

Investigação do desempenho do transporte ferroviário de cargas: um modelo de regressão com dados em painel

Investigation of the performance of rail freight: a regression model with panel data

Rendimiento de la investigación del transporte de cargas ferroviaria: un modelo de regresión con datos del panel

Recebido: 22/07/2019 | Revisado: 06/08/2019 | Aceito: 12/08/2019 | Publicado: 24/08/2019

Wallace Giovanni Rodrigues do Valle

ORCID: <https://orcid.org/0000-0003-4662-4129>

Universidade Federal do Rio Grande do Norte, Brasil

E-mail: wallacegrv@gmail.com

Resumo

O objetivo deste trabalho é investigar os fatores que afetam o desempenho do transporte ferroviário de cargas e averiguar a magnitude dos efeitos decorrentes. Para isso, elaborou-se um modelo de regressão linear múltipla com análise de dados em painel, considerando os efeitos fixos ao longo do tempo. Os dados coletados correspondem ao período de 2011 a 2018 e advêm da Agência Nacional de Transporte Terrestres (ANTT). As variáveis explicativas utilizadas são: velocidade, manutenções, acidentes e volume de carga (produção). Após formulação e execução do modelo, os índices obtidos foram testados quanto à significância estatística. Buscou-se atenuar erros heterocedásticos por meio do cálculo dos erros-padrão robustos e realizou-se um teste de especificação do modelo. Detectou-se que o volume de carga transportada e velocidade do trem possuem um impacto estatisticamente relevante sobre o desempenho. O modelo desenvolvido não apresentou evidência de má especificação e pode auxiliar no planejamento das atividades das empresas observadas.

Palavras-chave: Transporte ferroviário de carga; Análise de regressão; Dados em painel.

Abstract

The objective of this paper is to investigate the factors that affect the performance of rail freight and to ascertain the magnitude of the resulting effects. A multiple linear regression model was developed with panel data analysis, considering the fixed effects over time. The data collected correspond to the period from 2011 to 2018 and come from the National Land Transportation Agency (ANTT). The

explanatory variables used are: speed, maintenance, accidents and cargo volume (production). After formulating and executing the model, the obtained indices were tested for statistical significance. We attempted to attenuate heteroscedastic errors by calculating robust standard errors and performed a model specification test. It was detected that the volume of cargo transported and speed of the train have a statistically relevant impact on performance. The model developed showed no evidence of poor specification and can assist in the planning of the activities of the observed companies.

Keywords: Rail freight; Regression analysis; Panel data.

Resumen

El objetivo de este trabajo es investigar los factores que afectan el rendimiento del transporte ferroviario y determinar la magnitud de los efectos resultantes. Se desarrolló un modelo de regresión lineal múltiple con análisis de datos de panel, considerando los efectos fijos a lo largo del tiempo. Los datos recolectados corresponden al período de 2011 a 2018 y provienen de la Agencia Nacional de Transporte Terrestre (ANTT). Las variables explicativas utilizadas son: velocidad, mantenimiento, accidentes y volumen de carga (producción). Después de formular y ejecutar el modelo, los índices obtenidos se probaron para determinar su significación estadística. Tratamos de atenuar los errores heteroscedásticos calculando errores estándar robustos y realizamos una prueba de especificación del modelo. Se encontró que el volumen de carga transportada y la velocidad del tren tenían un impacto estadísticamente significativo en el rendimiento. El modelo desarrollado no presentó pruebas de una especificación deficiente y puede ayudar a planificar las actividades de las empresas observadas.

Palabras clave: Carga ferroviaria; Análisis de regresión; Datos del panel.

1. Introdução

O transporte de carga por meio de estradas de ferro tem sido preferível em alguns países devido às vantagens em diferentes perspectivas. Do ponto de vista ambiental, o desenvolvimento do transporte ferroviário ocasiona uma baixa emissão de gás carbônico (CO₂), mantendo uma capacidade de carga similar ou superior a outros modais de transporte. Ademais, em se tratando de riscos ao desempenho, as ferrovias estão livres de problemas como congestionamentos, além de oferecerem menor ameaça à segurança da carga (baixos índices de acidentes graves). Logo, a velocidade média de entrega é elevada, também devido a própria estrutura do veículo automotor (locomotiva).

