



*Citation for published version:*

Brace, CJ, Burke, RD & Moffa, J 2009, 'Increasing accuracy and repeatability of fuel consumption measurement in chassis dynamometer testing', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 223, no. 9, pp. 1163-1177. <https://doi.org/10.1243/09544070jauto1084>

*DOI:*

[10.1243/09544070jauto1084](https://doi.org/10.1243/09544070jauto1084)

*Publication date:*

2009

[Link to publication](#)

Made available with publisher permission. Always cite as follows: Brace, C. J., Burke, R., Moffa, J., 2009. Increasing accuracy and repeatability of fuel consumption measurement in chassis dynamometer testing. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 223 (D9), pp. 1163-1177. Link to published article: <http://dx.doi.org/10.1243/09544070jauto1084>

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Increasing accuracy and repeatability of fuel consumption measurement in chassis dynamometer testing

C J Brace, R Burke\*, and J Moffa

Powertrain and Vehicle Research Centre, Department of Mechanical Engineering, University of Bath, Bath, UK

The manuscript was received on 5 December 2008 and was accepted after revision for publication on 19 May 2009.

DOI: 10.1243/09544070JAUTO1084

**Abstract:** The aim of this paper is to identify and investigate the effect of small changes in test conditions when quantifying fuel consumption. Twelve test set-up variables were identified and intentionally perturbed from a standard condition, including the effect of removing the power-assisted steering pump.

Initially a design-of-experiments (DoE) approach was adopted and the results showed that most of the tested parameters had significant effects on fuel consumption. Most of these effects were greater than the effect of typical technology changes assessed on chassis dynamometer facilities. For example, an increase of 8.7 per cent in fuel consumption was observed following a 90 min battery discharge from vehicle headlamps. Similarly an increase of 5.5 per cent was observed when the rig was run 3 km/h faster over a drive cycle, and 2.6 per cent when using tyres deflated by 0.5 bar. As a consequence, statistical tolerancing was used to suggest typical tolerances for test rig set-up variables. For example it was recommended that the tyre pressure be controlled to within 0.1 bar and the test rig speed to 0.3 km/h.

Further investigations were conducted into the effect of battery discharge, coast-down time, and engine cooling. These highlighted the need for rigorous battery charge management as the battery voltage was found not to be an appropriate measure of the variation in the alternator loading. Coast-down time was found to be a good control measure for a number of set-up variables affecting the rolling resistance of the vehicle. Finally the variations in the engine cooling were quantified using a cumulative engine temperature over a drive cycle. This was found to correlate well with fuel consumption. For each of these subsequent investigations, results were compared with the DoE predictions and found to agree well when considering the relatively low number of tests compared with the number of factors.

**Keywords:** chassis dynamometer, engine testing, repeatability, reproducibility, design of experiments, fuel consumption

## 1 INTRODUCTION

The increased costs associated with crude oil and the suspected impact of human activity on global warming are pushing research in automotive powertrains to search for more areas for fuel economy gains [1]. Improvements are likely to be made as a result of a series of small measures producing fuel

consumption benefits of the order of 1 per cent [2]. Bannister *et al.* [3] measured 3.5 per cent reduction in fuel consumption when comparing two oils of different grades over the New European Drive Cycle (NEDC). Another example is the desire to compare and rate different auxiliary units such as different designs of oil or coolant pumps. The differences between units on a fuel consumption basis is likely to be small, as the previous studies in this area have shown that fuel consumption improvements resulting from the removal of these units is of the order of 3 per cent [4].

\*Corresponding author: Powertrain and Vehicle Research Centre, Department of Mechanical Engineering, University of Bath, Claverton Down, Bath BA2 7AY, UK.  
email: en3rdb@bath.ac.uk

Faced with this challenge, measurement accuracy and repeatability on a chassis dynamometer, in addition to the reproducibility over different facilities, must be improved to be able to achieve credible results to demonstrate real improvements in fuel economy. It is accepted that an aim is to achieve a repeatability of about 0.5 per cent at 95 per cent confidence level. Two international standards [5, 6] suggest tolerances for some set-up parameters which shall be detailed in the following section where appropriate, although other parameters remain uncontrolled and could be a source of reproducibility inaccuracies.

This paper will attempt to discover the reasons for variability in testing on a chassis dynamometer. To achieve this, the study concentrated on identifying the effects of key set-up parameters on measured fuel consumption. The results from these tests will be useful in two ways: they will help to identify key areas that need to be controlled to increase repeatability for a particular laboratory, but also will highlight reasons for inconsistencies between different testing facilities. Parameters were intentionally varied to simulate the difference in set-ups. The exercise was split into two stages: initially a design-of-experiments (DoE) approach was used in a screening exercise to assess the key factors affecting

fuel consumption measurements; second, further testing using one factor at a time (OFAT) was conducted to better explain the DoE model.

