

# From customer usage to a sustainable product offering through digitalization

Neda Abdolrashidi<sup>1,\*</sup> and Fredrik Öjjer<sup>1</sup>

<sup>1</sup> *Volvo Group Trucks Technology, Vehicle Engineering, 40508 Göteborg, Sweden*

*\*Corresponding author. neda.abdolrashidi@volvo.com*

## ABSTRACT

To develop sustainable vehicles understanding their intended usage is crucial. This knowledge allows us to translate usage patterns into specific requirements, enabling the design of optimal transport solutions. To gain a holistic understanding of the real usage of our trucks, it is essential to provide insights into both vehicle performance and the environment in which they are operated. Logged vehicle data provides a valuable source for understanding customer usage patterns. This data can be clustered into groups with similar usage profiles, and these patterns can then be translated into the requirements. To apply clustering, real-time tracking and data collection from vehicles are necessary. In this work, telematics technology has been utilized as a digitalization tool, enabling seamless transmission, storage, and reception of daily logged vehicle data. The data flow serves as the foundation for conducting a variety of analytics using advanced AI techniques. Amongst the collected data, a set of about fifteen parameters related to the vehicle usage and operating environment have been selected as the input to this study. In the data preparation phase, several preprocessing methods such as data cleaning, data transformation, feature creation and feature selection are applied. Clustering, which is one of the popular AI techniques, has been applied to group data points by implementing a multidimensional algorithm enabling us to handle daily data readings for several parameters over one year and for about two million trucks driving over the globe. In theory, data points that fall in the same group should have similar characteristics called features. In practice, each cluster of vehicles should represent a certain transport mission characterized by its unique usage. In the next step as a verification of the identified clusters in addition to statistical methods, a comparison of the output with information collected in interviews with customers in different transport applications around the world has been performed. Combining the clusters and

customer studies with vehicle domain knowledge resulted in identifying the Transport Application Descriptions (TAD). This study also shows how the proposed method can be applied to assess various transport solutions for different customer applications, with the aim of identifying the optimal transport solutions.

## 1. INTRODUCTION

Understanding the usage of vehicles and the environment in which they are operated is fundamental to developing transportation solutions (Edlund et al., 2004). By comprehensively analyzing usage patterns and incorporating direct customer feedback, it becomes possible to derive precise requirements that inform the design of optimized and efficient transport systems. The transport application descriptions (TAD) provided in this study offer a structured way of including customer usage patterns in product development. Having the transport applications defined, we can evaluate and design solutions that address real customer needs, enabling the development of more sustainable products.

Achieving a holistic understanding of truck usage requires not only evaluating vehicle performance and the operational environment but also integrating insights from customer experiences. This integrated perspective, enriched by customer input, is essential for creating solutions that are both effective and sustainable, ensuring that product development aligns closely with real-world needs and expectations.

In this context, digitalization has emerged as a powerful enabler for companies seeking to enhance efficiency and drive sustainable product development. Advanced digital tools and telematics systems facilitate real-time data collection and analysis, allowing the companies to monitor vehicle usage more accurately and respond to evolving customer requirements (Ghafaripasand et al., 2022). By utilizing digital technologies, the companies will be able to optimize design processes while reducing resource consumption, contributing to more sustainable and customer-centric transportation solutions.

## 2. OBJECTIVES/PROBLEM STATEMENT

Vehicle manufacturers are committed to developing the right solutions for their customers which are aligned with the actual usage patterns and

transportation applications. By thoroughly understanding the performance demands and operational environment conditions, the aim is to optimize performance, improve efficiency, and ensure long-term sustainability. This approach enables us to deliver solutions that meet industry needs while minimizing environmental impact.

New technologies such as battery electric vehicles and hydrogen vehicles face more significant energy limitations compared to diesel trucks, due to battery capacity limits and hydrogen production and storage challenges (Ajanovic, 2023). The energy restriction highlights the importance of understanding how the vehicles are utilized in real-world operations by the customers to select and develop the right technology solution based on the actual transport applications.