De acordo com a Associação Nacional dos Transportadores Ferroviários – ANTF (2019), esse conjunto de benefícios faz com que países com grandes extensões territoriais optem por ferrovias para transporte de carga, é o caso da China (37% do transporte de carga),

Estados Unidos (43%), Canadá (46%) e Rússia (81%). No Brasil, apesar das vantagens, o transporte ferroviário ainda não possui grande participação (cerca de 15%) na matriz de transporte de carga, sendo o segundo modal mais utilizado (perdendo para o rodoviário, extremamente caro, mas detentor de 65% da matriz). Dentre todas as cargas transportadas no país, as ferrovias são responsáveis por uma parcela considerável dos grãos sólidos agrícolas exportados (mais de 40%) e mais de 95% dos minérios chegam aos portos pelos trilhos (ANTF, 2019).

Dada a subutilização do potencial do transporte ferroviário, ao longo do tempo, alguns estudos vêm avaliando e evidenciando os benefícios e desempenho desse modal. Dentre as técnicas utilizadas, destacam-se as paramétricas e não-paramétricas. No grupo de técnicas não-paramétricas sobressai-se a Análise Envoltória de Dados (Yu, 2008, Luan & Zhang, 2012, Domernik, 2015, Zhou & Hu, 2017), enquanto a vertente paramétrica normalmente apresenta modelos de Regressão Linear Múltipla (Benishay & Whitaker Jr., 1966, Yang, 2015, Kecman & Goverde, 2015, Wen et al., 2017) como principais solucionadores de problemas.

Yang (2015) enfatiza que os modelos construídos com base na análise de regressão possuem a vantagem de poder ser usados na análise de vários problemas de transporte de carga. A escolha da ferramenta de regressão também possibilita uma análise estatística mais profunda das variáveis envolvidas (individualmente e em conjunto) e do modelo como um todo. A variável dependente (objeto de investigação) pode ser explicada por diversos fatores. Dessa forma, torna-se viável o desígnio de variáveis explicativas (independentes) em uma função produção com base em estudos anteriores, especialistas na área de interesse e até testes de especificação do modelo desenvolvido.

No caso do transporte ferroviário, observa-se que os principais objetos de investigação são variáveis dependentes relacionadas a: o tempo de viagem (Azambuja, 1995), o índice de acidentes (Austin & Carson, 2002, McCollister & Pflaum, 2007), o desenvolvimento econômico regional gerado (Da Silva, Jaime Jr & Martins, 2009, Reiter, 2015), o volume de carga (Fernandes, 2011), o tempo de parada do trem (Kecman & Goverde, 2015), o atraso (Wen et al., 2017) e a distância percorrida (Wang et al., 2019).

De acordo com Azambuja (1995), a alocação de recursos no setor de transportes ocorre basicamente de acordo com uma análise de custo-benefício, o que faz com que a verificação dos tempos de movimentação da carga seja um aspecto primordial na tomada de decisão. Isso ocorre porque todo modal possui fatores independentes que influenciam o tempo de viagem da mercadoria, sendo essa uma variável de medida comum a todos, mas afetada de

diferentes maneiras em cada tipo de transporte.

Tendo isso em vista, pode-se alcançar melhorias nesse setor por meio de uma análise simultânea das reduções nos tempos de viagem e aumento da confiabilidade dos serviços ofertados (Azambuja, 1995). Pereira (2009) corrobora com essa ideia ao afirmar que a seleção entre o modal de transporte deve se fundamentar em impactos do serviço e custos logísticos, isso inclui aspectos de nível de serviço como o tempo de transporte, consistência no tempo de entrega (conforme acordo), frequência e disponibilidade/flexibilidade.

À luz dessas ponderações, o objetivo deste trabalho é investigar os fatores que afetam o desempenho do transporte ferroviário de cargas, ao longo do tempo, e analisar os efeitos decorrentes. Essa variável é mensurada, em horas em movimento, pela Agência Nacional de Transportes Terrestres (ANTT), juntamente com outros fatores como: velocidade média do trem, índice de acidentes, manutenções realizadas, volume de carga transportado e distância percorrida. Portanto, este estudo também tem o intuito de analisar quais variáveis podem ajudar a explicar o desempenho obtido.

