A combined framework of Biplots and Machine Learning for real-world driving volatility and emissions data interpretation

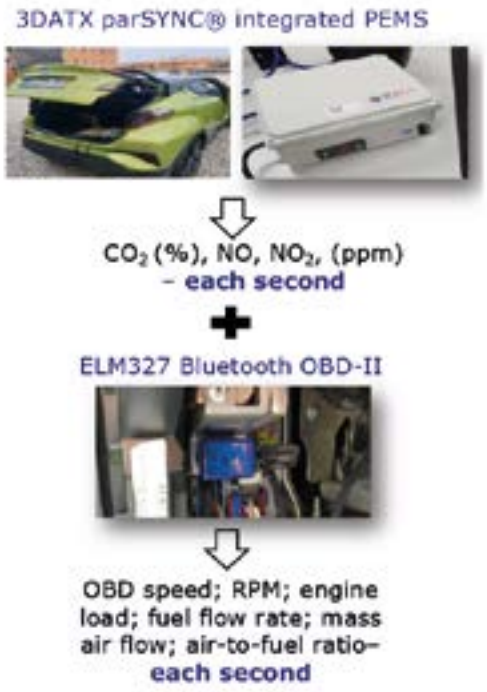
Elisabete Ferreira [1,2], Eloísa Macedo [1,2], Paulo Fernandes [1,2], Margarida Coelho [1,2]

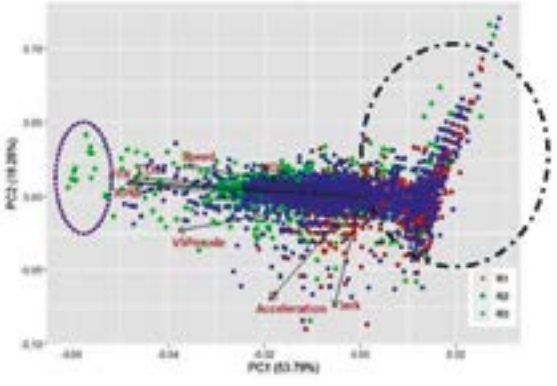
Advanced visualization techniques can be useful for a better understanding of driving behavior and vehicle emissions in real-time.

This study used classic and sparse HJ-biplots to examine the relationship between driving behavior, vehicle engine, exhaust emissions, and route type variables. Different Machine Learning classifiers were applied.

Second-by-second vehicle dynamic, engine, and emissions data were collected from three light-duty vehicles (hybrid electric (HEV), diesel, and gasoline) along three routes (one national (N109) road and two highways (A29 and A1). The dataset included a sample of 12,150 s of speed, acceleration, vehicular jerk (first derivate of acceleration), engine speed, engine load, fuel flow rate, vehicle-specific power mode, and carbon dioxide (CO_2) and nitrogen oxides (NOx) emissions.

The proposed methodology not only enables the distinction of driving styles, road types, and emissions profiles but also allows for revealing the correlation of variables in a single plot. The Random Forest algorithm showed to present the highest accuracy. This study can be useful in the context of road traffic emissions monitoring since it identifies hidden relationships in input data and reduces the redundancy in input parameters without compromising information.





1 – TEMA & Department of Mechanical Engineering, University of Aveiro.
2 – LASI – Intelligent Systems Associate Laboratory, Guimarães.

FIGURE 1
Passenger cars were equipped with an Integrated Portable Emission Measurement System (iPEMS), on-board diagnostic readers (OBD-II), and a Global Navigation Satellite System (GPS) data-logger.

FIGURE 2
Classical biplot for HEV, with 72% of the total explained variance, showing higher emissions values on R2 (A29), separation of acceleration and jerk from other variables.