The current operational fleet of customer vehicles is predominantly powered by diesel engines, which do not face the same energy limitations as battery electric vehicles and hydrogen vehicles. However, analyzing the usage of diesel vehicles is still considered to be relevant as the logistics system is likely to remain static as production units and logistics hubs are expected to remain unchanged at least in the short to medium term.

### 3. METHODOLOGY

To identify the transport application of the trucks through a comprehensive and real-time perspective, it is essential to continuously collect the data with a proper frequency covering the entire market in which customers utilize the trucks.

To meet this requirement, telematics technology is employed to secure wireless logging and transmission of data from all connected trucks to the company database on a continuous basis, with data updates occurring at a daily frequency.

When studying transport applications, two dimensions of usage must be considered: truck utilization and truck mission. Utilization refers to the operational context of the truck's transport activity, such as long haul, regional haul, or local delivery. Mission, on the other hand, pertains to the type of cargo that is being transported, for example, general freight, car carrier, tanker, and so forth.

In this study, based on the expertise of company specialists with extensive vehicle knowledge and the availability of data, about fifteen parameters were selected to provide insights into the mission and utilization of the trucks.

Besides the usage parameters, the body connection variant is also selected as a factor to identify the type of cargo being transported. These parameters were chosen to cover both operational and logistical aspects of truck usage, enabling a comprehensive analysis of transport applications.

The data collected in this study includes information from approximately two million trucks globally, with data transmitted daily over a one-year period.

To conduct the analysis, the initial and essential step is to secure the quality of the data, achieved through a combination of cleaning and pre-processing techniques. Data validation and cleaning techniques include range checks, consistency checks, missing data imputation and duplicate removal. The pre-processing techniques encompass normalization and standardization, data transformation, data integration and noise filtering (Garcia et al., 2015).

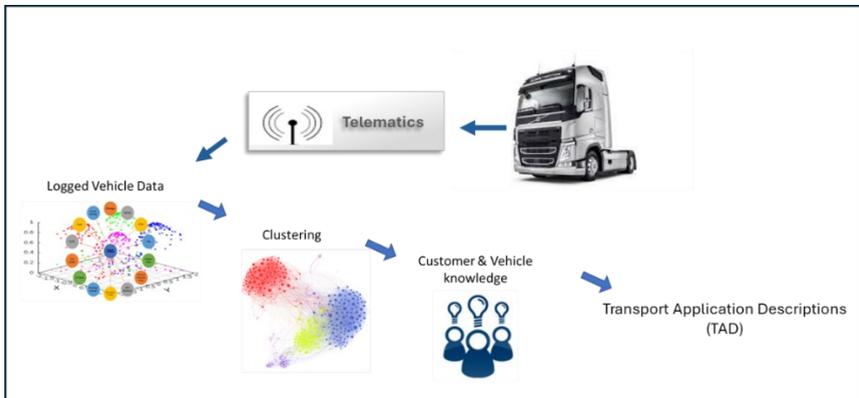


Figure 1: Schematic development process of Transport Applications Descriptions.

To cluster vehicles based on the selected usage parameters, the k-means algorithm which is a widely recognized and effective machine learning technique is employed. K-means clustering is known for its simplicity, efficiency, and ability to divide the big datasets into distinct groups of data by minimizing the variance between different data groups while maximizing the differences between them (Hastie et al., 2009), (Bishop, 2006). This method assigns the data points to clusters based on the nearest centroid and iteratively recalculates centroids until convergence which makes it suitable for clustering multi-dimensional vehicle usage data. Theoretically, data points within the same cluster should exhibit similar characteristics, referred to as features. Consequently, each vehicle cluster is expected to represent a distinct usage pattern.

In the subsequent step, to validate the identified clusters, a comprehensive verification process was undertaken. This process involved not only the application of statistical methods but also an empirical comparison of the clustering results with qualitative insights gathered from interviews conducted with customers operating in diverse transport applications globally. By integrating quantitative analysis with real-world experiential data, this approach ensures that the clusters accurately reflect practical usage patterns and transport missions. Such triangulation enhances the robustness and credibility of the findings, providing a well-rounded validation that aligns data-driven results with industry expertise and customer perspectives.