Nacionalmente, ainda não se tem registro de uma pesquisa investigativa considerando essas variáveis na área de transportes, principalmente tratando do segmento ferroviário ao longo do tempo, longitudinalmente. Espera-se, também, que o presente trabalho contribua com a intensificação do uso do transporte ferroviário de carga, destacando-o como uma boa alternativa a partir dos fatores que influenciam o desempenho desse modal.

O presente trabalho estrutura-se em cinco tópicos principais, incluindo a introdução ao objeto de pesquisa que aqui se conclui. O tópico 2 abrange o estudo do transporte ferroviário de carga e análise de regressão, esclarecendo as principais premissas envolvidas. Posteriormente, o tópico 3 estabelece o método adotado, elucidando todas as etapas de construção do conteúdo para alcance dos objetivos. Os resultados alcançados serão expostos e discutidos no tópico 4, seguido das implicações gerais e interpretação da pesquisa presentes no tópico 5.

2. Transporte ferroviário de cargas e análise de regressão

Em países desenvolvidos, o interesse de numerosas cidades grandes em sistemas ferroviários está conectado ao objetivo de promover o desenvolvimento de atividades econômicas (Wang et al., 2019). O uso de transporte de cargas por ferrovias é considerado o nível mais básico da vida social em algumas regiões, e o desenvolvimento desse tipo de transporte está fortemente correlacionado com elementos que motivam desempenho

econômico regional (He et al., 2004, Nuzzolo et al., 2013).

Fatores como aumento de produção, reestruturação econômica e estilo de crescimento econômico podem ser reflexos da capacidade dos sistemas ferroviários para condução de cargas (Yang, 2015). Isso porque essa forma de transporte consiste, essencialmente, em um serviço relativamente homogêneo, composto por demanda econômica externa e sistema de abastecimento de transporte (Benishay & Whitaker Jr., 1966, Zhao et al., 2004).

Por esse e outros motivos, na engenharia de transporte, estudos de previsão e impacto são a base para a formulação de políticas relevantes, preparando os programas de desenvolvimento de transporte e planejando o gerenciamento (Smith, 1974, Jiang & Yang, 2002). Nesse contexto, os métodos de análise de regressão são primordiais para análise de efeitos e previsão de curto e médio prazo, com possibilidade de aplicação em muitos campos de pesquisa (Yang, 2015).

No entanto, para uma avaliação mais consistente, é preciso observar os efeitos, causados ao objeto de estudo, ao longo de um determinado período. Desde a pesquisa de Benishay e Whitaker Jr (1966), há uma percepção de que a correspondência é alta entre os resultados das regressões empregando tempo para eliminar o efeito de tendência comum. Nessa perspectiva, o uso de dados em painel ajuda a obter estimativas mais eficientes, pois considera conjuntamente o tempo e as dimensões da seção transversal do conjunto de dados (Rosa et al., 2019).

Um conjunto de dados longitudinal, ou painel, é aquele que segue uma dada amostra de indivíduos ao longo do tempo e, portanto, fornece múltiplas observações sobre cada indivíduo na amostra (Hsiao, 2003). A dimensão do tempo, segundo Lončar et al. (2019), não pode ser muito pequena, uma vez que a questão da heterogeneidade pode não ser resolvida. A heterogeneidade de parâmetros (na especificação do modelo) consiste em especificar e estimar os efeitos individuais e/ou específicos de tempo que existem entre unidades seccionais ou séries temporais, mas não são capturados pelas variáveis explicativas incluídas (Ibid.). Box (1979) já indicava que testes de robustez formam um importante elemento de inferência estatística em dados observacionais.

No contexto de modelos de regressão linear, outros testes relevantes são o de determinação da forma linear desenvolvida. O teste de especificação geral mais utilizado é o Teste de Erro de Especificação de Regressão de Ramsey (1969) – RESET (Ramalho & Ramalho, 2012). O teste RESET emergiu como um teste geral de erro de especificação, o qual é projetado para detectar variáveis omitidas e forma funcional inadequada em modelos de equação única (Ramsey, 1969, Mantalos & Shukur, 2007).

