

REVISÃO SISTEMÁTICA SOBRE A UTILIZAÇÃO DE DADOS DA TELEMÁTICA EM PESQUISAS NA ENGENHARIA MECÂNICA

Eng^o Mecânico Fábio Ribeiro von Glehn, fabioglehn@discente.ufg.br¹
Prof. Doutor João Paulo da Silva Fonseca, jpsfonseca.ufg.br¹
Prof. Doutor Marlipe Garcia Fagundes Neto, marlipe@ufg.br¹

¹Universidade Federal de Goiás (UFG), Programa de Pós-Graduação em Engenharia Mecânica (PPGMEC) – Av. Ingá, Quadra B, s/n, Campus Samambaia, CEP: 74.690-900 – Goiânia – Goiás – Brasil

Resumo. *Este trabalho teve como objetivo buscar artigos que poderiam ajudar a solucionar problemas de consumo de combustível e emissões de poluentes em caminhões rodoviários, por meio da análise de dados obtidos por telemática veicular. A pesquisa realizada nas bases Scopus e Web of Science (WoS), usou a intersecção de quatro grandes conjuntos: caminhões, economia de combustível, telemática e aprendizado de máquina. Os resultados foram analisados por programas bibliométricos para selecionar trinta artigos relevantes. A telemática e a análise de seus dados podem oferecer soluções para problemas de manutenção, economia de combustível e assistência à direção.*

Palavras-chave: Caminhão. Telemática. Aprendizado de máquina. Economia de combustível. Manutenção.

Abstract. *This work aimed to search for articles that could help solve fuel consumption and pollutant emissions problems in highway trucks through the analysis of data obtained by vehicle telematics. The search was conducted on Scopus and Web of Science (WoS) databases, using the intersection of four major sets: trucks, fuel economy, telematics, and machine learning. The results were analyzed by bibliometric programs to select thirty relevant articles. Telematics and the analysis of its data can provide solutions to maintenance, fuel economy, and driving assistance problems.*

Keywords: Truck, Telematic, Machine Learning, Fuel save, Maintenance.

1. INTRODUÇÃO

Existe uma dificuldade em construir um conjunto de dados ideal para o treinamento de modelos de diagnóstico inteligentes, conforme mencionado por Zhang *et al.* (2022). Os autores destacam que, no contexto da engenharia, as máquinas geralmente operam em condições normais, com falhas raras, o que dificulta a obtenção de dados de falhas suficientes para treinar modelos de diagnóstico. Além disso, experimentos em laboratórios podem ser caros e a simulação computacional pode não ser prática o suficiente para obter dados de falhas.

Uma solução proposta é o uso de telemática veicular, que permite a captura de dados de funcionamento dos veículos em tempo real e em grande quantidade. Isso implica em um desafio para a aplicação de conhecimentos científicos na análise desses dados. As abordagens relacionadas a *big data*, inteligência artificial e *machine learning* têm sido cada vez mais utilizadas, conforme demonstrado pela crescente quantidade de publicações nos últimos anos, especialmente em *machine learning*.

Para reduzir o escopo do estudo e lidar com a grande diversidade de máquinas na Engenharia Mecânica, optou-se por focar na revisão sistemática que envolve a aquisição de dados pela telemática veicular em caminhões rodoviários, com o objetivo de reduzir o consumo de combustível e realizar diagnósticos inteligentes de falhas. Essa escolha se baseia na pesquisa de EPE and AIE (2021), que aponta que os caminhões pesados são responsáveis por cerca de 60% da energia consumida no transporte de cargas no Brasil.

A telemática fornecerá um enorme volume de dados de funcionamento dos caminhões em que o equipamento telemático estiver instalado. Esses equipamentos estão disponíveis para aquisição a um preço acessível, o que se traduz em vantagem da não necessidade de um laboratório físico nas instalações da universidade, visto que o dado é coletado em campo e transmitido e armazenado em nuvem pelo fabricante do equipamento. O trabalho consistirá em construir uma *Application Programming Interface* (API) e buscar nos dados coletados as respostas aos problemas que se deseja estudar.