By integrating the clustering results, customer feedback, and in-depth vehicle domain expertise, the study successfully identified distinct Transport Application Descriptions (TADs). This multidisciplinary approach enabled a comprehensive understanding of vehicle usage patterns, ensuring that the classifications are both data-driven and grounded in practical industry knowledge. The synthesis of quantitative cluster analysis with qualitative customer insights and technical expertise provided a robust framework for accurately defining and characterizing the various transport applications.

In this study, clustering analyses were conducted separately for distinct economic regions characterized by comparable legislative frameworks that influence vehicle usage patterns. Factors such as speed limits, load regulations, and allowable driving mileage, which vary across regions, have a significant impact on how trucks are operated and utilized. By segmenting the data according to these regulatory environments, the analysis accounts for regional differences that affect transport behavior, ensuring that the resulting clusters more accurately reflect the specific operational conditions within each area. This tailored approach enhances the relevance and precision of the clustering outcomes, enabling a more nuanced understanding of transport applications across diverse economic and regulatory contexts.

The names of the identified TADs have two parts; one that specifies how the transport is performed, such as Long Haul and another one that indicates the type of cargo being transported, such as Aggregates.

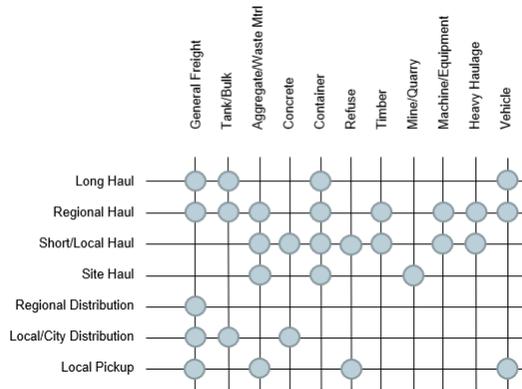


Figure 2: Identified transport applications.

#### 4. BASIC RESULTS AND FINDINGS

Through a comprehensive analysis of the transport applications identified in this study, the importance of conducting a multidimensional clustering approach becomes increasingly evident.

To illustrate this point more clearly, consider the comparison between a general freight regional haul and general freight regional distribution. While an analysis of the daily distances driven reveals significant overlaps, a comparison of their operating radius highlights a significant distinction: the operating radius in regional haul is much greater than that of regional distribution (see Fig. 3).

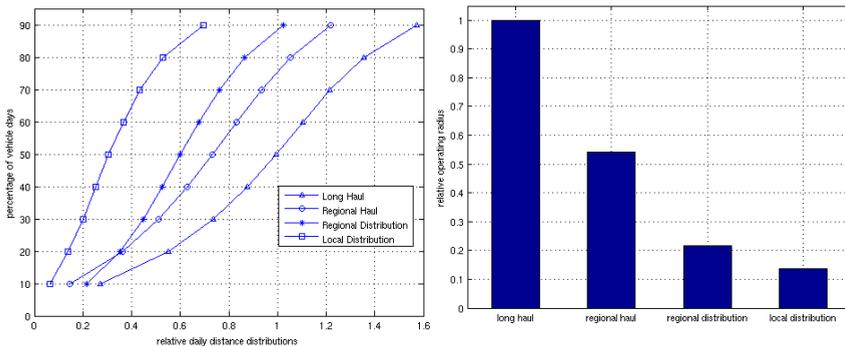


Figure 3: Daily distance (left) and average operating radius (right) for general freight applications in Europe

Moreover, additional operational parameters further differentiate these vehicle categories. Factors such as the average distance between stops, frequency of curve driving, and other operational characteristics exhibit clear variations that are not captured by mileage alone. This example underscores the necessity of a multidimensional clustering approach, which integrates multiple variables to accurately characterize and segment transport applications.

