

Capability Assessment Process for the Optimisation of Testing Facilities for Powertrain Development

Philip Lawson, John Houldcroft, and Andrew Neil
Jaguar Land Rover Ltd.

Andrea Balcombe, Richard Osborne, and Antonio Ciriello
AVL UK Ltd.

Wilhelm Graupner
AVL List Nordiska AB

ABSTRACT

A recent trend in powertrain development organisations has been to apply processes historically associated with manufacturing. The aim is to capitalise on the resulting productivity gains to contain the increasing test demand necessary to develop current and future product. Significant obstacles to the implementation of manufacturing derived methods include the lack of clarity of the engineering test requirements and existing working practices in the product development environment.

The System Optimisation Approach has been presented in previous work as a potential solution [1]. As an extension, this paper introduces a new concept closely related to the established manufacturing principle of Process Capability (Cp). Application of the resulting method quantifies the test facility's capability to provide a test result subject to a specified statistical confidence within a certain number of test repeats. In addition the process provides objective direction when an engineering target is unspecified or deemed unachievable.

As a generic, data-driven method the Capability Assessment Process is a tool to objectively reduce test effort whilst retaining sufficient confidence in test results and conclusions. On a wider scale, the process directly links engineering targets to potential facility investment.

CITATION: Lawson, P., Houldcroft, J., Neil, A., Balcombe, A. et al., "Capability Assessment Process for the Optimisation of Testing Facilities for Powertrain Development," *SAE Int. J. Engines* 9(3):2016, doi:10.4271/2016-01-0982.

INTRODUCTION

This paper documents processes derived in a recent joint project between AVL and Jaguar Land Rover at JLR's emissions test facilities. The aim of the project was to investigate current measurement uncertainty and testing processes in order to make improvement recommendations that would enable a reduction in testing effort.

Since the introduction of the stringent Euro5, and subsequent Euro6 emissions legislation, automotive manufacturers have come under significant and sustained pressure to deliver vehicles onto that market that meet these emissions targets without compromising on comfort and performance [2]. This has led to increasingly complex engine designs and control strategies in order to ensure emissions remain within the required levels in all operating conditions. The vast amount of effort required to calibrate all of the hundreds of maps contained within Engine Management Systems with all possible combinations has done more than simply increase the amount of

development testing; the added complexity makes it more difficult to ensure test repeatability and increases the likelihood of failure at any point in the testing process [3].

The rate at which this demand on test effort and sophistication has increased has out-run the cycle of investment and upgrades on testing hardware at many automotive manufacturers, and with the introduction of SULEV legislation and Real Driving Emissions (RDE) legislation on the horizon, this trend in the industry is set to continue. In recent years, several test departments from various manufactures have addressed aspects such as utilisation and data quality in order to increase the efficiency of their testing activities, but what is to be done in the case of a test facility that operates at very high utilisation and is also over-subscribed with work? Two extremes exist, 1; increase the capacity of the facility, or 2; reduce the amount of work required from the facility. If investment cost was no limitation on Option 1, there would remain other equally real limitations to facility expansion, for example; physical space and infrastructure to support the additional hardware, time to install and ramp-up new equipment to operating conditions, significant increase

in resource for operation and maintenance, etc. Also, the possibility exists that if you provide more test facilities for development teams, they will simply see an opportunity to conduct even more tests, as the pressure to manage the increasing complexities and emissions requirement puts a strain as much on powertrain development teams as it does on the test facility.

Therefore, this work explores one way to address the second option; to reduce demand on the test facility. This was done using established methods of statistical analysis, as well as using previously published methods of evaluating testing practices [1]. In addition, a new concept was developed, based on proven methods used in manufacturing. Core to this concept is the development of a metric to evaluate a test facilities capability in delivering the quality of results required, and also provides a means to match test facility capability with development requirements. The development of this concept forms the basis of this paper.

This paper begins with a review of measurement uncertainty in terms of statistical convention and the implications on powertrain development work. This will be followed by an explanation of the concept of Process Capability, how it is currently used in manufacturing, and which aspects can and cannot be carried over into a development environment. Finally, the methodology behind the development of the Test Facility Capability Assessment Process is explained, including examples of how it can be used to address some common questions that arise during product development activities.

