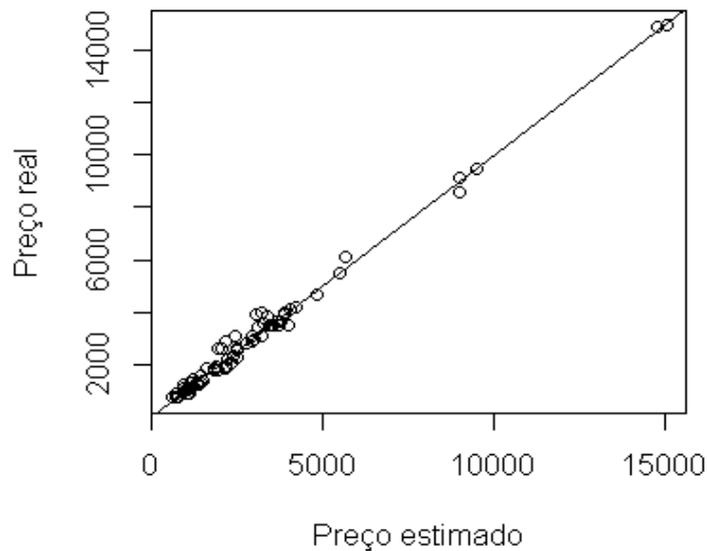
5 RESULTADOS

Esta seção é dedicada a análise dos resultados obtidos. É feita a validação do modelo, a comparação dos MSE e a análise de resíduos de novos valores preditos.

5.1 Validação do modelo

Para validar o modelo construído, foi criada uma nova base de dados com 100 observações. O modelo apresentado na Seção 4.4 foi utilizado para prever o valor desses novos fretes. Após, foi feita a comparação dos valores preditos com os valores reais. O gráfico apresentado na Figura 5 mostra a relação entre os valores preditos e os observados.

Figura 5 - Gráfico de valores preditos x valores observados



Fonte: Os autores.

Cada ponto no gráfico representa a combinação entre o valor real do frete (eixo y) e o valor estimado pelo modelo (eixo x) e a linha diagonal representa o conjunto de pontos que possuem o mesmo valor em x e em y. Sendo assim, quando um ponto está localizado exatamente na linha, isto indica que o valor previsto é o mesmo que o real. Caso o ponto se localize acima da linha, o valor real é maior que o estimado e se o ponto estiver abaixo o preço estimado é maior que o real.

Pode-se concluir que, de fato, o ajuste obtido pelo modelo desenvolvido é eficaz. Os preços estimados são satisfatoriamente próximos aos reais, com poucas variações aleatórias para mais ou menos o valor real do frete.

5.2 Comparação do MSE

Por ser um indicador de desvio, pode-se utilizar o MSE para avaliar a qualidade das novas previsões em relação ao modelo construído.

O MSE com a base do modelo, é 45315,22. Já o MSE da base de validação é 68238,65. O MSE das novas previsões é 50% maior que o primeiro. Entretanto, não se deve analisar este indicador isoladamente. O modelo, além de ser significativo, possui um alto coeficiente de determinação, resíduos que indicam que ele é válido e previsões com valores bem próximo dos observados.

Para aferir a qualidade do modelo, foi calculado também o MSE percentual das duas bases, através da fórmula:

$$MSE \% = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2 \quad (12)$$

Para analisar o erro médio percentual, foi calculada $\sqrt{MSE \%}$. A base utilizada no modelo teve um erro médio percentual de 10,57% enquanto a base utilizada para predição teve de 10,61%. Apesar dos valores serem altos, eles são muito próximos, o que é um bom indício. Além disso, foi considerado que um erro de até 5% é aceitável e foi então, calculada o percentual de observações que possuem o erro percentual inferior a 5%. Os resultados se encontram na Tabela 2.

Tabela 2 - Análise MSE

Base	Observações com erro inferior a 5 %	Observações com erro superior a 5 %
Modelo	95,99 %	4,01 %
Predição	95,00 %	5,00 %

Fonte: Os autores.