Another noteworthy finding of this study is that, although environmental parameters such as road conditions and topography were not explicitly included in the clustering process, a subsequent comparison of the identified transport applications against these factors reveals a logical and meaningful alignment. For instance, long haul vehicles mainly operate on well maintained motorways and travel between regions, which means they are likely to encounter a mix of flat and hilly terrains. In contrast, local or regional vehicles tend to stay within a region meaning that they are more likely to encounter either flat or hilly terrains. They also operate in cities and on smaller roads of lesser quality to a larger extent than long haul vehicles. Therefore, local and regional vehicles will have a higher proportion of vehicles running in rough road conditions and in flat or hilly terrain as shown in Fig. 4.

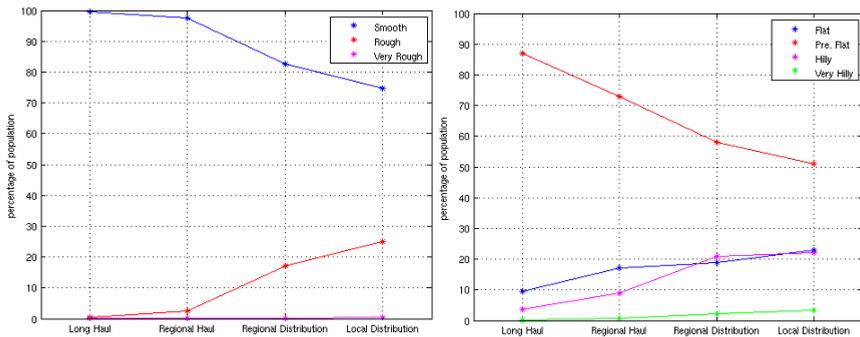


Figure 4: Road condition (left) and Road topography (right) for general freight vehicles in Europe

This observed correlation suggests that the methodology effectively captures underlying operational distinctions, even without directly incorporating environmental variables in the clustering. It also highlights the importance of considering such external factors in future analyses to further refine the understanding of transport applications. Considering the environmental parameters could enhance the precision of vehicle design and operational strategies,

contributing to improved performance and sustainability across different transport applications.

### 5. EVALUATION OF SOLUTIONS

The Transport Application Descriptions (TADs) can be used to evaluate how future solutions would perform if they were to run in the same transport missions as current diesel vehicles.

This is exemplified by evaluating two hypothetical but realistic battery electric vehicle solutions in two different applications, General Freight Regional Haul and Liquid Bulk Regional Haul. The daily distance distribution is very similar in the two applications but the weight that is transported differs as shown in Fig. 5. In the liquid bulk application, the vehicles operate either full or empty whereas the general freight vehicles operate with varying load.

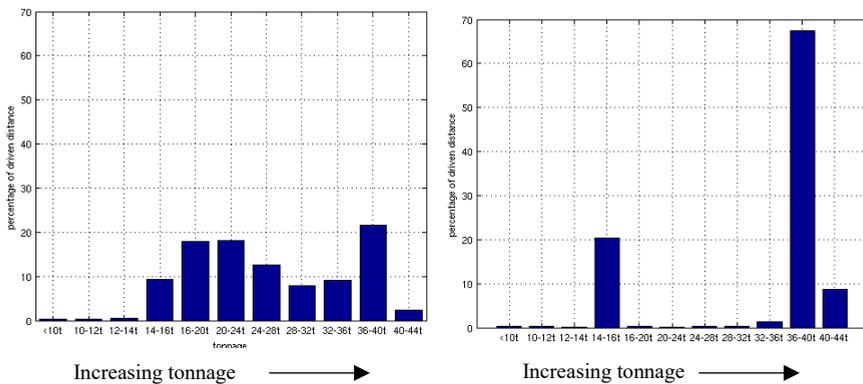


Figure 5: Weight distribution: general freight (left), liquid bulk (right)

The difference between the two evaluated vehicles is that one is focused on range (more batteries) and the other is focused on payload (less batteries). The vehicles are evaluated with respect to two parameters: payload and range, which are crucial for understanding transport mission fulfilment. The vehicles are assumed to start every day with fully charged batteries and are allowed one opportunity charge per day if necessary. Moreover, it is assumed that charging is available when needed and that using it does not interfere with the overall transport mission.