BACKGROUND

Measurement Uncertainty

The main purpose of the project at JLR was to quantify the measurement uncertainty of the emissions test facility. Such an activity is common in powertrain test facilities and is well documented in the literature, particularly in the area of emissions testing and with good reason [2,4]. As permitted emissions levels continue to decrease, it puts pressure on development teams to calibrate engines to emit ever smaller quantities of the individual pollutants and an equal pressure on the test facility to ensure these ever decreasing quantities are measured repeatably. It is widely acknowledged in the industry that measuring such low levels of pollutants is difficult even if the most sophisticated instrumentation and equipment currently on the market is available, as routine calibration and correct set-up for each class of powertrain is absolutely necessary in order to measure such low levels [2]. Furthermore, complexities of the sample handling systems necessary to conduct chassis-dynamometer based emissions testing introduce systematic uncertainties that dominate the repeatability performance of modern analytic technology. Therefore, understanding the uncertainty of measurement becomes critical at this stage of powertrain development, with the implication being wider reaching than simply 'pass or fail' on certification and compliance tests.

The act of measurement is widely misunderstood [5], and the concept of accuracy more so. To measure is to use an arbitrary unit (e.g.; kilograms, meters, litres) to describe the properties of something, such as its mass, length, volume, etc. In order to take a measurement

an instrument must be used, and as such, a measurement can never be exact as the instrument can only approximate the property being measured. The accuracy of that approximation (how close it is to being exact), as well as the precision (repeatability) are dependent on both the quality of the instrument and the processes used to perform the measurement [6]. As an exact measurement can never be obtained, accuracy and precision depend on eliminating all possible sources of systematic error in an effort to achieve a measurement as close as possible to the true value: a dependable result [5].

Systematic errors affect measurement results in a predictable and constant manner, that is, they exert the same influence on the results at each measurement, such as incorrect zeroing or calibration of measurement devices, incorrect testing processes, or undesired environmental conditions. Statistical error on the other hand describes uncertainties that arise from random fluctuations of the data within a finite sampling population. While repeated testing (increasing the size of the sampling population) can reduce statistical error, sources of systematic uncertainty cannot be reduced via repeated measurements. Therefore, for an unaltered process, the precision and accuracy of a measurement are maximised when the measurement is repeated a sufficient amount of times to reduce the statistical error such that the intrinsic systematic errors of the system dominate the measurement uncertainty [5]. The process outlined in this paper specifically addresses random statistical error and results from repeated testing, however systematic errors were addressed during the project using related processes.

Knowing the combined influence of all sources of systematic and random statistical error allows one to estimate the uncertainty of a measurement. The outcome of this estimation is frequently referred to as an 'uncertainty budget,' and it is a powerful method of understanding the quality of a measurement [4], as it expresses a measurement plus its uncertainty (to a given confidence interval) in much the same way as tolerances are specified on the dimension of a part to be manufactured. However, the process of determining an uncertainty budget has several disadvantages in a product development environment;

- The calculations used can be cumbersome
- The resulting uncertainty budget only applies to the specific measurement process investigated
- A significant number of repeated tests are required to build statistical confidence
- Uncertainty budgets are quoted in the units measured, i.e.; they are not non-dimensional.

This approach may be practical in a manufacturing environment, where parts to be made are highly specified, enough can be manufactured and measured in a short period to build statistical confidence, and processes are well defined and strictly adhered to. But these conditions are far removed from a powertrain development environment. In the experience of the authors, processes for engine calibration and development activities are often roughly defined due to the fact that each programme task requires greater or lesser innovation to meet engineering targets, therefore there must be flexibility to adapt tests. Even in the case of emissions testing, where

drive cycles are strictly defined, the vehicle being tested is rarely stable during a development phase, with regular changes of calibration and components. Unlike manufacturing for a part to a specific dimension, engineering requirements in powertrain development are usually a minimum or maximum limit, and development engineers can and do try to improve upon those limits. Frequently, it is necessary to find an optimum trade-off between emissions quantities that are easier to improve versus quantities that are difficult to improve in order to meet all legislative limits. Thus, the desired measurement result is a moving target. Finally, repeated testing of a development vehicle to build the required statistical confidence demanded by uncertainty budget calculations are unfeasible for automotive manufactures, either in terms of time or resource. Many sources state that thirty samples are generally required in order to obtain a normal distribution of results [7], and in manufacturing where samples sizes typically range between 50 and 200 results, a sample size of thirty is considered an absolute minimum. Indeed, it is suggested a minimum of ten samples is required to calculate a reasonable uncertainty budget for a given factor [5], but even examples in the literature of investigations into emission measurement uncertainty show that the research teams struggle to complete more than five or six repeat tests [2,4]. If it is not possible in a research environment, how can it be conceivable in a development environment? Finally, uncertainty budgets are quoted in the units of the quantity being measured, therefore an uncertainty budget for every critical quantity would have to be carried out and individual specification set.