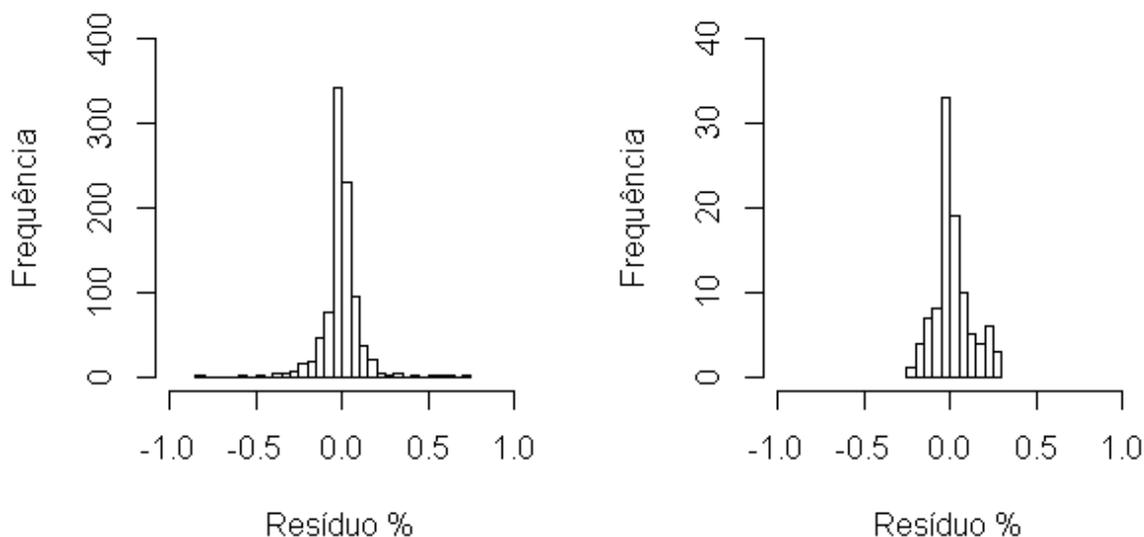
Conforme a Tabela 2, em ambas as bases, pelo menos 95% das observações obtiveram um erro inferior a 5%. Esse fato reafirma a qualidade do ajuste.

5.3 Comparação dos resíduos

Com o intuito de avaliar o comportamento de tendência central e dos resíduos, a Figura 6 apresenta o histograma tanto dos resíduos da predição da base utilizada na modelagem quanto da predição base de validação do modelo. Como um resíduo de determinado valor em reais pode ser significativo ou não, dependendo do valor total do frete, optou-se por analisar os resíduos em percentuais. O cálculo do resíduo percentual apresentado no gráfico se deu pela seguinte fórmula:

$$\varepsilon_i = \frac{(y_i - \hat{y}_i)}{y_i} \quad (13)$$

Figura 6 - Histograma dos resíduos para a base utilizada na modelagem e a base da predição



Fonte: Os autores.

Observa-se que para os resíduos da base do modelo (gráfico à esquerda da Figura 6), a faixa com maior frequência é entre 0 e -10% de erro, pode-se observar o mesmo comportamento para os resíduos da base utilizada na predição (gráfico a direita). Quando o resíduo possui um valor negativo, isto indica que o valor predito foi maior do que o observado, nesse caso o cliente da empresa estudada poderia optar por outro fornecedor com um valor de frete menor. Apesar de o percentual de erro ser pequeno, é recomendada a avaliação de um analista quanto ao valor fornecido pelo modelo.

Erros acima de 0 implicam que o valor predito é menor que o observado, o que pode fazer com que o cliente opte pela empresa estudada, mas a mesma terá um lucro menor do que o poderia ter caso cobrasse o valor observado.

De toda forma nota-se que, em ambos os gráficos, a distribuição e dispersão são similares, o que é mais um indício de que o modelo está bem ajustado.

6 CONCLUSÃO

O presente estudo propõe um modelo de precificação de fretes rodoviários baseado em regressão linear múltipla altamente explicativo. Sua utilização pode facilitar a rotina do analista da equipe de vendas e tornar o processo de precificação mais preciso. Apesar de a precificação um serviço ser menos trivial em comparação

à precificação de um bem de consumo, as técnicas utilizadas forneceram um modelo com coeficiente de determinação $R^2 = 0,9966$.

Os resultados apontaram que, de fato, diversas variáveis são importantes para explicar o comportamento preço do frete. Estas variáveis já eram consideradas na precificação pelos analistas com o peso que eles atribuíam individualmente, porém de maneira intuitiva. Com o modelo desenvolvido, cada variável tem sua participação na composição do preço final calculada estatisticamente e não empiricamente. Como a precificação também deve incorporar fatores locais e conjunturais, conforme Martins (2008), e estes fatores não foram considerados no modelo, cabe ainda ao analista avaliar o preço fornecido pelo modelo e as atuais conjunturas para definir qual o valor final a ser cobrado.

Na prática, seria necessário implantar a utilização do modelo com a atualização constante dos dados. Seria interessante planejar e desenvolver procedimentos com o software parceiro da empresa ou por meio de planilhas eletrônicas, que teriam a base de dados atualizadas periodicamente, de modo a garantir validade do mesmo.

