Measurement 130 (2018) 467–481

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

A novel use of multivariate statistics to diagnose test-to-test variation in complex measurement systems

R.D. Burke[a,*], K.A. Burke, E.C. Chappell[a], M. Gee[b], R. Williams[b]

[a] Powertrain and Vehicle Research Centre, Dept. of Mechanical Engineering, University of Bath, Bath BA2 7AY, UK
[b] Shell Global Solutions (UK), London, UK

ARTICLE INFO

Keywords:
Vehicle Testing
Multivariate Statistics
Multi-way PLS
Partial Least Squares
Testing precision
Batch Process

ABSTRACT

Vehicle testing is critical to demonstrating the cost-benefits of new technologies that will reduce fuel consumption, CO_2 and toxic emissions. However, vehicle testing is also costly, time consuming and it is vital that these are conducted efficiently and that the information they yield is maximised. Vehicles are complex systems, but it is straightforward to install intensive instrumentation to record many data channels. Due to costs, relatively few repeated test cycles are conducted. Identifying correlations within these datasets is challenging and requires expert input who ultimately focus on small subsets of the original data. In this paper, a novel application of Partial Least Squares (PLS) regression is used to explore the complete data set, without the need for data exclusion. Two approaches are used, the first collapses the data set and analyses all data channels without time variations, while the second unfolds the data set to avoid any information loss. The technique allows for the systematic analysis of large datasets in a very time efficient way meaning more information can be obtained about a testing campaign. The methodology is used successfully to identify sources of imprecision in four different case studies to analyse sources of imprecision in vehicle testing on a chassis dynamometer. These findings will lead to significant improvements in vehicle testing, allowing both substantial savings in testing effort and increased likelihood confidence in demonstrating the cost-benefit of new products. The measurement analysis technique can also be applied to other fields where repeated testing or batch processes are conducted.

1. Introduction

Experimental work on many large-scale systems is a sophisticated multidisciplinary challenge involving observing, understanding, and managing many opposing physical effects. To understand the behaviour of these complex systems, many facilities employ comprehensive measurement suites that can easily result in a large total number of data channels. Vehicle test facilities, or chassis dynamometers, are an example of such complex installations where electrical, chemical, thermal, and mechanical processes are working together with time-dependent variations inherent to the testing process.

The chassis dynamometer is a complex facility for testing full vehicle systems over transient duty cycles [1]. The simple concept is to install the vehicle on two large rotating drums that replicate the road and provide a resistance to the vehicle, whilst a robot or human drive the vehicle according to a target speed. The facility is installed in a climatic chamber and an array of instrumentation devices are used to measure emissions, fuel economy and many other vehicle operating metrics (see Fig. 1). Such facilities are widely used by scientists, development engineers and legislative bodies to determine the performance metrics (fuel consumption, emissions, range...) to benchmark different technologies (engines, catalysts, fuels, lubricants) and model variants.

As the full vehicle system is tested, there are many sub-systems operating and interacting with each other (such as fuel supply, after-treatment, electrical system, air conditioning...) [2]. Therefore, it becomes both interesting and relatively straightforward to install a large number of sensors measuring many different data channels. Modern data acquisition systems allow logging of data at high frequencies, but for vehicle-based work over certification driving cycles lasting 20–30 mins, 1–10 Hz is usually sufficient. Nevertheless, this results in a large number of data points.

The major difficulty in chassis dynamometer work is that a single test requires a significant amount of vehicle conditioning before each test [3]. Both older standard (NEDC) and new standard (WLTC) [4,5] cycles require the vehicle to be soaked for multiple hours prior to

* Corresponding author at: Dept. of Mechanical Engineering, University of Bath, Claverton Down, Bath BA2 7AY, UK.
E-mail address: R.D.Burke@bath.ac.uk (R.D. Burke).

https://doi.org/10.1016/j.measurement.2018.07.059
Received 12 February 2018; Received in revised form 8 June 2018; Accepted 21 July 2018
Available online 29 July 2018
0263-2241/