This does not mean that attempting to understand measurement uncertainty in a development environment is futile, on the contrary, it is critical. The individual pollutant quantities emitted by modern engines is now so low that it approaches the resolution limits of measurement devices; understanding the measurement uncertainty provides an indication that observed differences between multiple results is due to real physical variation as opposed to random fluctuations of the measurement system. The opposite applies to engine ECU calibration activities; has a calibration change made a real improvement on emissions output or are the results different simply because of inherent random error in the system? This situation is all the more striking as it is common practice to make calibration decisions on very few repeat measurements, sometimes only the one measurement, as the luxury of repeating tests does not exist. Understanding the measurement uncertainty can help answer this question.

Most importantly, understanding measurement uncertainty has an implication on setting engineering targets in order to pass certification. Engineering targets in a development environment are typically set significantly lower than the legislative limits in order to account for the random error due to normal build variation in mass manufactured powertrains, as compliance of vehicles for sale in Europe dictates testing vehicles randomly pulled from the production line. A poor understanding of the measurement uncertainty during testing can result in engineering limits being set too high, and production vehicles might fail compliance tests, or the engineering limits may be set too low, resulting in wasteful development effort to meet artificially low emissions targets.

So how does a development organisation set realistic engineering targets and ensure that the measured results of each test are accurate enough to build a robust calibration that can meet emissions certification and compliance? Traditionally, measurement accuracy for engineering targets were set as a percentage of the required value, i.e.; the result had to be within plus/minus x% of the engineering target. However, as legislative targets have decreased, the absolute values of some of these percentage tolerances would equate to values so low that they are within the resolution of the measurement devices. Setting a standard deviation target (e.g.; within +/- 3stddev) or confidence interval target (95%) has also been used, but as with uncertainty budgets, these targets are dimensional and would have to be calculated and set for each individual quantity.

Ultimately, all of these methods fail to provide one vital piece of information; is the test facility capable of measuring to the specified tolerances?

So the challenge of this project was to find a method that could link engineering targets with test facility capability in the simplest way possible.

Process Capability

Manufacturing within the automotive industry has undergone revolutionary changes in operations in the last several decades, as manufacturers have almost universally adopted and adapted the highly methodical and process driven practices established so successfully in Japan [8]. As these operational processes developed, they have become more and more standardised across the industry, as have the methods of evaluating those processes. One such process stood out as having potential to be adapted into a development environment to assess test facility capability.

The concept of Process Capability, frequently referred to by its indices, Cp, is a quality metric used to estimate the performance of a process [7,9,10]. It directly compares the actual variation of a process to the desired specifications, in a manufacturing environment this often implies the question; 'does this manufacturing process produce parts within the desired specifications?' The formula for Cp is [7,9];

$$Cp = \frac{\text{allowable process variation}}{\text{actual process variation}}$$

$$Cp = \frac{USL - LSL}{6StdDev}$$

Where;

USL = Upper Specification Limit

LSL = Lower Specification Limit

StdDev = standard deviation of the sample

The Upper and Lower Specification Limits define the maximum and minimum value allowable for *any individual result*. For example, in a process that cuts wires within an Upper Specification Limit of 10mm

and a Lower Specification Limit of 19mm, every individual produced must between 21mm and 19mm in length. Every single wire outside of these limits would be deemed defective.

The standard deviation used to calculate C_p is the standard deviation for the sample of measurements taken. Frequently it is not possible to have results for the entire population, especially in manufacturing where it's not feasible to measure the hundreds, thousands or millions of parts that a process can produce. Therefore a sample is used to represent the entire process. This is a reasonable practice as for process variation with a normal distribution, six times the standard deviation represents 99.73% of the observations [11].