Como limitação deste trabalho, destaca-se a impossibilidade de gerar a predição do valor a ser cobrado com o modelo escolhido para fretes com origem ou destino que não constam na base de dados. Uma possível solução seria utilizar um novo modelo sem a inclusão dessas variáveis preditoras. Este modelo não seria tão preciso e explicativo quanto o escolhido, mas em contrapartida ofereceria valor de parâmetro para o analista da empresa precificar o serviço. Vale ressaltar a existência de técnicas mais avançadas que a utilizada na modelagem. Espera-se, portanto, que futuros trabalhos possam dar prosseguimento no estudo da precificação de fretes através da utilização de técnicas mais sofisticadas, como o *multivariate adaptive regression splines* (ver Friedman (1991) e Martinez *et al.* (2015) para detalhes).

REFERÊNCIAS

ADRIAN, Tobias; CRUMP, Richard K.; MOENCH, Emanuel. Pricing the term structure with linear regressions. **Journal of Financial Economics**, v. 110, p. 110-138, 2013. <https://doi.org/10.1016/j.jfineco.2013.04.009>

ARAÚJO, M. P. S.; BANDEIRA, R. A. M.; CAMPOS, V. B. G. Custos e fretes praticados no transporte rodoviário de cargas: uma análise comparativa entre autônomos e empresas. **J. Transp. Lit.**, Manaus, v. 8, n. 4, out. 2014. Disponível em: http://www.scielo.br/scielo.php?pid=S2238-10312014000400009&script=sci_arttext. Acesso em: 20 mar. 2020. <https://doi.org/10.1590/2238-1031.jtl.v8n4a8>

CAIXETA FILHO, J.; MARTINS, R. **Gestão logística do transporte de cargas**. São Paulo: Atlas, 2011.

CARNEIRO, Jorge M. T. *et al.* **Formação e administração de preços**. 4. ed. Rio de Janeiro: FGV Editora, 2018.

CHRISTOPHER, Martin. **Logística e gerenciamento da cadeia de suprimentos: criando redes que agregam valor**. Trad. Mauro de Campos Silva. 2. ed. São Paulo: Thomson Learning, 2007.

CNT – Confederação Nacional do Transporte. **Anuário CNT do transporte: estatísticas consolidadas – 2018**. [S. l.]: CNT, 2018 Disponível em: <http://anuariodotransporte.cnt.org.br/2018/>. Acesso em: 30 mar. 2020.

CNT – Confederação Nacional do Transporte. **Boletim Estatístico – CNT – Fevereiro 2019**. [S.l.]: CNT, 2019. Disponível em: <https://www.cnt.org.br/boletins>. Acesso em: 20 fev. 2020.

CYPRIANO, L. A. *et al.* Formação do frete no Brasil: subsídios para estratégias de negociação em cadeias de suprimentos. In: CONGRESSO DA SOBER, 44., 2006, Fortaleza. **Anais...** Fortaleza: SOBER, 2006. Disponível em: <https://docplayer.com.br/35117942-Formacao-do-frete-no-brasil-subsidios-para-estrategias-de-negociacao-em-cadeias-de-suprimentos.html>. Acesso em: 30 jan. 2020.

DEMIR, Emrah *et al.* A selected review on the negative externalities of the freight transportation: Modeling and pricing. **Transportation Research Part E**, v. 77, p. 95-114, mar. 2015. <https://doi.org/10.1016/j.tre.2015.02.020>

FLEURY, P. F. **Supply Chain Management: conceitos, oportunidades e desafios da implementação**. Rio de Janeiro: UFRJ/COPPEAD, 2002.

FLEURY, P. F.; FIGUEIREDO, K. F.; WANKE, P. **Logística empresarial: a perspectiva brasileira**. São Paulo: Atlas, 2000. (Coleção COPPEAD de Administração).

FRIEDMAN, Jerome H. Multivariate Adaptive Regression Splines. **The Annals Statistics**, v. 19, n. 1, p. 1-67, mar. 1991. Disponível em: https://projecteuclid.org/download/pdf_1/euclid.aos/1176347963. Acesso em: 10 mar. 2020. <https://doi.org/10.1214/aos/1176347963>

HAO, Fei *et al.* An efficient pricing strategy of sensing tasks for crowdphotographing. **Journal of Supercomputings**, v. 75, p. 4443-4458, ago. 2019.
<https://doi.org/10.1007/s11227-019-02808-7>

HOLGUÍN-VERAS, José. The truth, the myths and the possible in freight road pricing in congested urban areas. **Procedia Social and Behavioral Sciences**, v. 2, p. 6366-6377, 2010. <https://doi.org/10.1016/j.sbspro.2010.04.045>