3. Metodologia

Com base na delimitação do objeto de estudo, este trabalho executa as seguintes fases: (a) pesquisa bibliográfica; (b) coleta de dados; (c) estruturação do modelo; e (d) análise de dados e modelo. A natureza deste trabalho é quantitativa, uma vez que, como consideram Pereira et al. (2018), há o emprego de valores, números, porcentagens e métodos estatísticos para solução de problemas. Quanto à finalidade, esta pesquisa é aplicada, pois visa gerar conhecimentos para aplicação prática, concentrando-se em torno dos problemas presentes nas atividades das instituições, organizações, grupos ou atores sociais (Gerhardt & Silveira, 2009, Fleury & Werlang, 2017).

A busca na literatura teve o propósito de encontrar referências teóricas de regressão linear múltipla aplicada ao transporte ferroviário de cargas. As informações observadas inicialmente foram as variáveis adotadas e técnica de análise. Posteriormente, coletou-se os dados a partir do Anuário Estatístico presente no endereço eletrônico da ANTT. A amostra a ser utilizada continha 13 concessionárias autorizadas a prestar o serviço, mas somente 11 possuíam dados suficientes para este estudo. Isso porque notou-se que o histórico disponível, de 2006 a 2018, não continha determinados dados - como manutenção e acidentes - em alguns casos mais antigos. Logo, reduziu-se o período para 2011 a 2018, quando o controle estava padronizado para a maioria das observações da amostra.

As variáveis apresentadas no Anuário se referem aos trens formados, gerais, às locomotivas (veículo automotor responsável pela condução) e aos vagões (acoplados à locomotiva e onde estão as cargas). As principais identificadas estão detalhadas no Quadro 1.

Quadro 1 – Variáveis coletadas para as 12 empresas de 2010 a 2018

	Variáveis	Unidade de medida
Gerais	Índice de acidentes	Acidentes por milhão de trem.km
	Velocidade Média de Percurso	Km/h
	Movimentação de trens	Trem.km
Locomotiva	Desempenho da frota	Horas
	Disponibilidade e utilização	Porcentagem (%)
	Consumo combustível	L/milhões de TKU
	Percurso	Km
Vagões	Desempenho	Horas
	Disponibilidade e utilização	Porcentagem (%)
	Manutenção em vagões	Quantidade
	Percurso vagão	Km
	Produção (carga) por vagão	Tonelada útil (TU) e Tonelada Quilômetro útil (TKU)

Fonte: elaborado pelo autor.

Por ainda ser uma área pouco explorada, muitas dessas variáveis não foram utilizadas anteriormente. Dentre as variáveis da amostra, foram analisadas (como variável dependente ou independente) apenas: desempenho (Azambuja, 1995), índice de acidentes (Austin & Carson, 2002, Mccollister & Pflaum, 2007), produção ou volume de carga (Fernandes, 2011) e distância percorrida (Wang et al., 2019).

Tendo em vista esses fatores e avaliando empiricamente quais outros da amostra podem afetar/explicar o desempenho dos vagões (DV), inicialmente, determina-se as seguintes variáveis independentes (ou explicativas): índice de acidentes (ACDNT), velocidade média de percurso (VMP), manutenção em vagões (MV) e produção (volume de carga em TKU) por vagão (PTKUV). A unidade de medida da produção foi escolhida a partir da literatura, onde verificou-se que estudos como os de Fernandes (2011) e FitzRoy e Smith (1995) mensuraram o volume de carga em tonelada por quilômetro útil (TKU), em vez de somente tonelada útil (TU). Isso também elimina a necessidade de mais uma variável para captar o efeito da distância percorrida isoladamente.

As variáveis de produção e manutenção, assim como a própria variável a ser explicada (desempenho), são referentes aos vagões porque é nesse espaço que são armazenadas as cargas (objeto de interesse do transporte). No entanto, velocidade média do percurso e índice de acidentes são incluídos devido ao impacto que podem causar no tempo gasto para entrega.