Therefore, C_p compares an ideal normal distribution (based on the specification limits) against the actual normal distribution. So a process with $C_p = 1$ is theoretically capable, where 0.27% of the process would be defective. In other words, the allowable process variation is equal to the actual process variation. However, in manufacturing, targeting $C_p = 1$ is considered an inappropriate value as this would equate to an expected defect rate of 2,700ppm [9,11]. So recommended C_p targets are typically 1.33 (8σ), 1.66 (10σ), or for a defect rate of 0.002ppm a C_p of 2 (12σ) is targeted [7,9].

Despite its simplicity, the vast majority of the literature [7,9,10,12] stresses that some fundamental aspects of the process being measured must be understood before C_p can be successfully used as a metric. Many experts have noted a tendency to jump straight to the calculation of C_p without performing a prerequisite analysis of the process, resulting in misleading assessment of process capability that can have a significant operational and financial impact.

Before calculating C_p , the literature suggests the following must be true for the process [7,9,10,12]

- The process is consistent for the sample taken: There is a common misconception that C_p applies to short term evaluation only. What the literature makes clear is that C_p should only be applied to processes that are completely repeated, i.e.; the same actions are taken, the same tooling is used, and the same measurement equipment is used, etc., irrespective of the time taken to gather results.
- The process is in control: Experts encourage the use of simple control charts to visually assess that the process is in control. Evidence of steps, drifts or outliers in the data can indicate special causes of error in the process that should be corrected. Not doing so can falsely indicate a less capable process, thus leading to wasteful improvement actions.
- The distribution of the sampled data should be approximately normal: The Process Capability metric is specifically used for a process with normal, Gaussian distribution. It is not valid for other types of distribution.
- The mean of the sampled data should be approximately centred: The calculated average of the sample should be approximately centred in the distribution of the data. Using a histogram can visually indicate where the mean lies. If the mean is significantly off centre, than the index C_p should not be used.

Ensuring that these conditions are met is quick and easy, relying on graphical data that is quickly plotted and involves minimal calculations [7,12].

Therefore, several characteristics make Process Capability an ideal candidate to apply in a development environment;

- It is a tool that is very well established and widely used in manufacturing. Information about this concept and how to apply it is prolific and easily obtainable.
- The calculation of the indices is extremely easy, involving only three parameters.
- The indices are unitless, making it a convenient metric to use for comparison and evaluation, especially when many differing quantities are involved.

In addition, the application of Process Capability complements processes already developed and used by the authors, thereby strengthening the potential application of the methodology.

METHODOLOGY

System Optimisation Approach

The System Optimisation Approach is a process developed by AVL, in conjunction with Jaguar Land Rover. It is a set of systematic actions that assess development requirements in relation to test facility capability, and highlight areas for improvement actions, including investment in testing [1]. This process formed the framework for the project being discussed, guiding all actions taken, and will be briefly outlined here.

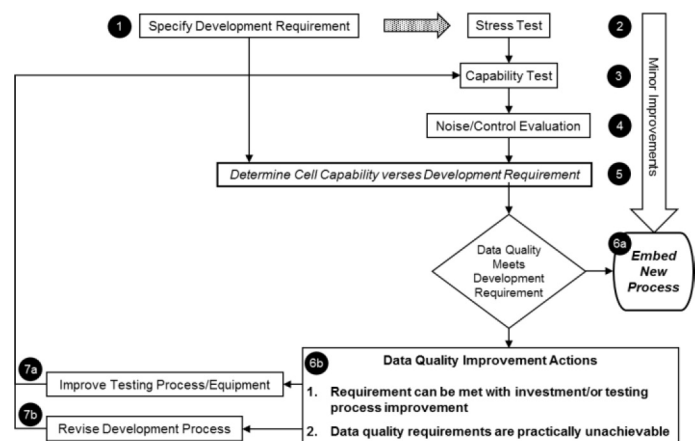


Figure 1. The System Optimisation Approach.

The System Optimisation Approach (Fig.1) comprises of seven major steps, which are followed in a strict sequential order;

Step 1. - Specify the Development Requirement: The very first step is to understand the development requirement in order to relate this directly to testing capability. As development teams are typically the internal customers of the test facility, it is vital to determine exactly what data and of what quality is required to complete their development tasks so that the test facility is able to provide value to their internal customer. Frequently, the development requirements are