## 2 EXPERIMENTAL APPROACH

### 2.1 Experimental set-up

Experiments were conducted on a chassis dynamometer where the test vehicle was a Ford Mondeo with common-rail diesel injection. A robot driver was used to minimize driver-induced variations. The tests employed six separate techniques for quantifying fuel consumption as described below and shown in Fig. 1:

- bag analysis: an industry standard method which consists of performing a carbon balance on collected exhaust gases over the cycle (the measurement of fuel consumption using feed-gas and tailpipe emissions were conducted while taking into account time alignment issues exposed in previous work by Hawley *et al.* [7] and Bannister *et al.* [8]);
- feed gas carbon balance: similar to the bag test, only performed continuously on pre-catalyst gases;

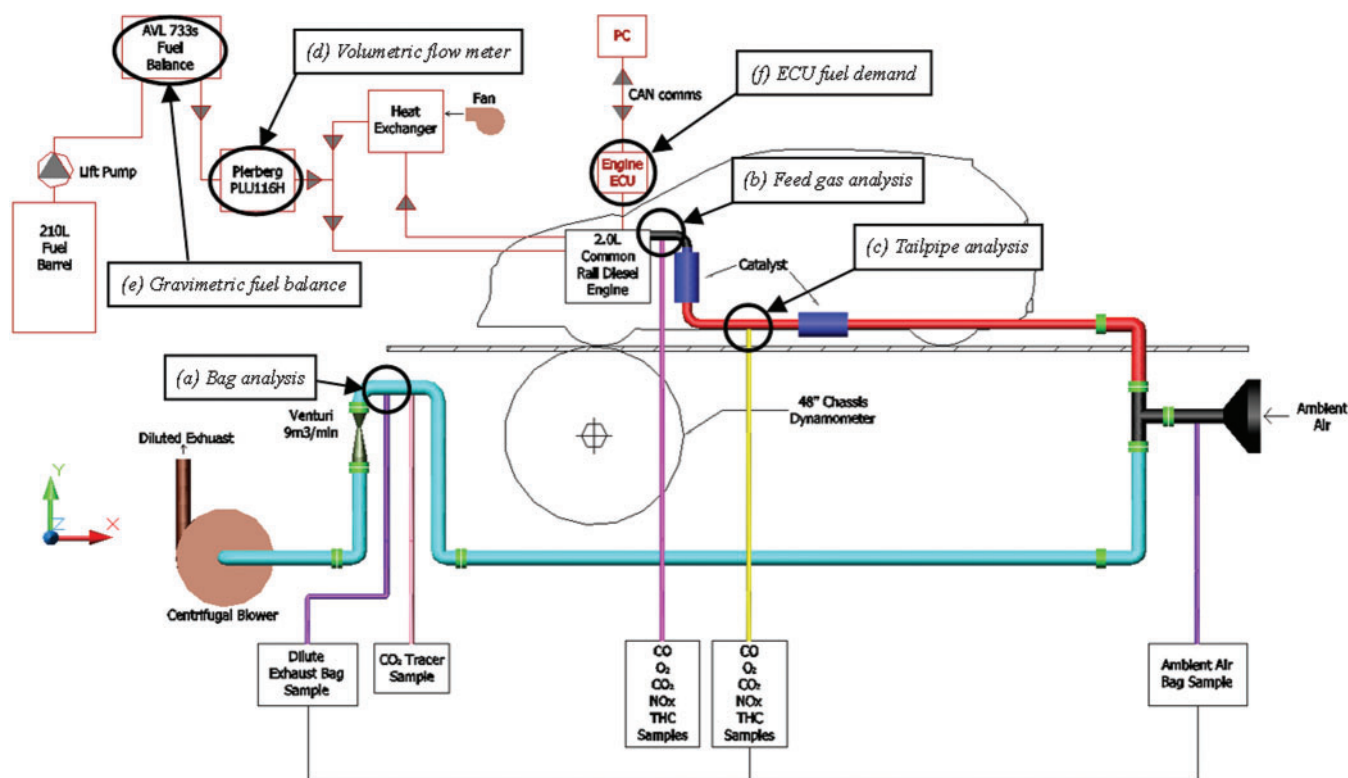


Fig. 1 Test cell arrangement (PC, personal computer)

**Table 1** Summary of the experimental factors and their two settings

Identification number	Description	Value for the following levels	
		Standard	Perturbed
V1	Battery state of charge	Normal	Headlamps on 90 mins
V2	Engine start temperature	-7 °C	-4 °C
V3	Engine oil level	Upper dipstick mark	Remove 2.5l
V4	Pedal busyness	Normal	Double-pedal activity
V5	Speed error	None	3 km/h fast on cruises
V6	Road speed fan	Normal	+40% overspeed
V7	Vehicle alignment	0	75 mm offset
V8	Tie-down straps	Horizontal	Angled
V9	Tyre type	Production	Sports
V10	Tyre pressure	Normal	Low
V11	Vehicle mass	1479 kg	1617 kg
V12	PAS pump	Production	Removed

- (c) tailpipe gas carbon balance: as above but on post-catalyst gases;
- (d) volumetric fuel flow meter (Pierburg PLU 116H);
- (e) gravimetric fuel mass flow meter (AVL 733s);
- (f) electronic control unit (ECU) fuel demand.

## 2.2 Testing factors

Following an initial phase to establish the accuracy of various fuel consumption measurement techniques, a series of tests was conducted to assess the error in fuel consumption measurement induced by a poorly controlled set-up of the test rig. To do this, 12 set-up parameters were intentionally varied and the effect on fuel consumption over the NEDC was measured. Because of the large number of factors to test, a DoE approach was adopted. A two-level fractional factorial design was used to reduce experimental effort. Each parameter was assigned a nominal level and a perturbed setting that could result either from an error in set-up or from differences between laboratory standards. The 12 factors and their levels are described below and summarized in Table 1. While the perturbations may seem excessive in some cases, these describe the conditions that allow better exploration of the test design space.

### 2.2.1 Battery state of charge

The state of charge was determined by measuring the voltage at the start of the test. Tests were performed with a full state of charge, or following a discharge having headlamps on for 90 min before testing.