If range is not enough after one opportunity charge the mission is said to have failed on range, and if the vehicle could not have enough payload capacity for the mission it is said to fail on payload. The results are displayed in Fig. 6.

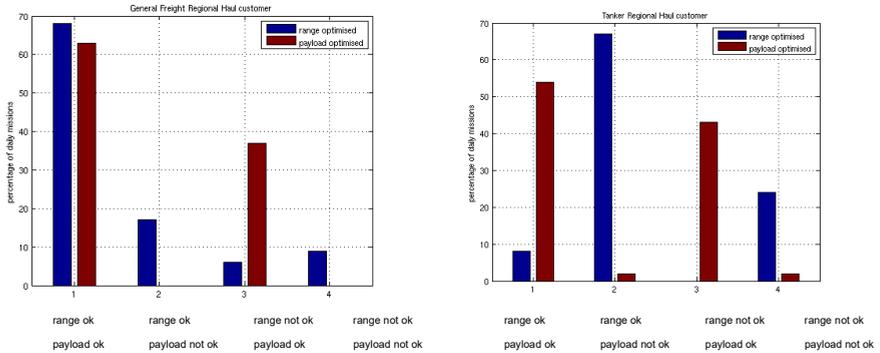


Figure 6: Vehicle performance plots: general freight (left), liquid bulk (right)

For the general freight application, the range-optimized vehicle performs well and succeeds in 70% of the missions and for 20% more of the missions it has sufficient range but must leave some cargo behind. The payload-optimized vehicle succeeds in 63% of the missions but lacks the range for the rest.

For the liquid bulk application, the payload-optimized vehicle performs similarly as for general freight, with only a slight drop in successful missions due to a higher energy need because of the higher average payload. The range-optimized vehicle has the range for 75% of the missions but must leave payload behind for most of them, and for 25% it lacks both range and payload capacity.

Of the two vehicle solutions, the range-optimized vehicle performs better for the general freight application. For the liquid bulk application, it is not as clear which vehicle solution performs the best but with realistic assumptions on the cost of opportunity charging and the reduced income due to lack of payload capacity, the payload-optimized vehicle will be the cheaper to operate. This highlights the importance of understanding transport applications in order to develop the right solutions for customer needs.

## 6. CONCLUSIONS

Digitalization plays a significant role in sustainable development by enabling data driven decision making. In this study, telematics technology has served

as a powerful enabler for data collection, which, when combined with analytics and AI techniques, significantly enhanced our understanding of vehicle usage patterns.

Regarding the usage patterns, this study highlights the critical importance of employing multidimensional clustering rather than relying on a single parameter. Analyzing usage from multiple perspectives is essential to effectively distinguish between the utilization of different applications.

In addition, this study revealed that the operating environment of trucks can vary across different transport applications. Recognizing these differences provides valuable insights and a solid foundation for effective solution design. A realistic understanding of the truck usage enables an accurate evaluation of solutions, ensuring that they are well-aligned with actual needs and remain sustainable by keeping the customer at the center of the analysis.

Moreover, the degree of flexibility available in solution design can vary significantly depending on the vehicle technology. While diesel trucks often allow for greater operational margin, electric and hydrogen trucks require more precise optimization due to energy storage and charging limitations, making it essential to thoroughly understand the real needs of the customer.

## REFERENCES

- Ajanovic, A. (2023). Electricity vs hydrogen in the transition towards sustainable mobility. *Oxford Open Energy*, 2, oiad013. <https://doi.org/10.1093/oen-energy/oiad013>
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Edlund, S., & Fryk, P.-O. (2004). The right truck for the job with Global Truck Applications. *SAE Technical Paper 2004-01-2645*.
- Garcia, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining*. Springer.
- Ghaffarpasand, O., Burke, M., Osei, L. K., Ursell, H., Chapman, S., & Pope, F. D. (2022). Vehicle telematics for safer, cleaner and more sustainable urban transport: A review. *Sustainability*, 14(24), 16386. <https://doi.org/10.3390/su142416386>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.