UKACC International Conference on Control 2012
Cardiff, UK, 3-5 September 2012

USE OF CONFIDENCE LIMITS IN THE SETTING OF ON-BOARD DIAGNOSTIC THRESHOLDS

Paul J King
Powertrain Control Systems and Calibration
Jaguar Land Rover Limited
Coventry,
CV3 4BJ, U.K

Keith J Burnham
Control Theory and Applications Centre, Coventry University,
Coventry,
CV1 5FB, U.K.

Abstract—**This Vehicles sold in the US and Europe have to be equipped with a Diagnostics, called On-Board Diagnostics (OBD), which monitor the performance of various elements of the emission control system. The driver is informed as to any failures by the use of a Check Engine Light on the dashboard of the vehicle and then should return the vehicle to the dealership for rectification. The vehicle manufacturer's aim is to ensure that the Check Engine Light is only illuminated for legitimate failures. For the calibration of an On Board Diagnostics there needs to be sufficient separation between the response of a good sensor and a failed sensor, the setting of this threshold should be based upon a statistical model of the data so that the predicted failures rate can then be determined. By applying confidence limits to the models allows the engineer to understand how additional data points will effect the calculation of the failure threshold. This gives the engineer the ability to determine the tradeoff between the number of data points and a confidence in the estimated statistical model.**

Keyword-components: ***Automobiles, Detection, Diagnosis, Engines***

I. INTRODUCTION

Vehicles sold throughout the world are subject to an increasingly stringent set of emission thresholds. To achieve certification, all sensors and vehicle sub-systems that may affect vehicle exhaust emissions have to be monitored by an On-Board Diagnostic (OBD) system that is part of the Engine Management System (EMS) or any other embedded controller [1]. This requirement was first introduced in the US in 1988 for OBD1, for open and short circuit faults, and in 1994 for OBD2, for changes in sensor and actuator responses [2]. For Europe this legislation, denoted EOBD, has been introduced for all vehicles built after January 2000 [3]. Both sets of legislation link the performance of the different diagnostics to emission thresholds. In the event of component or sub-system failure, a 'check engine' light must be illuminated as an indication to a driver that there is a problem, so corrective action can be taken to minimise the pollution caused by such a fault. As the emissions thresholds are continually reduced, more sophisticated techniques are required to be employed to meet these increasingly tightening thresholds.

As the Vehicle Emission Legislation drives down vehicle pollution the impact on the diagnostics is that they have had to become increasing more complex to determine a failed system. As a consequence the diagnostics are becoming have become increasing more like models of the functions that they are monitoring, if only over a restricted set of operating conditions, and as such the diagnostic results that they produce are becoming more likely to fit to a Gaussian Distribution.

OBD Diagnostic Calibration engineers develop test plans that invoke the worse conditions for the diagnostics, by introduction of variety of test conditions. These are typically different fuel specifications, operations at different ambient conditions (hot, cold and altitude), with different driving styles and tolerance sensors. Each diagnostic will have it's own set of worse case test conditions which have been developed through the experience of the engineer and lessons learnt. The approach used by the calibration engineer is to collect data for these conditions and then fit a Gaussian distribution to the data to set thresholds to ensure that 'normal' systems does not false flags and that 'failed' system flag in a timely manner.

By making use of these Gaussian models and through the information obtained by the use of confidence interval of these models allows a more conservative and robust threshold to be set. Since the variation in the models reduces as the more data is collected they also provide the Engineer with a way of gauging whether collected any more data will significantly change their results. This is especially useful when collecting Fault condition data which where it is difficult to generate a large amount of data either because it requires specialist hardware setup or a specific environmental condition for which there might be a limited amount of testing time available.

The paper is organised as follows: Section II Problem Formulation which outlines the diagnostic used within this paper, Section III defines the Diagnostic Specification which outlines the set of conditions for setting the diagnostic thresholds, Section IV which defines the calculation of the Confidence Interval, Section V Process Monitoring and

Analysis details the implementation and analysis and finally, Section VI details further work and concludes the paper.