Desta forma, muito embora a redução do escopo do estudo limite a generalização dos resultados para outros tipos de máquinas, a resolução do problema da raridade de dados de falhas pode permitir a aplicação da abordagem em outras máquinas, além de contribuir para o entendimento do consumo de energia de forma mais ampla.

2. METODOLOGIA

A metodologia proposta consiste em responder a seguinte pergunta: 'Como a utilização de dados da telemática tem sido usada em pesquisas relacionadas à Engenharia Mecânica, em especial em caminhões rodoviários?' Para coletar os dados necessários, seriam utilizadas as bases Scopus - Elsevier (2022) e WoS - Clarivate (2022), e para o tratamento dos dados seriam realizados por meio das ferramentas Bibloshiny de Aria and Cuccurullo (2017) e pybibx de Pereira (2022).

Inicialmente foram criados quatro grupos para organizar as palavras-chave utilizadas na busca, sendo eles:

- A) a região limitante do problema;
- B) o problema que se deseja estudar;
- C) como os dados seriam adquiridos; e
- D) as técnicas de análise dos dados.

A "Fig. 1" ilustra a interação entre esses quatro grupos, e o objetivo é encontrar o conjunto de documentos que corresponda à intersecção desses grupos.

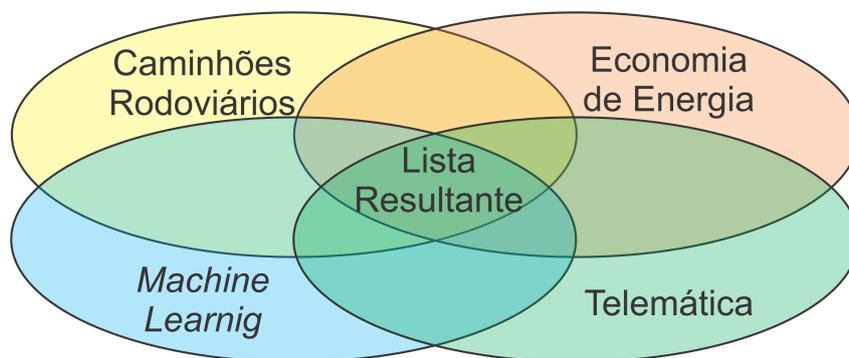


Figura 1. Diagrama de Venn dos conjuntos de palavras-chave utilizados na busca de artigos

Após a definição dos grupos de palavras, a busca pode resultar em alguns artigos em duplicidade devido a inconsistências na indexação, tais como a identificação do *Digital Object Identifier* (DOI) em apenas uma das bases. Essas inconsistências devem ser identificadas e removidas. Além disso, é necessário identificar e contar como apenas um os artigos que aparecerem em mais de uma base.

Após essa organização, os artigos deverão ser classificados como elegíveis ou não. Entradas do tipo *Conference Review* em bases de dados bibliográficas como Scopus e WoS são índices que listam os artigos publicados em uma determinada conferência e, muitas vezes, trazem uma breve descrição do evento em si. Ou seja, essas entradas não se referem a um artigo específico, mas sim a uma lista de artigos publicados em um determinado evento. Portanto, ao pesquisar em bases de dados, é importante verificar o tipo de entrada e o que ela representa, para evitar confusões na interpretação dos resultados.

2.1 As palavras de formação de cada conjunto e o tamanho de cada um deles

Foram definidas as seguintes palavras-chave para cada grupo, e realizadas buscas nas bases Scopus e WoS, agrupadas por *Article title*, *Abstract*, *Keywords*. Os resultados obtidos em 03/01/2023 são apresentados na "Tab.1". O Scopus retornou uma lista com 100 resultados e a WoS, 45.