### 2.2.2 Engine start temperature

The ECU-measured engine temperature at the start of the test was used. The cell temperature was used as a nominal setting and a setting 3 °C higher as a perturbed value.

### 2.2.3 Engine oil level

The engine was filled to upper dipstick mark for the standard setting, and the effect of removing 2.5 l was assessed.

### 2.2.4 Pedal busyness

This is defined as the cumulative rate of change in the pedal position over a complete cycle. It may be calculated by first taking the derivative of the pedal position to give the rate of change in the pedal position. The absolute values of this derivative are then summed up over the test cycle to give pedal busyness. Although the value has little physical meaning, it gives insight into any oscillatory behaviour in the pedal activation. This value would be influenced by the driver or by the robot driver control algorithm. It is not obvious how to set a standard for every test installation for this parameter, and so the tolerance will be presented as the percentage change from the optimum set-up. As it was not clear how much of an effect this parameter would have on fuel consumption, the control algorithm was modified to induce twice as much pedal activity as in the baseline set-up.

### 2.2.5 Speed error

If the rig has a speed error, the vehicle will be driven more quickly or more slowly than the desired speed

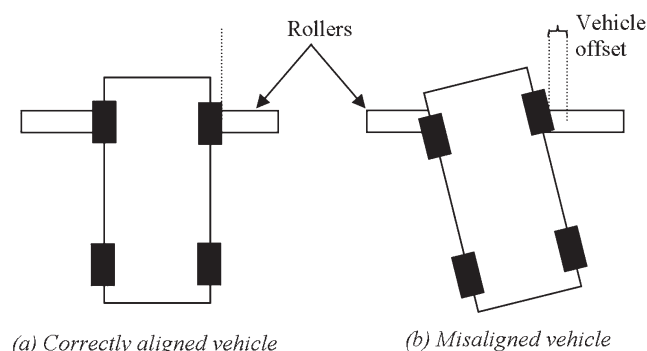
and will burn more or less fuel. A speed error could be the result of a poorly controlled test rig but may also be the result of comparing results from different laboratories that both work within the standard testing tolerance. The international standard [5] specifies a tolerance of  $\pm 0.5$  km/h or  $\pm 1$  per cent, whichever is greater.

### 2.2.6 Road speed fan

The road speed fan used to simulate the flow of air over the vehicle and to ensure adequate engine cooling was run at 40 per cent over speed. (In the standard condition, the cooling fan was calibrated to match the measured top hose coolant temperatures during road tests.) This would induce excessive cooling and was expected to increase fuel consumption. To the best of the authors' knowledge, no standards exist on setting this parameter, which could cause inconsistencies when comparing results from different facilities.

### 2.2.7 Vehicle alignment

This refers to the alignment of the vehicle with the rollers about the yaw axis (Fig. 2). Should the vehicle not be aligned correctly it was expected that the rolling resistance would increase and hence fuel consumption. While the international standard [5] specifies that the vehicle be restrained in a safe manner, there is no reference to the alignment of the vehicle.



**Fig. 2** Measurement of vehicle misalignment. Vehicle offset is measured by the difference between (a) the position of the vehicle tyre wall in the correctly aligned case and (b) the position of the vehicle tyre wall in the misaligned case. This value is measured in millimetres

### 2.2.8 Tie-down straps

The tension in these straps will have an effect on the rolling resistance. It is common to tie these horizontally to ensure that no extra downward force is applied to the vehicle. The effect of angling these straps was to be investigated. It was expected that the increased force would increase the fuel consumption by increasing the rolling resistance.

### 2.2.9 Tyre type

Fuel saver tyres offering increased fuel economy are commonplace on vehicles today. Low-profile sports tyres were to be used to assess the increase in fuel consumption.

### 2.2.10 Tyre pressure

The effect of a low tyre pressure was to be tested. It was expected that this would increase the fuel consumption by increasing the rolling resistance. Requirements on tyre pressures [5] refer only to the safe operation of the test rig and not to the accuracy of results.

### 2.2.11 Simulated vehicle mass

The simulated vehicle inertia was raised by 138 kg by adjusting dynamometer settings. This was designed to show the impact of an inaccurate calibration or the use of approximate inertias on older flywheel systems. (The simulated vehicle mass is a parameter programmed into the dynamometer for correct simulation of accelerations and braking manoeuvres. On a modern test rig this will be achieved by electrical control of the rolling road dynamometer. Some older systems use a series of flywheels to achieve the inertia value, and an exact inertia value may not be achievable.)

### 2.2.12 Power-assisted steering pump

As a reference, the power-assisted steering (PAS) pump was removed. It is unlikely that this would be done in error, but, as a common accessory to most vehicles, it serves as an interesting reference point for comparison.

Table 1 shows a summary of the factors and their two levels: standard, representing the expected set-up on a test rig, and the perturbed set-up, causing the error in fuel consumption.



### 2.3 Design-of-experiments approach

A full factorial design for the 12 factors would have required 4096 ( $2^{12}$ ) separate experiments. This is obviously not practical and a fractional factorial design was used. It was decided to produce an experimental design using 32 tests. This was chosen as it is the smallest number of experimental runs required to be able to estimate the main effects of each factor independently from two-way interactions. The design obtained is called a  $2^{(12-7)}$  design, which is of resolution IV. (The resolution of a fractional factorial design is a description of the generating relation which then defines the alias structure of the design. In a resolution IV design, main effects are confounded with three-way interactions and higher; however, two-way interactions are confounded between themselves [9].) The alias structure of this design is such that the main effects are confounded only with three-way interactions or greater. This means that it is not possible to distinguish between the measurement of a main effect and high-level interaction terms from the experiments. However, by assuming that third-order interactions are negligible, the main effects can be identified. With this design, the second-order interactions are confounded between themselves, and so it is not possible to estimate any interactions without further experimental work.