O modelo composto neste estudo será avaliado e investigado por meio de dados em painel. Essa ferramenta permite uma análise abrangendo vários períodos, mantendo efeitos fixos ao longo do tempo (representados por α_i). Dessa forma, o modelo é capaz de abordar cada observação (i) da amostra (n), em cada período de tempo (t). De acordo com Da Silva, Jayme Jr e Martins (2009), uma das vantagens da estimação com dados em painel é a consideração da heterogeneidade individual das observações. O modelo teórico elaborado está representado na Equação 1.

$$\log(DV) = \beta_0 + \beta_1 ACDNT_{it} + \beta_2 VMP_{it} + \beta_3 VMP_{it}^2 + \beta_4 \log(MV) + \beta_5 \log(PTKU_{it}) + \alpha_i + u_{it} \quad (1)$$

O termo de erro u_{it} representa outros fatores não observados (erros idiossincráticos). Os coeficientes β correspondem às inclinações das retas, com exceção de β_0 que consiste no intercepto da regressão.

A variável VMP foi acompanhada do seu termo quadrático porque entende-se que o excesso de velocidade pode causar acidentes e até aumentar o número de manutenções. Sendo assim, a partir de um dado momento o aumento da velocidade não mais diminuiria o tempo de transporte, mas o aumentaria devido aos problemas decorrentes. Em relação às variáveis inseridas em logaritmos, compreende-se que nessa forma funcional há uma melhor interpretação dos resultados (em porcentagem), uma vez que tratam de unidades difíceis de idealizar o significado prático, quando não se tem familiaridade com a área de transporte ferroviário de cargas.

A verificação do modelo e dados é realizada em etapas, as quais consistem em: (a) estatística descritiva da amostra; (b) erro-padrão robusto para heterocedasticidade; (c) teste de significância estatística das variáveis explicativas e; (d) teste RESET para revelar se há má especificação de variáveis no modelo. Todo o modelo, testes e análises estatísticas foram desenvolvidos no software *Stata 13.0*.

A estatística descritiva consiste em especificar as médias amostrais, assim como os valores máximos e mínimos coletados, além do desvio-padrão da amostra. O intuito dessa primeira etapa é, basicamente, promover o entendimento da amplitude e densidade da amostra. O erro-padrão robusto tem a função de atenuar a heterocedasticidade, se o valor da estatística F calculado na forma usual for muito diferente do calculado na forma robusta, há sinais de que o modelo não apresenta variância constante.

O teste de significância estatística individual das variáveis explicativas escolhidas é executado por meio do teste t-student (teste t), o qual apresenta valores tabelados de acordo com as características da amostra. Se o valor calculado for maior que o tabelado, rejeita-se a hipótese nula de que a variável em questão é estatisticamente não significativa. Além disso, será efetivada uma avaliação conjunta dessas variáveis, por intermédio do teste F. Se o p-valor advindo do teste F for menor do que o nível de significância estabelecido (normalmente 5%), rejeita-se a hipótese nula.

O teste de erro de especificação de regressão de Ramsey (RESET) tem o objetivo de testar a forma funcional do modelo. Primeiramente, inclui-se termos de ordem mais elevada na equação de regressão da variável dependente. Segundamente, realiza-se um teste da estatística F para exclusão desses termos inseridos, observando o p-valor. Determina-se, então, uma hipótese nula de que os coeficientes estimados para os termos incluídos são iguais a zero. Se tal hipótese for rejeitada, há evidência de relevância dos termos inseridos. Logo, há evidência de que o modelo está mal especificado.

4. Resultados

Inicialmente, após serem plotados os dados e geradas as variáveis em logaritmo e termo quadrático, obteve-se a estatística descritiva da amostra. Tendo em vista o período de tempo analisado (8 anos), pode-se obter as informações estatísticas de maneira global (2011-2018), entre empresas e considerando uma mesma empresa ao longo do tempo, conforme apresentado na Tabela 1.