II. PROBLEM FORMULATION

On modern engines the air fuel ratio (AFR) has to be tightly controlled so that the Three Way Catalyst is operating at it's optimum and providing the correct emissions control [4]. AFR is measured as the ratio between the mass of air and the mass of fuel for pure octane the stoichiometric mixture is approximately 14.7:1. The exact composition of fuels varies seasonally and geographically so modern engines use a more convenient measure of lambda (λ) rather than AFR to allow them to control combustion process. Lambda (λ) is defined as the ratio of measured AFR to stoichiometric AFR for that given mixture. A Lambda of 1.0 is stoichiometry, rich mixtures are less than 1.0, and lean mixtures are greater than 1.0.

In this paper the calibration of an Individual Cylinder Air Fuel Ratio Diagnostic is investigated. This is a diagnostic which was first introduced for vehicles sold as 2010 Model Year vehicles [5,6]. Since this is still a relatively new diagnostic it was decided to carry out a more detailed analysis of the diagnostic results. The data analysed in this paper was collected from a cold ambient environment trip, this condition being deemed as being the worse case condition potentially giving the smallest separation between a fault free and the failed set of data.

Prior to the test trip the failure condition for both a rich and lean shift in λ for individual cylinders was determined by carrying out tests over an Emissions Drive Cycle. A failure condition being determined when the fuelling shift resulted in an emissions test result being 1.5 times the certify emissions standard. The failure threshold is set as the amount of shift in λ. The diagnostic infers the amount of Fuelling shift that each cylinder is experiencing from λ = 1 to determine a failure.

III. DIAGNOSTIC THRESHOLD SETTING

Ideally the failure threshold should be set so you can capture all of the diagnostic results. However, in practice this may not always be possible so a minimum target of 90% of the data which contains a fault condition should result in the diagnostic bring on the Check Engine Light. To put on the Check Engine Light in USA requires that a failure is detected on two successive diagnostic operations. So to achieve the 90% detection on two successive tests requires that for a single test the failure threshold should be set at a level approximately 95% (100√0.9). To determine the Failure Thresholds in Section 4 it is more convenient to convert this figure into a Standard Deviation. In the case we are assuming that the threshold will only occur on one side of the distribution closest to the Fault Free set of data. On the Failure Threshold side of the distribution the point at which 45% of the population is represented by a Standard Deviation of 1.64, the other 50% of the data being containing the other half of the distribution.

For the fault free set of data we need to ensure that the fault thresholds are greater than 3×Standard Deviations from either side of the distribution we will refer to this as the Rich or Lean Robustness Threshold. This will then allow 99.74% of the data to be correctly identified as being fault free for a single diagnostic result. For the two successive diagnostic results this would then lead to the possibility of flagging a fault free system as having a fault as being 1 in 148,000 tests.

IV. PROBLEM FORMULATION

The set of data collected from the diagnostic cannot precisely define the characteristics of the population. The sample can only define a range of values for both the probable mean position and the probable standard deviation value. Confidence interval calculations are used to define a probable range of values for the population mean $\bar{x}_U$ (Upper) and $\bar{x}_L$ (Lower) and the population standard deviation σ_U and σ_L. This then generates a range of possible statistical models that will include the population model with a given confidence.

Equation (1) [7] is used to Calculation of the Confidence Interval for the Mean

$$\mu = \bar{x}_S \pm t_{\beta,n-1} \frac{\sigma_S}{\sqrt{n}} \tag{1}$$

$\bar{x}_S$ is the sample mean t is the Confidence Factor based upon the Student's t-Distribution, σ_S is the Sample Standard Deviation, n is the Sample Size and β is either α for a single sided confidence Interval or $\alpha/2$ for two sided confidence interval. Where α is the significance level and for this paper a value of 5% or 0.05 will be used.

Calculation for the Confidence Limit for Standard Deviation is given by

Lower Limit Upper Limit

$$\sigma_S \sqrt{\frac{n-1}{\chi^2_{\beta,n-1}}} \le \sigma \le \sigma_S \sqrt{\frac{n-1}{\chi^2_{1-\beta,n-1}}} \tag{2}$$

χ^2 is the Confidence Factor based upon the Chi-Squared Distribution.