KIM, Jiowon *et al.* Valuing Recreational Beaches: A Spatial Hedonic Pricing Approach. **Coastal Management**, v. 48, p. 118-141, 2020.
<https://doi.org/10.1080/08920753.2020.1732799>

KING, David A.; GORDON, Cameron E.; PETERS, Jonathan R. Does road pricing affect port freight activity: Recent evidence from the port of New York and New Jersey. **Research in Transportation Economics**, v. 44, p. 2-11, 2014.
<https://doi.org/10.1016/j.retrec.2014.04.002>

MARTINEZ, D. L. *et al.* A Convex Version of Multivariate Adaptive Regression Splines. **Computational Statistics and Data Analysis**, v. 81, p. 89-106, jan. 2015.
<https://doi.org/10.1016/j.csda.2014.07.015>

MARTINS, Ricardo S. Estudo da formação do frete rodoviário e potencial de conflitos em negociações em cadeias do agronegócio brasileiro. **Organizações Rurais & Agroindustriais**, Lavras, v. 10, n. 1, p. 73-87, 2008. Disponível:
<http://revista.dae.ufla.br>. Acesso em: 10 fev. 2020.

MEDINA, Jovane M. **Cadeia de abastecimento no Comércio Eletrônico sob a ótica de redes flexíveis** – um método de estruturação. 2002. 289 f. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis, 2002.

MILAN, Gabriel S. *et al.* As estratégias de precificação e o desempenho das empresas. **REAd**, Porto Alegre, v. 84, n. 2, p. 419-453, maio/ago. 2016. Disponível em: <http://www.scielo.br/pdf/read/v22n2/1413-2311-read-22-02-00419.pdf>. Acesso em: 10 mar. 2020. <https://doi.org/10.1590/1413-2311.0982015.57273>

MONTGOMERY, Douglas C.; PECK, Elizabeth A.; VINING, G. Geoffrey. **Introduction to Linear Regression Analysis**. 5th ed. New York: John Wiley & Sons, 2012.

MOZAFARI, Marzieh; KARIMI, Behrooz. Pricing for freight carriers in a competitive environment: A game theory approach. **International Journal of Industrial Engineering Computations**, v. 2, p. 467-478, 2011.
<https://doi.org/10.5267/j.ijiec.2011.04.004>

ORTÚZAR, Juan D.; WILLUMSEN, Luis G. **Modelling Transport**. West Sussex, Inglaterra: John Wiley, 1994.

PEREIRA, Júlio C. ; GARSON, Salomão ; ARAÚJO, Elton G. Construção de um modelo para o preço de venda de casas residenciais na cidade de Sorocaba-SP. **Gestão da Produção, Operação e Sistemas**, Ano 7, n. 4, p. 153-167, dez. 2012.

PEREIRA, Luiz A. G.; LESSA, Simone N. O processo de planejamento e desenvolvimento do transporte rodoviário no Brasil. **Caminhos de Geografia**, Uberlândia, v. 12, n. 40, p. 26-46, dez. 2011. Disponível em: <http://www.seer.ufu.br/index.php/caminhosdegeografia/article/view/16414/9175R>. Acesso em: 10 mar. 2020.

PLATT, Allan A.; KLAES Luiz S. Utilizando o Sistema Integrado de Gestão (ERP) no apoio ao ensino de logística e gestão da cadeia de suprimentos. **Revista de Ciências da Administração**, v. 12, n. 28, p. 224-241, set./dez. 2010. Disponível em: <https://periodicos.ufsc.br/index.php/adm/article/view/2175-8077.2010v12n28p224/17432>. Acesso em: 10 mar. 2020. <https://doi.org/10.5007/17394>

RAUPP, Fabiano M.; BEUREN, Ilse M. Metodologia da pesquisa aplicável às ciências sociais. In: BEUREN, Ilse M. (Org.). **Como elaborar trabalhos monográficos em contabilidade: teoria e prática**. 3. ed. São Paulo: Atlas, 2006. Cap. 3, p. 76-97.

SOUZA, Paulo T. **Logística interna para empresas prestadoras de serviço**. Guia Log, 2002. Disponível em: <http://guialog.com.br/ARTIGO350.htm>. Acesso em: 10 mar. 2020.

WANG, Dan ; NICOLAU, Juan L. Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com. **International Journal of Hospitality Management**, v. 62, p. 120-131, 2017. <https://doi.org/10.1016/j.ijhm.2016.12.007>



Artigo recebido em: 17/04/2019 e aceito para publicação em: 05/12/2020
DOI: <http://dx.doi.org/10.14488/1676-1901.v20i4.4041>