Tabela 1 – Estatística descritiva da amostra em painel

Variável	Modo de análise	Média	Desvio-padrão	Mínimo	Máximo	Observações
log(dv)	Total	8.35	1.53	4.16	9.95	N = 88
	Entre empresas		1.55	5.60	9.87	n = 11
	Empresa no tempo		0.36	6.91	9.24	T = 8
acdnt	Total	19.74	27.05	1.81	196.43	N = 88
	Entre empresas		23.96	2.55	87.56	n = 11
	Empresa no tempo		14.27	-20.05	128.61	T = 8
vmp	Total	23.22	10.41	10.92	48.53	N = 88
	Entre empresas		10.27	12.39	40.05	n = 11
	Empresa no tempo		3.38	8.84	31.69	T = 8
log(mv)	Total	8.55	1.79	2.40	10.59	N = 88
	Entre empresas		1.81	4.30	10.19	n = 11
	Empresa no tempo		0.47	6.66	10.27	T = 8
log(ptkuv)	Total	14.18	1.01	12.18	16.02	N = 88
	Entre empresas		0.98	12.74	15.82	n = 11
	Empresa no tempo		0.37	13.06	15.92	T = 8

Fonte: elaborado pelo autor.

Observando a Tabela 1, nota-se que a variável ACDNT é a que apresenta maior variação e inconsistência, principalmente em se tratando dos valores mínimos e máximos das observações – tanto de modo geral, quanto entre empresas e cada empresa ao longo do tempo. O modelo executado de forma usual evidenciou as consequências da natureza dessa variável, sendo esta a única estatisticamente insignificante (teste t calculado de 1.12, inferior ao tabelado com 95% de confiança). O mesmo modelo retornou um teste F total no valor de 69.25, R^2 total de 16.65%, R^2 entre empresas de 30.28% e R^2 ao longo do tempo de 82.79%, para um total de 88 observações.

Para diminuir possíveis imprecisões decorrentes de heterocedasticidade, mensurou-se os erros-padrão robustos, o que culminou em novos valores para os testes estatísticos, conforme Tabela 2.

Tabela 2 - Resultados do modelo considerando efeitos fixos no tempo (dados em painel)

ldv	Coefficiente	Erro-padrão usual	Erro-padrão robusto	t-student	P > t
acdnt	0.0014304	0.0012771	0.000804	1.78	0.106
vmp	-0.0026701	0.0391869	0.0469332	-0.06	0.956
vmp2	-0.0001798	0.0005373	0.0006366	-0.28	0.783
lmv	0.123435	0.042852	0.1069314	1.15	0.275
lptkuv	-0.8078335	0.0537656	0.066615	-12.13	0.000
Intercepto	18.90226	1.16937	1.83256	10.31	0.000

Fonte: elaborado pelo autor.

A diferença de valor entre a estatística F usual (69.25) e a do modelo robusto (102.73) indica que havia possibilidade de heterocedasticidade. Apesar de maior que o anterior, o teste t para variável ACDNT continuou demonstrando não significância estatística. O teste F conjunto para as variáveis VMP e VPM² apontou um p-valor de 0.0019, sendo essas variáveis estatisticamente significantes para o modelo. O logaritmo das manutenções realizadas nos vagões passaram a ser estatisticamente insignificantes (1.15, inferior ao t calculado com 95% de confiança), ao contrário do logaritmo da produção nos vagões, com maior significância estatística.

O efeito do logaritmo de produção no desempenho dos vagões é o que mais se destaca. A interpretação é de que 1% a mais no volume de carga reduz o tempo de entrega em 0.81%. Isso pode ser decorrente número de viagens para transporte de uma determinada carga, interferência de outro modal de transporte no caminho (multimodal), ausência de uma especificação correta para o peso por vagão ou até de fiscalização quanto a ocupação do vagão. Logo, seria necessário um estudo mais aprofundado mapear o impacto causado com maior precisão.

Por outro lado, o aumento de 1% no logaritmo de manutenção dos vagões aumenta em 0.12% o tempo gasto no transporte. O aumento unitário do índice de acidentes (em acidentes por milhão de trem.km) também contribui com o aumento do tempo de viagem, em 0.14%. O efeito da variável VMP, devido a forma funcional quadrática, pode ser obtido por meio de derivada, conforme Equação 2. Dessa forma, uma velocidade de 50 km/h ocasionaria uma redução de apenas 2.1%.

$$\frac{\partial \log(\widehat{dv})}{\partial vmp} = -0,0027 - 2 * 0,00018 vmp \quad (2)$$

Todos os efeitos das variáveis explicativas sobre o desempenho são verificados em média e considerando as demais variáveis fixas/constantas. O teste de especificação da forma funcional, conforme indicado no método, revelou que não há indícios de má formulação do

modelo. Após inclusão dos termos de ordem maior, realizou um teste F que resultou em um p-valor igual a 0.0703 (maior que o nível de significância a 5%). Portanto, os termos inseridos poderiam ser excluídos da regressão (não são relevantes para o modelo) – não se rejeita a hipótese nula.