The results of equation (1) and (2) will produce a range of standard deviation and means values which will include the population model with a 95% confidence. This confidence increases and the range of these values reduces as there is an increase in the amount of data, n, as it is collected. Using this range of values it is then possible to choose a combination which will give the worse case Failure or Robustness thresholds. In terms of the Standard Deviation it is the Lower

Limit calculation in (2), σ_L, which produces the maximum value of sigma and this will generate a model which produces a greater range of λ values. The means, $\bar{x}_L$ or $\bar{x}_U$, are chosen so as to give the smallest separation between the Failure threshold and Fault free condition in each case.

V. Process Monitoring and Analysis

From testing three specific sets of data were collected a the Fault Free, represented by under score F, a set of data in which a rich shift has been introduced to represent the emission failure threshold, represented by an underscore R, and a set of data for the lean shift, represented by an underscore L. This resulted in 6294 sets of results with a normal engine, 102 diagnostic results for the rich shift 106 results for the lean shift.

The first thing to check is that all of the data sets have a Gaussian distribution. This was done using the LilleTest function within MATLAB which performs a Lilliefors test [8, 9] of the default null hypothesis that the samples comes from a Gaussian distribution, against the alternative that it does not come from a Gaussian distribution. The test returns a 1 if it rejects the null hypothesis at the 5% significance level. For the fault free condition the set of data failed this test to future investigate this data is plotted on a Quartile-Quartile (QQ) Plot. This shows the raw data as blue crosses plotted against an ideal Gaussian model which is the red dashed line. This shows that for the bulk of the data between the 5% and 95% percentiles it fits a Gaussian distribution. It is only the tail information of the distribution which does not match this statistical model. From the shape of the tails in Figure 1 it indicates that the data has come from a 'fat tailed distribution' where the data extends further than for the tails of a Gaussian distribution. This can be seen by considering the 0.999 and the 0.001 percentile points which lie at points 0.925 and 1.07 respectively and should, for a Gaussian Distribution, lie at 0.94 and 1.06. Even though this statistical model does not accurately fit the data it was decided to assume that the fault free set of data could be considered to be Gaussian initially and then reviewed when the final thresholds for the Rich and Lean have been determined. The risk is by making this assumption that the variance used for the Fault Free Data will under estimate the true risk.

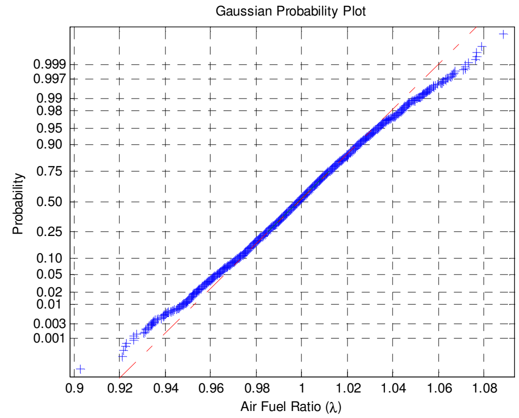

Figure 1: QQ plot for the Fault Free Data

	L_F	S_F	U_F
$\bar{x}$	0.9980	0.9985	0.9990
Σ	0.0217	0.0213	0.0209

Table 1: Statistical Model for Fault Free Data (β = 0.025)

Table 1 has been calculated for $\beta = \alpha/2 = 0.025$ in Equations (1) and (2). The Standard Deviation σ_{L_F}, from Table 1, gives the largest value of sigma and is used to calculate the Robustness Threshold for both the rich and lean sides of the distribution. For the Lean Robustness Threshold the mean $\bar{x}_{U_F}$ was used to give a value of 1.0640, the Rich Robustness Threshold was obtain using $\bar{x}_{L_F}$ which gives a value of 0.9330.

Table 2 shows the Rich Failure model information note that since we are considering a single side threshold then β = 0.05. Figure 2 shows the Histogram of the raw Rich Failure data and the red solid line shows the worse case statistical model using $\bar{x}_{U_R}$ and σ_{LR}. Using this model the Rich Failure threshold is calculated as 0.8579 and is shown on the graph by the solid vertical red line. The dotted blue vertical line shows the sample Rich Failure threshold of 0.8507 calculated by making use of σ_{SR} and $\bar{x}_{SR}$.