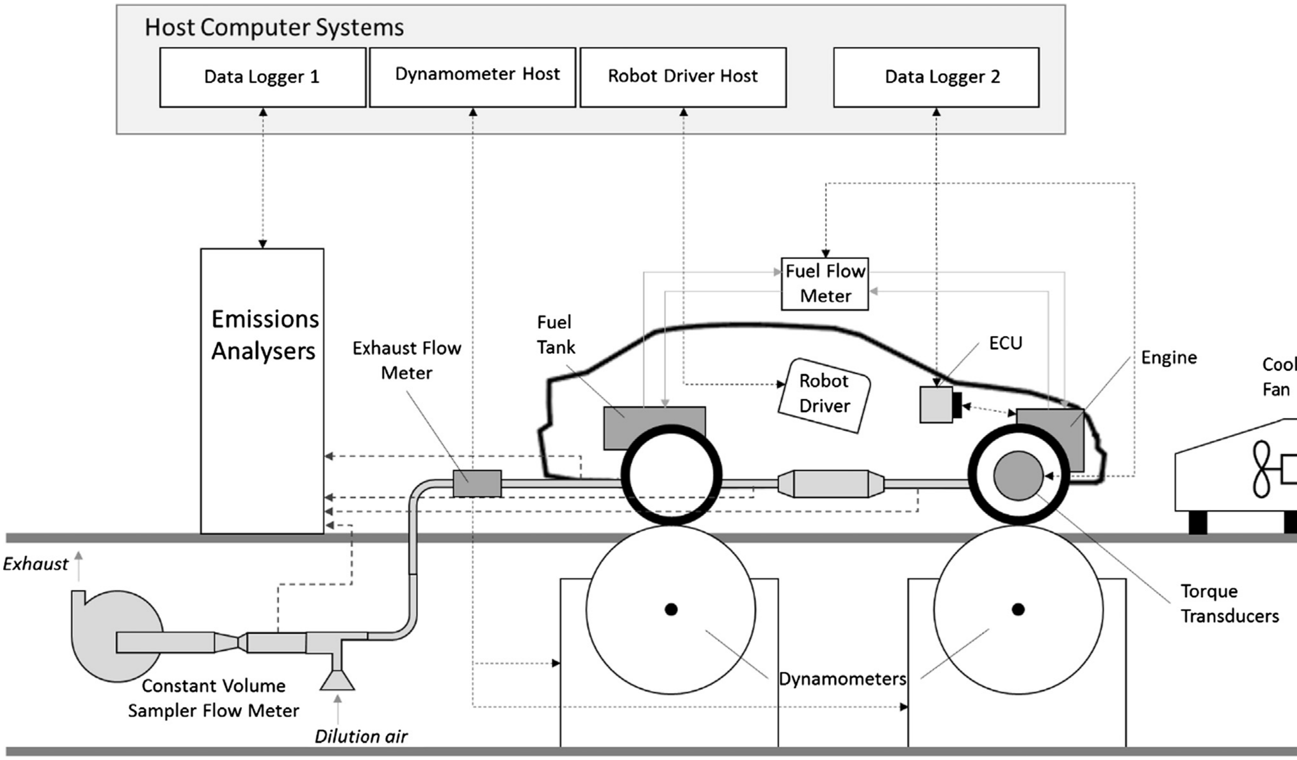

Fig. 1. Typical Chassis dynamometer layout and number of measurement channels

testing to ensure thermal equilibrium. This thermal soaking is also essential for testing precision. In an industrial setting with shift working, it is unlikely more than three cycles will be conducted per working day on a single vehicle.

The consequence of such an experimental setup is a relatively small number of repeated tests (< 10), each containing many data logs (> 1000) for a large number of data channels (> 100). It is useful at this stage to visualise this as a 3-dimensional data set shown in Fig. 2.

As with any experimental work, there will be variability inherent to the test vehicle [6], inherent to the measurement system [7] and inherent to the test methodology [8]. This causes undesirable variation in the key measurements which can mask real benefits that are being evaluated. As the source of variation is unknown, the analyst must isolate potential causes by identifying trends within the measured data. This is typically conducted sequentially via hypothesis testing, however when the number of measurements becomes large, it is not practical to analyse the complete dataset.

The aim of this work is to provide a method for easily identifying the sources of variability within the data sets whilst making use of the full richness of available data. Firstly, independent reviews of vehicle testing and multivariate statistical applications will be presented.

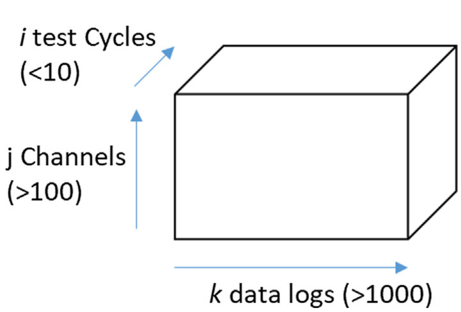

Fig. 2. Typical 3-dimensional data set obtained from repeated batch processes such as chassis dynamometer tests.

Secondly the experimental facilities and data analysis methodology will be described. Finally, the four case studies will be presented where the new technique is applied to identify potential sources of variability.

2. Background

2.1. Data analysis techniques for vehicle powertrain testing

Two common approaches are adopted for the analysis of data from engine and vehicle test facilities:

1. Expert led analysis uses the experience of the engineer to interrogate the data based on previous experience and knowledge of the physical processes to identify and test the likely causes of variability. This is a very time-consuming process and relies on physical trends being sufficiently distinct for the expert engineer to identify. It is important to note that where there are multiple competing physical effects these can be very hard to identify.
2. The statistical approach makes use of multiple linear regression (MLR) modelling to interrogate the data set [9]. This is most commonly used in conjunction with Design of Experiments where a test plan has been determined in advance and the regression modelling is required to interpret the resultant data. Nevertheless, regression modelling has been used to assess random variation within data sets. Regression models have the benefit of being able to analyse data with correlations to more than one independent variable.

Most data analysis from engine and vehicle test facilities are expert driven and rely on the experience but also the time an individual can devote to the particular problem. Although there are benefits to the regression models, the mathematics require that there be more test points than there are input variables (variables suspected of correlating with the particular output).

This is not compatible with the typical data set from vehicle testing, and to apply the regression model approach, a small number of