5. Considerações finais

Este trabalho elaborou um modelo de regressão linear múltipla para investigar e analisar o impacto de alguns fatores sobre o desempenho do transporte de carga nos vagões. Após executado, verificou-se que a equação desenvolvida representa uma forma razoável de verificar o que pode influenciar o desempenho, mas não pode ser usada para explicar esse desempenho por completo, dada a magnitude dos valores encontrados.

Quanto à técnica utilizada, a análise de dados em painel auxiliou na incorporação de um vasto histórico de períodos de tempo, ao passo que viabilizou uma avaliação de natureza dinâmica. A consideração dos efeitos fixos aumentou a precisão e a certeza das informações geradas a partir dos dados. Portanto, as variações e tendências naturais já são observadas no modelo, facilitando a elaboração de planos de ação mais concisos e verossímeis.

A partir dos resultados obtidos, também é possível as empresas ajustarem a operação de transporte (incluindo volume de carga transportado) e o planejamento de manutenções que, embora baixo, demonstrou impactar o desempenho. Isso ocorre porque boa parte das manutenções são corretivas, demandando um maior tempo de parada da frota para realização dos ajustes. O ideal é que as manutenções preventivas sejam suficientes para evitar ao máximo paradas desnecessárias.

O uso da velocidade e a respectiva forma quadrática, conjuntamente, se mostrou relevante, assim como o logaritmo do volume de carga nos vagões (produção). Isso sugere que outros modelos podem melhorar seus resultados incorporando essas variáveis. No entanto, para modelagens futuras, recomenda-se uma análise extensa do comportamento e posicionamento das variáveis explicativas. Para isso, é indicado o acompanhamento da operação de transporte presencialmente ou uma investigação em outras bases de dados, a fim de identificar e testar demais fatores que possam afetar o tempo de condução da carga.

Como a literatura nessa temática é escassa, recomenda-se que novos modelos formulados também sejam submetidos a mais testes de especificação, promovendo uma investigação densa acerca das possíveis combinações de variáveis a utilizar. Esse processo, realizado de forma consistente, deve ser capaz de evitar vieses e atenuar a chance de se obter

erros heterocedásticos.

Referências

Associação Nacional do Transportadores Ferroviários – ANTF (2019). Informações gerais. Retirado em 24 junho, de www.antf.org.br/informacoes-gerais/.

Agência Nacional de Transportes Terrestres – ANTT (2019). Anuário estatístico. Retirado em 29 de junho, de www.antt.gov.br/ferrovias/arquivos/Anuario_Estatistico.html.

Austin, R. D., & Carson, J. L. (2002). An alternative accident prediction model for highway-rail interfaces. *Accident Analysis & Prevention*, 34(1), 31-42.

Azambuja, A. M. V. (1995). *Estimação de modelos comportamentais utilizando a técnica de preferência declarada: o caso de variabilidade dos tempos de viagem no transporte de grãos no Rio Grande do Sul*. 1995. Dissertação (Programa de Pós-Graduação em – PPGEP) – Escola de engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Benishay, H., & Whitaker, G. R. (1966). Demand and supply in freight transportation. *The Journal of Industrial Economics*, 243-262.

da Silva, G. J. C., Jayme Jr, F. G., & Martins, R. S. (2009). Gasto público com infraestrutura de transporte e crescimento: uma análise para os estados brasileiros (1986-2003) 1. *Revista Economia & Tecnologia*, 5(1).

Fernandes, R. (2011). *A procura de transporte ferroviário de mercadorias na Europa*. 2011. Dissertação (Mestrado em economia) – Faculdade de economia, Universidade de Coimbra, Coimbra.