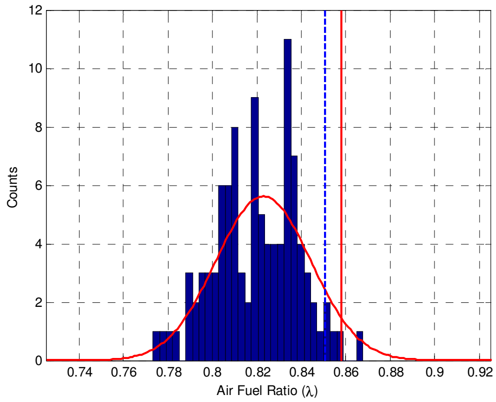

Figure 2: Statistical Results for Rich Failure Data

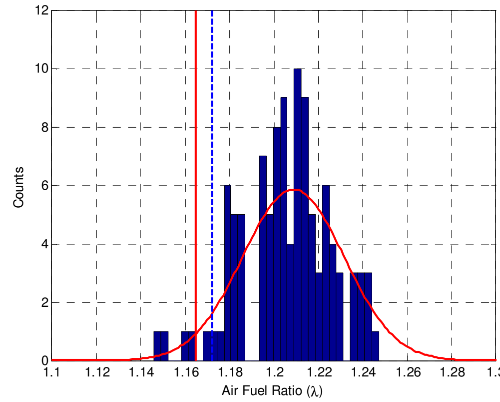

Figure 3: Statistical Results for Lean Failure Data

	L_R	S_R	U_R
$\bar{x}$	0.8166	0.8197	0.8229
Σ	0.0214	0.0189	0.0170

Table 2: Statistical Model for Rich Failure Data ($\beta = 0.05$)

	L_L	S_L	U_L
$\bar{x}$	1.2022	1.2055	1.2087
Σ	0.0229	0.0203	0.0182

Table 3: Statistical Model for Lean Failure Data ($\beta = 0.05$)

Since we most concerned with the issue of falsely flagging a Fault Free System it is useful to take the difference between the Rich Failure Threshold at 0.8579 and the Rich Robustness Threshold at 0.9330 and then determine the amount of sigma separation there is between them. This difference is normalised by dividing by σ_{L_F} to determine the amount of separation in terms of sigma and results in a separation of 3.46 σ_{L_F}

This then provides a clear robust threshold in terms of provide a detection which meets the requirements and provides a significant safety margin against falsely flagging and we have also taken into consideration the variability in the models by making use of the confidence interval

Figure 3 shows the response after introducing a lean failure shift again the graph shows the histogram of the raw data and the solid red line shows the distribution of the worse case statistic model derived from $\bar{x}_{L_L}$ and σ_{L_L} in Table 3. The red vertical line shows the Lean Failure Threshold of 1.1647 from this model and the dotted blue vertical line at 1.1723 for derived from $\sigma_{S\,L}$ and $\bar{x}_{S\,L}$. Using the same metric as previously derived the differences between the Lean Failure Threshold and the Lean Robustness Threshold gives a separation of 4.64 σ_{L_F}.

From the discussion earlier in the paper it has been highlighted that it can difficult to collect Fault condition so it is useful to be able to assess whether enough data has been collected to ensure that the real distribution of the Fault Condition has been captured. To determine this we have to make the assumption that the Sample model information that we have is the best estimate of the distribution model. So in the Rich Failure case the Sample values of $\sigma_{S\,R}$ and $\bar{x}_{S\,R}$ from Table 2 have been used and the value of n and the subsequent changes made to the t-distribution and Chi-Squared, in equations (1) and (2), can then be used to determine the effect on the threshold calculation.

In Figure 4 the blue solid line shows how the difference between the Rich Failure Threshold and the Nominal Rich Failure Threshold, normalised against the value of $\sigma_{S\,R}$ determined at n = 102, decays as the amount of data increases. The Red dot shows the current point Rich Failure Threshold Delta. In Figure 5 shows the same metric as in Figure 4 but for the Lean Failure Condition.