Tabela 1. Palavras de busca para formação de cada conjunto e quantidade de documentos obtida em cada base

Grupo	Palavras	Scopus	WoS
A	truck* OR heavy-duty OR vehicle* OR fleet	1.090.198	601.203
B	"fuel consumption"OR maintenance OR "mass detection"OR "fault detection"OR "intelligent fault diagnosis"	950.798	583.678
C	telematic* OR obd OR few-shot OR (sample AND (small OR few OR limited OR unbalance OR imbalance*))	780.436	1.004.776
D	"machine learning"OR "deep learning"OR "neural network"OR "artificial intelligence"	1.532.583	702.675
$A \cap B \cap C \cap D$	#A AND #B AND #C AND #D	100	45

2.2 Análise bibliométrica, unificação dos dados e remoção de duplicatas

Para realizar a análise bibliométrica, utilizamos o Biblioshiny, uma ferramenta de código aberto para pesquisa quantitativa em cienciometria e bibliometria. Com o auxílio dessa ferramenta, foram obtidos resultados do Scopus e WoS individualmente, que foram utilizados para gerar mais de uma dezena de gráficos, mapas e tabelas. As “Fig. 2 e 3” exemplificam alguns dos resultados obtidos.

O trabalho de unificação das bases dos dados foi realizada com o pybibx e o conjunto se solidificou conforme a “Tab. 2”.

Tanto Biblioshiny quanto o pyBibx fornecem uma lista das palavras mais frequentes. A “Tab. 3” mostra o resultado.

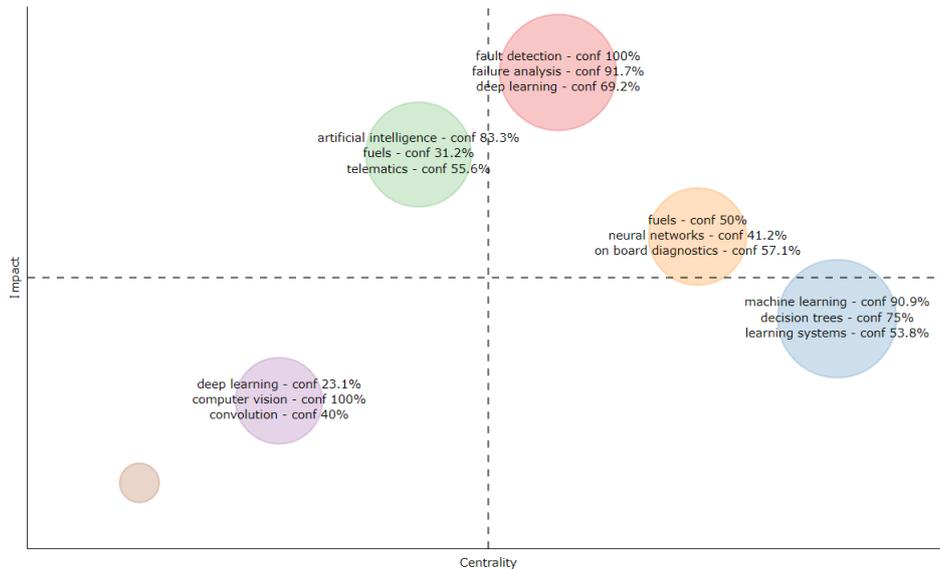


Figura 2. Clustering by Coupling - Measured by Keywords Plus do arquivo oriundo da base Scopus