Climatic conditioning dictated the engine start temperature, and the 32 tests were split into two identical series of 16 tests: 16 at  $-7^{\circ}\text{C}$  starting temperature and 16 at  $-4^{\circ}\text{C}$  starting temperature. Table 2 describes the 16 test schedules for the remaining 11 factors. For each factor a blank entry indicates the standard setting and a letter P indicates the perturbed setting, as described in Table 1. The test sequence was not chosen at random, meaning that any effects arising over time could not be identified; however, this did allow hardware changes to be minimized. Test 1 was repeated halfway through the programme and at the end to check for drifting during the programme and to give an idea of variability in the measurements.

Because of the nature of the factors described above, it was not possible to achieve the exact settings described in Table 1. This meant that measured values of factors were used in the subsequent analysis as opposed to the desired values listed above.

### 2.4 Further testing

Following the DoE approach, further tests were conducted to confirm some of the findings and to

investigate the effects further. This approach used an OFAT approach and the factors studied are listed as follows.

1. Increased alternator load. Two levels of battery discharge were to be studied following a 45 min and a 90 min discharge from headlamps. Two current clamps were fitted to the vehicle, one measuring alternator current and one measuring net current to the battery.
2. Increased rolling resistance. The effect of tyre type, tyre pressure, and vehicle alignment were combined to investigate the overall effect on rolling resistance. The tests were to be compared on the basis of coast-down time, measured for each test.
3. Excessive engine cooling. This was to be assessed by a combination of increasing the road speed fan and opening the car bonnet. To increase further the effect on fuel consumption, the fan was run at 180 per cent speed up to a vehicle speed of 100 km/h, after which the fan was saturated.

As in the case of the DoE testing, full NEDCs were conducted on the same vehicle and the tests compared on the basis of total fuel consumption.

## 3 FITTING THE RESPONSE MODEL

### 3.1 Model description

Owing to the limited number of tests and the large changes in test set-up, very little information can be obtained by looking at the raw data, and hence a response model was fitted to the data. As described previously, the experimental design limited the analysis to only the main effects, although sufficient data were available to give an idea of the error associated with the measurements. Four models were generated using the MATLAB model-based calibration toolbox for the response of different fuel consumption estimates. Each model assumed a linear relationship and no interaction terms. The following analysis concentrates on the gravimetric fuel consumption measurement, although the processes were similar for all responses.

### 3.2 Statistical analysis of responses

The main effects of the factors of the experiment when changed from the standard condition to the perturbed condition are shown in Fig. 3 with 95 per cent confidence intervals. Negative changes in fuel consumption show cases where the perturbed

**Table 2** Experimental test matrix (blank entries represent a factor in the standard condition; entries labelled P represent the perturbed condition)

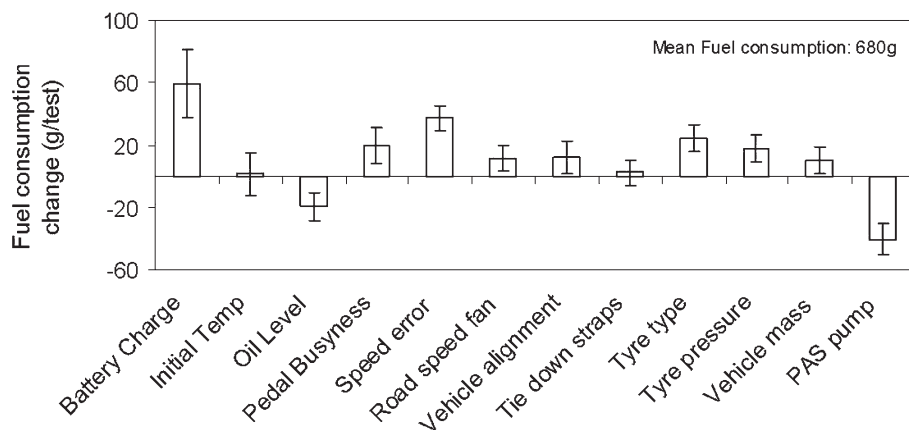
Test	Standard or perturbed condition for the following factors										
	V1	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
1											P
2			P	P					P	P	P
3	P	P	P	P			P	P		P	P
4	P	P					P	P	P		P
5	P	P		P	P	P					P
6	P	P	P		P	P			P	P	P
7				P	P	P	P	P	P		P
8			P		P	P	P	P		P	P
9	P		P		P			P			P
10	P			P	P			P	P	P	
11		P	P		P		P		P		
12		P		P	P		P			P	
13		P				P		P	P	P	
14		P	P	P		P		P			
15	P					P	P			P	
16	P		P	P		P	P		P		
17											P

set-up resulted in reduced fuel consumption compared with the baseline. In addition, the effects are tabulated with 95 per cent and 99 per cent confidence intervals in Table 3. This table also shows both the change in fuel consumption as a percentage of the mean and the statistical significance of each effect based on standard error.

The regression model shows that only two of the considered factors are insignificant: the engine start temperature (V2) and tie-down straps (V8). All other variables were significant at 95 per cent, and seven factors were found to be significant at 99 per cent (see Table 3).