Fleury, M. T. L., & da Costa Werlang, S. R. (2017). Pesquisa aplicada: conceitos e abordagens. Anuário de Pesquisa GVPesquisa.

FitzRoy, F., & Smith, I. (1995). The demand for rail transport in European countries. *Transport Policy*, 2(3), 153-158.

Gerhardt, T. E., & Silveira, D. T. (2009). Métodos de pesquisa. [e-book]. Porto Alegre, Editora da UFRGS. Retirado em 10 de agosto, de www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf.

González, R. M., Marrero, G. A., Rodríguez-López, J., & Marrero, Á. S. (2019). Analyzing CO2 emissions from passenger cars in Europe: A dynamic panel data approach. *Energy policy*, 129, 1271-1281.

Gu, S., & Lu, X. (2015, July). Analysis of China railway passenger volume's influence factors based on principal component regression. In 2015 International Conference on Logistics, Informatics and Service Sciences (LISS) (pp. 1-5). IEEE.

- He, Z. G., Shuai, B., & Liao, W. (2004). Freight Demand Forecasting for Logistics Centers [J]. *Logistics Technology*, 1.
- Jiang, H. Y., & Yang, D. M. (2002). The forecast methods of volume of water freight for comparison. *Oper. Res. Manag. Sci*, 11(3), 74-79.
- Kecman, P., & Goverde, R. M. (2015). Predictive modelling of running and dwell times in railway traffic. *Public Transport*, 7(3), 295-319.
- Lončar, D., Paunković, J., Jovanović, V., & Krstić, V. (2019). Environmental and social responsibility of companies cross EU countries–Panel data analysis. *Science of The Total Environment*, 657, 287-296.
- Mantalos, P., & Shukur, G. (2007). The robustness of the reset test to non-normal error terms. *Computational Economics*, 30(4), 393-408.
- Nuzzolo, A., Coppola, P., & Comi, A. (2013). Freight transport modeling: review and future challenges. *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti*, 151-181.
- McCollister, G. M., & Pflaum, C. C. (2007). A model to predict the probability of highway rail crossing accidents. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of rail and rapid transit*, 221(3), 321-329.
- Pereira, O. C. (2009). Soluções de otimização da eficiência energética de uma ferrovia de carga. Dissertação (Mestrado em Engenharia Industrial) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.
- Pereira, A.S. et al. (2018). Metodologia da pesquisa científica. [e-book]. Santa Maria/RS, Ed. UAB/NTE/UFSM. Retirado em 6 de agosto, de repositorio.ufsm.br/bitstream/handle/1/15824/Lic_Computacao_Metodologia-Pesquisa-Cientifica.pdf?sequence=1.
- Ramalho, E. A., & Ramalho, J. J. (2012). Alternative versions of the RESET test for binary response index models: a comparative study. *Oxford bulletin of economics and statistics*, 74(1), 107-130.
- Reiter, G. R. (2015). *Infraestrutura de transportes no Brasil: uma análise com dados em painel no período de 1995 a 2008*. 2015. Trabalho de conclusão (Ciências Econômicas), Universidade Federal do Rio Grande do Sul, Porto Alegre.
- Smith, P. L. (1974). Forecasting freight transport demand–The State of the Art. *Logistics and Transportation Review*, 10(4), 311-326.
- Wang, X., Tong, D., Gao, J., & Chen, Y. (2019). The reshaping of land development density through rail transit: The stories of central areas vs. suburbs in Shenzhen, China. *Cities*, 89, 35-45.

Wen, C., Lessan, J., Fu, L., Huang, P., & Jiang, C. (2017, August). Data-driven models for predicting delay recovery in high-speed rail. In *2017 4th International Conference on Transportation Information and Safety (ICTIS)* (pp. 144-151). IEEE.

Yang, Y. (2015). Development of the regional freight transportation demand prediction models based on the regression analysis methods. *Neurocomputing*, *158*, 42-47.

Zhao, C., Liu, K., & LI, D. S. (2004). Research on application of the support vector machine in freight volume forecast. *Journal of the China Railway society*, *26*(4), 10-14.

Porcentagem de contribuição de cada autor no manuscrito

Wallace Giovanni Rodrigues do Valle – 100%