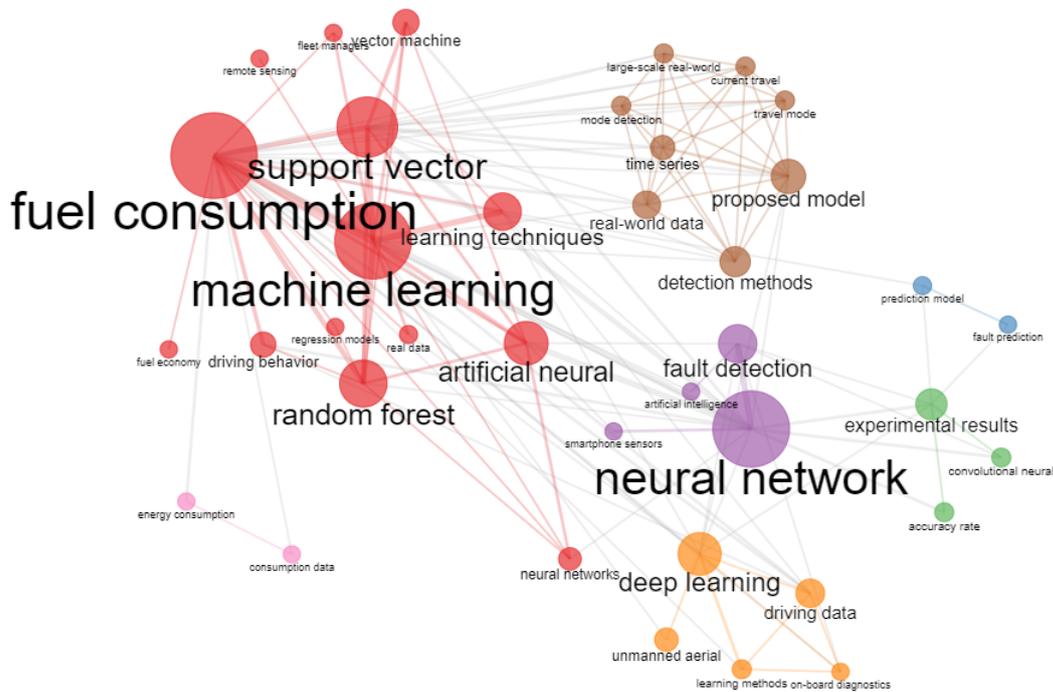


Figura 3. Gráfico de Co-occurrence Network dos títulos do arquivo oriundo da base WoS

Tabela 2. Relatório após a unificação das bases

Tipo de documento	Exclusivo Scopus	Exclusivo WoS	Comum a Scopus e WoS	Total
<i>Article</i>	24	6	21	51
<i>Article; Early Access</i>		1	1	2
<i>Book Chapter</i>	2			2
<i>Conference Paper / Proceedings Paper</i>	32	4	11	47
<i>Conference Review</i>	8			8
<i>Editorial</i>		1		1
<i>Review</i>	1			1
Duplicado	0	0	0	0
Total	67	12	33	112

Tabela 3. Relatório de palavras-chave mais frequentes

Local de pesquisa	Uma palavra - ngram = 1	Par de palavras - ngram = 2	Trio de palavras - ngram = 3
<i>resumo</i>	data (285)	fuel consumption (89)	support vector machine (12)
<i>título</i>	learning (34)	machine learning (81)	vehicle fuel consumption (4)
<i>palavras-chave</i>	systems (91)	machine learning (42)	support vector machine (10)
<i>palavras-chave do autor</i>	learning (31)	fuel consumption (16)	eco driving fuel (6)

* Entre parêntesis a frequência de repetição das palavras

2.3 Identificação e remoção de artigos não elegíveis através de palavras-chave

Os documentos do tipo *Conference Review* foram avaliados cuidadosamente para evitar que um artigo fosse selecionado pelo fato das palavras de busca que originaram o resultado na base Scopus ou WoS aparecerem no resumo da Conferência, porém em artigos distintos. Após essa análise, a lista inicial de 112 artigos foi reduzida a 78 documentos elegíveis.

2.4 Leitura e classificação da lista resultante

Após a leitura dos 78 resumos, e com a ajuda das palavras-chave mais frequentes relacionadas pelo Biblioshiny (expansão da “Tab. 3”), iniciou-se um processo de classificação destes resumos e chegou-se em quatro grupos:

1) Pesquisas relacionadas à busca de causas ou estimativas para explicar o consumo de combustível ou a emissão de poluentes, com as seguintes palavras em destaque: *fleet operations* e *fuel consumption*;