The level of fit of a regression model can be assessed by the coefficient of determination,  $R^2$ , which is a measure of the differences between the

fitted model and the measured data points: the closer the  $R^2$  value is to 1, the better is the fit. In this case the fit was very good and gave an  $R^2$  value of 0.95, meaning that 95 per cent of the variability in fuel consumption could be described by the 12 factors considered. An analysis of the predictive power of the model was also conducted through the predicted residual sum of squares (PRESS). (PRESS analysis is conducted by fitting a series of regression models to 31 of the 32 data points and assessing the error of that model's prediction with the removed value. The computed PRESS  $R^2$  value assesses the potential of the model to predict points that have not been measured. This analysis is also often used to identify over-fitting where inaccurate models can be obtained with high  $R^2$  values but low PRESS  $R^2$



**Fig. 3** Main effects of factors on gravimetric fuel consumption measurement and 95 per cent error bars. Positive effects represent an increase in fuel consumption in the perturbed condition, and negative effects a reduction in the perturbed condition. The mean fuel consumption is the statistical mean from the regression model and serves only to put into context the measured changes

**Table 3** Main effects and confidence intervals of the response model (factors in bold represent effects significant at the 95 per cent confidence level)

Identification number	Factor	Fuel consumption change (g)	95% confidence level range (g)	99% confidence level range (g)	Fuel consumption change (%)	Significance	
	Description					95%	99%
V1	<b>Battery discharge (V)</b>	59.2	±21.6	±29.5	8.7	Yes	Yes
V2	Engine start temperature	1.2	±13.6	±18.5	0.2	No	No
V3	<b>Engine oil level</b>	-19.6	±8.9	±12.2	-2.9	Yes	Yes
V4	<b>Pedal busyness</b>	19.4	±11.1	±15.2	2.8	Yes	Yes
V5	<b>Speed error</b>	37.1	±8.3	±11.4	5.5	Yes	Yes
V6	<b>Road speed fan</b>	11.5	±8.4	±11.5	1.7	Yes	No
V7	<b>Vehicle alignment</b>	11.8	±10.5	±14.4	1.7	Yes	No
V8	Tie-down straps	2.0	±8.4	±11.5	0.3	No	No
V9	<b>Tyre type</b>	24.3	±8.4	±11.5	3.6	Yes	Yes
V10	<b>Tyre pressure</b>	17.6	±8.3	±11.4	2.6	Yes	Yes
V11	<b>Vehicle mass</b>	9.9	±8.4	±11.5	1.5	Yes	No
V12	<b>PAS pump</b>	-40.6	±9.7	±13.2	-6.0	Yes	Yes

values.) While the  $R^2$  approach quantifies how well the regression model fits the acquired data, the PRESS  $R^2$  value quantifies how well the model performs in predicting responses, which is key to the current study. The present model has a PRESS  $R^2$  value of 0.84, indicating good predictive power of the model. (As with the  $R^2$  value, there is no absolute definition of a key value for the PRESS  $R^2$ . The closer the value is to 1, the better the model can predict the data.)

To have confidence in the regression results it is important that all the variables are independent. A way of estimating the dependence of variables is the correlation coefficient, and these are shown in Table 4 for all independent variables. All coefficients in this work are below 0.45 and the highest values are given in bold. (The correlation coefficient is a measure of the relationship between two variables. If there is a strong correlation between two inputs, it will not be possible to distinguish which factor the resulting effect is attributable to [9]. While the significance of an absolute value of correlation coefficient is highly dependent on context [10], and analysis of absolute value is not often discussed by most researchers, it is often accepted that a correlation coefficient greater than 0.8 or less than -0.8 indicates a strong relationship; greater than 0.5 or less than -0.5 a fair amount of correlation; and below 0.2 or above -0.2 a very weak correlation [11].)

The variables with the highest correlations are as follows:

- battery state of charge and engine oil level;
- battery state of charge and vehicle alignment;
- battery state of charge and PAS pump;
- vehicle alignment and PAS pump.

There is no obvious physical reason why there should be any correlation between these factors, and their higher correlation coefficients could be simply down to chance.

Finally the standard deviation of the residuals was found to be 1.6 per cent, which shows there is still a degree of random error in the process, or effects attributable to factors not included in the model. This was larger than the standard deviation of the measurement process, which was known to be 0.8 per cent, but was still an acceptable precision. This suggests that, when a large number of factors are varied, the system is subjected to larger random errors and this should be taken into consideration when quoting measurement accuracy. Random errors could be suppressed by the use of repeated runs, but this has to be considered against increased testing costs, time, and drift in the test set-up. The methods explained in this paper do not help to reduce systematic error, which can only be improved by good practice and rigorous testing procedures and good baseline checks.

Figure 4 shows the main effects for all four models for the different fuel consumption estimation techniques. These are all coherent and thus improve confidence in the test results and models. Table 5 shows the summary statistics for each model. This shows that all models are equivalent, with a small offset in average fuel consumption for the feed gas carbon balance method.