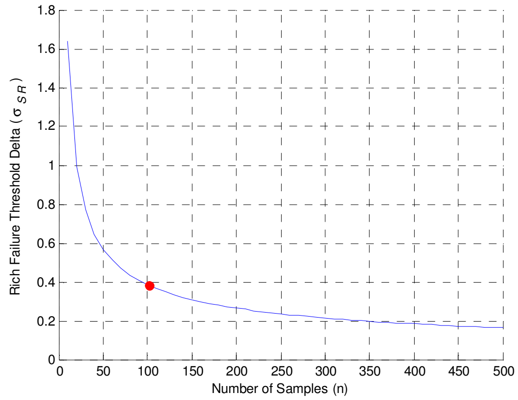

Figure 4: Rich Failure Threshold Delta

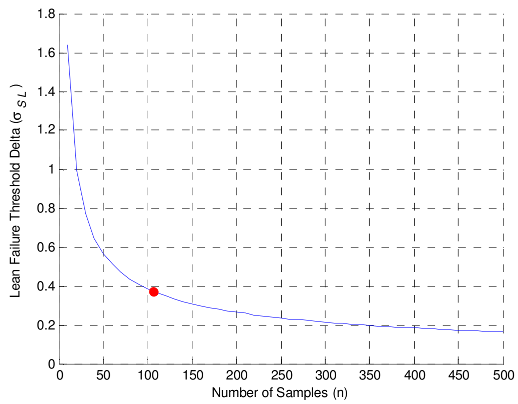

Figure 5: Lean Failure Threshold Delta

From the results shown in Figure 4 and 5 the current set of data obtained at around 100 samples gives an adequate confidence and subsequent testing would not necessarily lead to a significant change in the values of the Failure Thresholds. For example to reduce the current threshold, in both cases, by 50% would require an addition 250 tests.

	Robust Threshold	Failure Threshold	Separation
Rich	0.9330	0.8579	$3.46\,\sigma_{L_F}$
Lean	1.0640	1.1647	$4.64\,\sigma_{L_F}$.

Table 4: Summary of the Threshold Information

VI. CONCLUSION

The results in Table 4 show the calculated Failure Thresholds and the Robustness Thresholds for Rich and Lean conditions and for this diagnostic there is a significant amount of separation between these two sets of thresholds has enabling the requirements laid out in Section 3 to be met. This amount of separation has reduced the risk of making use of a Gaussian Model to derive the information for the Fault Free set of data. If it were the case that there was not such a separation then further investigation would have to undertaken to understand the reason or determine perhaps a more valid statistical model.

The use of the Confidence Interval in the statistical models and the threshold setting enables the calibrators to determine when they have collected enough data to provide a robust calibration.

ACKNOWLEDGEMENT

To Brian Varney and James Anderson for the inspiration and the data used in this work

REFERENCES

[1] P BALTUSIS. ON BOARD VEHICLE DIAGNOSTICS, SAE TECHNICAL PAPER 2004-21-0009, 2004.

[2] OBD-II TITLE 13, CALIFORNIA CODE REGULATIONS, SECTION 1968

[3] EOBD DIRECTIVE 98/69/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL OF 13 OCTOBER 1998

[4] R. STONE, INTRODUCTION TO THE INTERNAL COMBUSITON ENGINE, SOCIETY OF AUTOMOTIVE ENGINEERS, U.S., 3RD REVISED EDITION EDITION, ISBN 978-0-7680-0495-3

[5] J. FANTINI, J.F. BURQ, EXHAUST-INTAKE MANIFOLD MODEL FOR ESTIMATION OF INDIVIDUAL CYLINDER AIR FUEL RATIO AND DIAGNOSTIC OF SENSOR-INJECTOR, ELECTRONIC ENGINE CONTROLS 2003 (SP-1749), 2003-01-1059

[6] J.C. SMITH, C.W. SCHULTE, D.D. CABUSH, INDIVIDUAL CYLINDER FUEL CONTROL FOR IMBALANCE DIAGNOSIS, SAE INT. J. ENGINES, VOLUME 3 ISSUE 1, 2010-01-0157

[7] CONFIDENCE INTERVAL PRESENTATION - BLACK BELT TEACHING MATERIAL, 6 SIGMA ACADEMY, 2003

[8] H. W. LILLIEFORS. ON THE KOLMOGOROV-SMIRNOV TEST FOR NORMALITY WITH MEAN AND VARIANCE UNKNOWN. JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION. VOL. 62, PP. 399–402, 1967.

[9] H. W. LILLIEFORS. ON THE KOLMOGOROV-SMIRNOV TEST FOR THE EXPONENTIAL DISTRIBUTION WITH MEAN UNKNOWN. JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION. VOL. 64, PP. 387–389, 1969