2) Pesquisas relacionadas à determinação ou medição da influência do estilo de direção no desempenho do caminhão ou atuação como assistente de direção, onde as palavras de destaque são: *eco-driving*, *fuels*, *driving styles* e *fuel economy*;

3) Pesquisas relacionadas à utilização de dados para aprimoramento de planos de manutenção, cujas palavras *predictive maintenance*, *maintenance* e *commercial vehicles* são os destaques; e

4) Pesquisas com interesse em engenharia elétrica, sistemas de comunicação, diagnóstico por imagem, pesagem de veículo em movimento e outras. A análise das palavras mais frequentes permitiu identificar aqueles artigos que não eram elegíveis para o estudo de engenharia mecânica, apareceram palavras como *autonomous vehicle*, *battery*, *child*, *peach*, *photovoltaic*, *railway*, *road*, *smartphone*, *UAV*, *vegetation* e *weather* para indicar a não elegibilidade de um artigo.

3. RESULTADOS

As “Tabs.5, 6 e 7” identificam respectivamente os títulos relacionados aos clusters 1, 2 e 3, ou seja, pesquisas relacionadas à busca ou estimativas para explicar o consumo de combustível ou a emissão de poluentes; pesquisas relacionadas a determinar ou medir a influência do estilo de direção no desempenho do caminhão ou atuar com assistente de direção e pesquisas relacionadas à utilização de dados para aprimoramento de planos de manutenção. Os títulos foram ordenados do mais antigo para o mais novo e em ordem alfabética.

Entre esses 30 artigos selecionados a produção científica anual oscila entre 1 e 3 artigos entre 2013 e 2020 e salta a 9 em 2021 e 2022.

A SAE Technical Papers é a fonte mais relevante - 3 artigos publicados. A maioria das fontes locais citadas, 21, é a Transportation Research Part D: Transport and Environment, seguida da Applied Energy e Sensors ambas com 10 citações cada uma.

Os autores mais relevantes foram CHEN CC e YEN MH, ambos com 3 artigos.

Dentre as afiliações, a National Chin-Yi University of Technology de Taiwan se destacou com 4 artigos e a University of Nottingham da Inglaterra com 3 artigos.

Na produção científica por país sobressai-se a China com 15 autores, seguido da Espanha com 9.

O artigo com maior número de citações é Application of machine learning for fuel consumption modelling of trucks com 59 citações totais.

Tabela 4. Títulos de artigos identificados com consumo de combustível ou emissão de poluentes

<i>Título</i>	Ano
Intelligent prediction model of vehicle fuel consumption	2014
Application of machine learning for fuel consumption modelling of truck	2017
Accessing and constructing driving data to develop fuel consumption forecast model	2018
Data driven analysis and forecasting of medium and heavy truck fuel consumption	2019
Methodology to recognize vehicle loading condition - an indirect method using telematics and machine learning	2019
Route-sensitive fuel consumption models for heavy-duty vehicles	2020
A deep learning method for monitoring vehicle energy consumption with gps data	2021
Combining a universal obd-ii module with deep learning to develop an eco-driving analysis system	2021
Estimation of pollutant emissions in real driving conditions based on data from obd and machine learning	2021
Neural models to quantify the determinants of truck fuel consumption	2021
Remote noxemission prediction model based on lstm neural network	2021
Estimating truck fuel consumption with machine learning using telematics, topology and weather data	2022
Integration of second-generation on-board diagnostics data via deep learning to develop eco-driving analysis system applicable to large and small car	2022
Interpretable machine learning models for predicting and explaining vehicle fuel consumption anomalies	2022
Mass detection for heavy-duty vehicles using gaussian belief propagation	2022
Predicting gasoline vehicle fuel consumption in energy and environmental impact based on machine learning and multidimensional big data	2022
Traffic emissions clustering using obd-ii dataset based on machine learning algorithms	2022