#### 4 ANALYSIS OF RESULTS AND TOLERANCE SETTINGS

Figure 5 presents the results as a percentage change and sorts them into descending order. Absolute



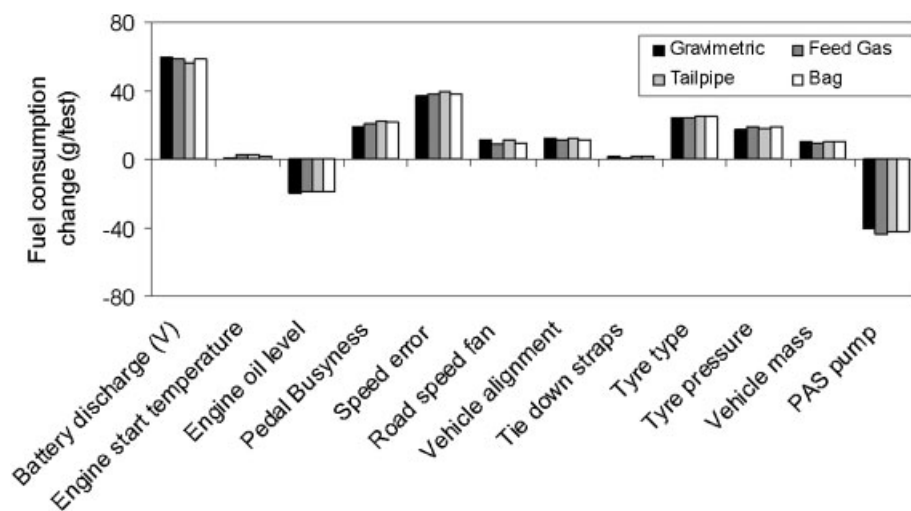
**Table 4** Correlation coefficients between independent variables (strongest correlations are in bold)

Correlation coefficient											
V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
V1	-0.03	<b>-0.36</b>	0.10	0.05	-0.07	<b>-0.36</b>	-0.20	-0.19	0.05	0.08	<b>-0.44</b>
V2		-0.07	-0.14	0.16	-0.01	-0.14	-0.08	-0.10	0.09	-0.07	0.13
V3			0.02	-0.03	0.05	0.03	0.08	0.08	-0.02	-0.03	0.11
V4				-0.06	0.03	-0.01	0.02	-0.06	-0.03	0.16	-0.11
V5					-0.01	-0.04	-0.03	-0.02	0.03	-0.01	0.00
V6						-0.17	0.01	0.03	0.00	-0.02	-0.03
V7							0.09	0.03	-0.02	0.09	0.39
V8								0.05	-0.01	-0.01	0.08
V9									-0.01	-0.02	0.06
V10										0.01	0.00
V11											-0.02
V12											

values of fuel consumption are used here rather than the measured values as it was assumed that the effects would be symmetrical about the mean for the bands over which tolerancing was to be applied. Also included in this figure is the result from a previous study on various oil properties, showing the effect of an increase in the high-temperature high-shear (HTHS) value of 0.6 cP. (The HTHS value is a measure of the oil viscosity at a temperature of 150 °C and a shear of  $10^6 \text{ s}^{-1}$ . One method of measuring this value is by studying the flowrate and pressure drop of a flow of oil through a capillary tube. This measure of viscosity in these conditions is thought to be representative of an automotive bearing engine under a high load [12]. As this value increases, it is expected that fuel consumption will increase as a result of increased friction in the engine. In this case, the effect serves as a good example of a typical desired measurement.)

It can be seen that, apart from the two factors not deemed statistically significant, all factors have an effect greater than the change in the oil HTHS value. It is also of interest to note that the effects of the battery voltage, speed error, tyre set-up, oil level, and pedal control are similar to the fuel consumption induced by the PAS pump.

Figure 5 clearly shows that all the areas investigated require careful control if accurate and repeatable testing is to be achieved. This implies that care should be taken when setting the parameters. One way of achieving this is through setting tolerance boundaries for the set-up of each factor. This was to be achieved by statistical tolerancing based on the DoE results. The DoE model was a linear model where the resultant output is the summation of the main effects of each of the inputs. The standard deviation (SD)  $\sigma_y$  of the output may therefore be expressed as a function of the SDs of the output under the effect of the input parameters ( $\sigma_1, \sigma_2, \dots$ ,



**Fig. 4** Main effects for four fuel consumption measurement techniques. Results are quoted in grams per test change, with respect to the mean fuel consumption for each method (which differs slightly depending on the method; see Table 5)

**Table 5** Summary statistics for all four fuel consumption measurement techniques

Measurement technique	Mean	R <sup>2</sup>	PRESS R <sup>2</sup>	SD of errors (%)
Gravimetric	680	0.95	0.84	1.6
Feed gas carbon balance	659	0.95	0.86	1.6
Tailpipe carbon balance	676	0.95	0.86	1.6
Bag test	677	0.95	0.86	1.6

$\sigma_n$ ) according to [13]

$$\sigma_y = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2} \tag{1}$$

It was then assumed that each factor would be toleranced to the same level, meaning that

$$\sigma_1 = \sigma_2 = \sigma_n = \sigma \tag{2}$$

Equation (1) therefore reduces to

$$\sigma_y = \sqrt{n\sigma^2} \tag{3}$$

where  $n$  is the number of parameters and  $\sigma$  is the desired SD of the output, under the effect of the deviation of the input.

It was desired to achieve a repeatability of 0.5 per cent at 95 per cent confidence level; therefore it was assumed that two SDs of the output are 0.5 per cent. (This is based on the assumption that two SDs include 95 per cent of the population in normal distribution.) Hence,  $\sigma_y$  is required to be no more than 0.25 per cent. Since the analysis assumes that the inputs are also normally distributed, Bender [13] suggested reducing this by a factor of 1.5, and  $\sigma_y$  becomes 0.17 per cent to take into account any underestimates. This is referred to as *Benderizing* and is applicable to situations where the process will vary over the target value over a long period of time, as will be the case in this testing [13].

The SD of the inputs may then be obtained for each of the appropriate variables. The *statistical tolerance* imposed on these variables is then calculated as three SDs of the input  $\sigma_y$ . (It is important to bear in mind that statistical tolerance implies not only that the value is between the two limits but also that it is normally distributed about the mean or nominal setting.) Wherever applicable, the resultant tolerances of the inputs are expressed in Table 6.