Tabela 5. Títulos de artigos identificados estilo ou assistente de direção

<i>Título</i>	Ano
The impact of using gamification on the eco-driving learning	2014
Assessing the impact of driving behavior on instantaneous fuel consumption	2015
Artemisa: A personal driving assistant for fuel saving	2016
Drivingstyles: A mobile platform for driving styles and fuel consumption characterization	2017
Lot sensing for reality-enhanced serious games, a fuel-efficient drive use case	2021
Truck driving assistance system	2021
Driving style analysis by studying pid's signals for determination of its influence on pollutant emission	2022

Tabela 6. Títulos de artigos identificados com utilização de dados para aprimoramento de planos de manutenção

<i>Título</i>	Ano
Overview of telematics-based prognostics and health management systems for commercial vehicles	2013
On-board predictive maintenance with machine learning	2019
Preventive maintenance for heterogeneous industrial vehicles with incomplete usage data	2021
Scenario uncertainty modeling for predictive maintenance with recurrent neural adaptive processes (rnaps)	2021
Handling imbalanced data for aircraft predictive maintenance using the bache algorithm	2022
Predictive maintenance of vehicle fleets using lstm autoencoders for industrial iot datasets	2022

4. CONCLUSÃO

A utilização dos recursos disponíveis na busca avançada do Scopus e da WoS se mostraram muito eficientes nessa seleção de artigos para leitura. A separação em quatro grupos de pesquisa, aliada à busca pela interseção entre estes, apontou-se relevante e trouxe um número reduzido de respostas, apesar da abrangência das palavras de busca. Isso permitiu reduzir de mais de 4 milhões de artigos a um grupo inicial de apenas 112.

A utilização do Biblioshiny e do PyBibx foram muito eficazes, tanto para unir os resultados das pesquisas Scopus e WoS, quanto na análise por palavras que facilitou o trabalho de exclusão e classificação da bibliografia. O Biblioshiny trouxe importantes observações que foram apresentadas no resultado.

A telemática pode ser uma oportunidade de custo acessível aos laboratórios para solucionar duas das razões citadas por Zhang *et al.* (2022) para se construir um conjunto de dados ideal para treinamento: "*É caro realizar experimentos de simulação de falhas e os dados de falha obtidos pela simulação computacional não são práticos.*". Os equipamentos de telemática estão disponíveis no mercado e podem fornecer dados de estudo para a outra grande questão citada por Zhang *et al.* (2022) que afeta diretamente a manutenção, o estudo sobre o consumo de energia e o diagnóstico de falhas: "*máquinas geralmente trabalham em uma condição normal e as falhas são raras.* Ou seja, a problemática dos dados desbalanceados ou de número reduzido de amostras com a falha.

A telemática e a análise dos dados entregues poderão revelar soluções interessantes para problemas relacionados à economia de combustível, redução da emissão de poluentes, criação de novos comandos de necessidade de manutenção mais apropriados à condição de utilização da máquina, atuar com informações para o treinamento de motoristas ou auxiliá-los a obterem uma melhor performance.

5. AGRADECIMENTOS

Os autores agradecem ao PPGMEC/ UFG, CAPES e à Constituição Federal, artigo 206, inciso IV.

6. REFERÊNCIAS

- Aria, M. and Cuccurullo, C., 2017. "bibliometrix: An r-tool for comprehensive science mapping analysis". *Journal of Informetrics*.
- Clarivate, 2022. "Web of science". www-webofscience.ez49.periodicos.capes.gov.br.
- Elsevier, B., 2022. "Scopus". www.scopus.com.
- EPE, E.d.P.E. and AIE, A.I.d.E., 2021. "Transporte rodoviário de cargas – brasil 2021 – benchmarking internacional". *Ministério de Minas e Energia*.
- Pereira, V., 2022. "Project pybibx". <https://github.com/Valdecy/pyBibX>.
- Zhang, T., Chen, J., Li, F., Zhang, K., Lv, H., He, S. and Xu, E., 2022. "Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions". *ISA Transactions*, Vol. 119, pp. 152–171.

7. RESPONSABILIDADE PELAS INFORMAÇÕES

Os autores são os únicos responsáveis pelas informações incluídas neste trabalho.