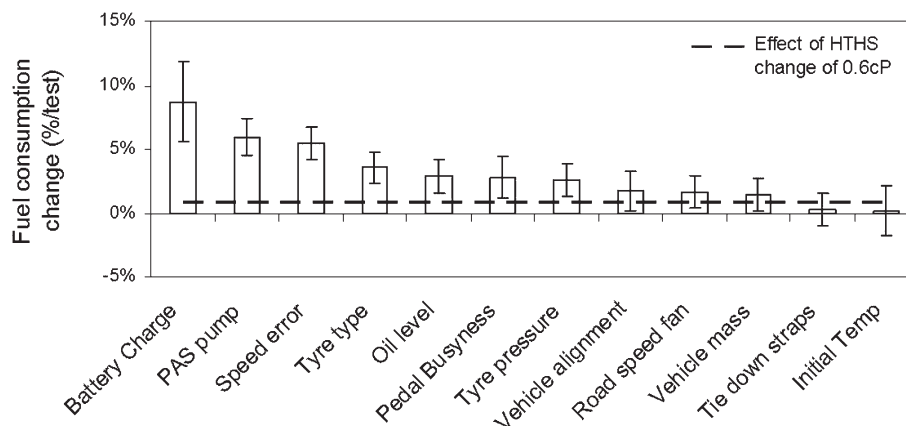
It is interesting to compare the results with the tolerances suggested in international standards.

1. The tolerance on the speed error is  $\pm 0.3$  km/h in this study compared with the standard of  $\pm 0.5$  km/h.
2. The tolerance on the tyre pressure is  $\pm 0.1$  bar and  $\pm 0.5$  bar in this study and the international standard respectively.

It is important to note that the tolerances in this study are tighter than those in common practice, which means that the current tolerances may be compromising repeatability and reproducibility.

### 5 FURTHER INVESTIGATIONS USING OFAT METHODS

Following the discoveries from the DoE screening exercise, further investigation was conducted on the effect of the battery charge, the combined effect of



**Fig. 5** Absolute effect on fuel consumption measurement ordered and compared with the typical effect of change of oil’s high-temperature high-shear value (HTHS) from 2.9 cP to 3.5 cP

**Table 6** Recommended set-up tolerances for parameters studied

Factor		
Identification number	Description	Tolerance
V1	Battery discharge	$\pm 0.2$ V
V2	Engine start temperature	Insignificant*
V3	Engine oil level	$\pm 0.45$ l
V4	Pedal busyness	$\pm 25\%$ <sup>†</sup>
V5	Speed error	$\pm 0.3$ km/h
V6	Road speed fan	$\pm 20\%$ <sup>‡</sup>
V7	Vehicle alignment	$\pm 25$ mm <sup>§</sup>
V8	Tie-down straps	Insignificant*
V9	Tyre type	N/A <sup>¶</sup>
V10	Tyre pressure	$\pm 1.4$ lbf/in <sup>2</sup> (0.1 bar)
V11	Vehicle mass	$\pm 50$ kg
V12	PAS pump	N/A <sup>¶</sup>

\*Parameters labelled Insignificant do not require tolerancing as their effect on fuel consumption was not found to be statistically significant.

<sup>†</sup>Pedal busyness is assessed as a percentage of the nominal value measured for this test rig. By expressing the value as a percentage, the results may be transferred to other facilities where this phenomenon may be measured by other means.

<sup>‡</sup>Fan speed is set in accordance with the road speed, and the speed of the fan has been calibrated to achieve realistic top hose coolant temperatures. This tolerance indicates the permitted variation on this setting throughout the drive cycle.

<sup>§</sup>This refers to the offset of tyre wall compared with a correctly aligned vehicle as described by Fig. 2.

<sup>¶</sup>Parameters labelled N/A are qualitative parameters that cannot easily be toleranced using the methods described in this paper. Methods for quantifying their effects need to be implemented before sensible tolerances can be applied.

parameters affecting rolling resistance, and the effect of excessive engine cooling.

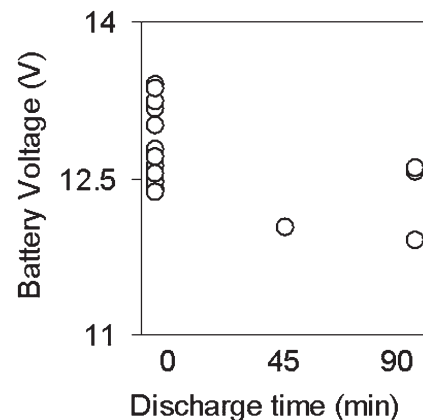
### 5.1 Battery discharge

Figure 6 shows a typical spread of battery voltage before each test for three states of charge of the battery. This clearly shows that the spread of battery voltage can be very large for similar states of charge, but also that a large reduction in battery charge may not be shown by a linear change in the battery voltage. This shows that the battery voltage is not an accurate measure of the state of charge, and other methods should be used to quantify this. This also suggests that the tolerance recommended for the battery voltage should be used with caution. This could explain the higher correlation coefficients associated with the battery state of charge, as the large degree of variation in the battery voltage, combined with the small number of experimental runs and large number of variables, could produce more accidental relationships.

In order to assess the effect of the state of charge of the battery, the currents to and from the battery were measured throughout a test cycle. This, together with the battery voltage measured by the ECU, allowed the electrical energy supplied to the battery over a test to be quantified. The results from this are shown in Fig. 7. A linear fit to the data appears to describe the relationship well with an  $R^2$  value of 0.92. In addition to this, the model predicts the fuel consumption with a 45 min discharge to within less than 1 per cent.

The observed increase in fuel consumption due to a 90 min discharge is 49 g (7.3 per cent) which is slightly lower than that predicted by the DoE approach, which was 59 g (8.7 per cent). However, this is within the error resulting from the DoE model. Another reason for this is that the DoE model was based on the battery voltage which has since been shown to be an inaccurate measure of the state of charge; therefore, taking this into account, the two methods agree well.

The power rating of the headlamps was 130 W, meaning that, over a 90 min discharge, around 700 kJ would be discharged from the battery. However, Fig. 7 shows that the energy supplied to the battery is in the region of 950 kJ. This shows that the penalty



**Fig. 6** Battery voltage at the start of the test for various discharge times. These results are from a separate series of tests to the DoE approach, but on the same test facility and set-up

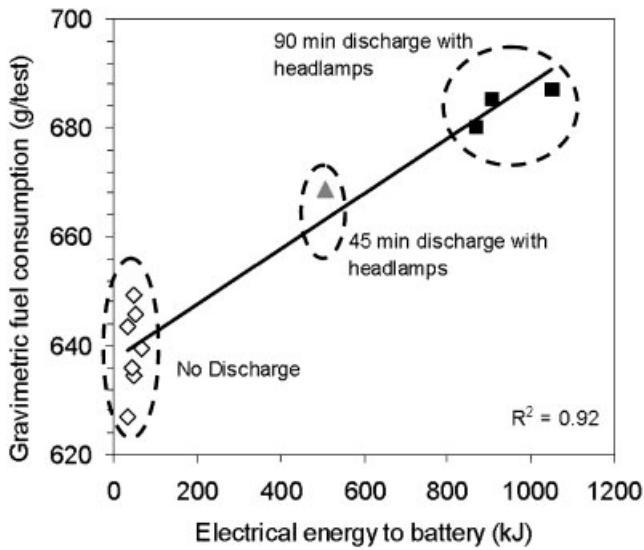


Fig. 7 Effect of energy supplied to the battery on gravimetric fuel consumption over the NEDC

for the discharge exceeds the apparent energy usage, which can be explained by the calibration of the charging control algorithm, battery irreversibility, and battery heating effects.

Figure 8 shows the instantaneous current to the battery over the NEDC for three different initial states of charge. It is clear the current to the battery is higher throughout the test as the initial discharge is greater. It is of interest to note that in both discharged cases the current to the battery is still higher at the end of the test, showing that the battery is still under charge. This has important implica-

tions, as the effect from the discharge could persist over multiple tests if no corrective action is taken.

This study has shown not only that the effect of battery discharge exceeds the apparent energy use, but also that it persists over a long period and hence possibly over multiple tests. In addition, the only precise way of assessing this is through the use of current clamps, which will only give a result after the test. This reinforces the need to maintain an effective battery management regime as a pre-emptive measure to ensure that the battery is fully charged for all tests.

### 5.2 Rolling resistance

To allow a direct comparison between the effects of vehicle alignment, tyre types, and pressures, the fuel consumption has been assessed against coast-down time in Fig. 9. The points are grouped into the results from the variations in the different parameters and explained on the figure. The results express a clear trend in increasing fuel consumption for reducing coast-down time, and the data are best fitted by a third-order polynomial, giving an  $R^2$  value of 0.97.

An increase in the fuel consumption of about 12 per cent is seen by the use of underinflated sports tyres with excessive tie-down force and misaligned vehicle. While it is unlikely that such conditions would occur in normal testing, it is interesting to note that in the area of normal operation a 10s reduction in coast-down time results in a 2 per cent change in the fuel consumption.

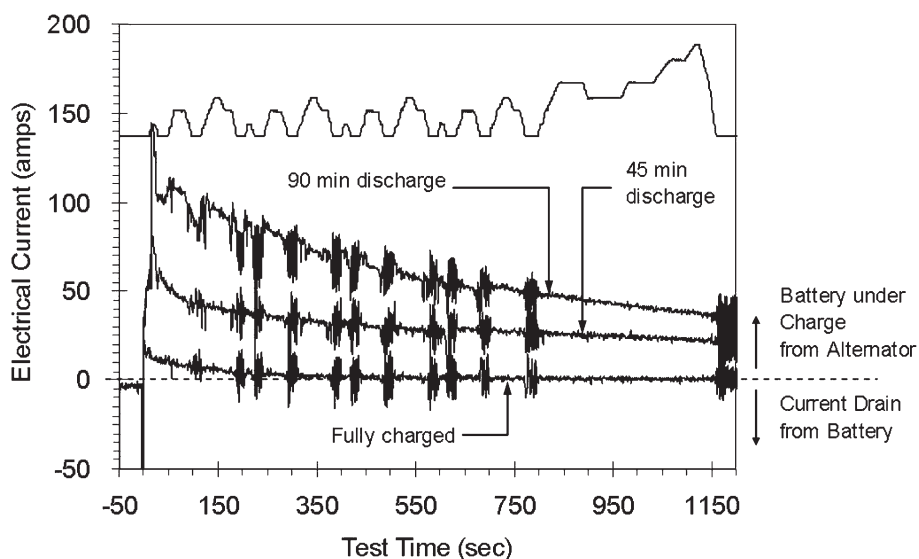


Fig. 8 Instantaneous battery current over the NEDC for three levels of initial battery charge. The NEDC is shown at the top for reference

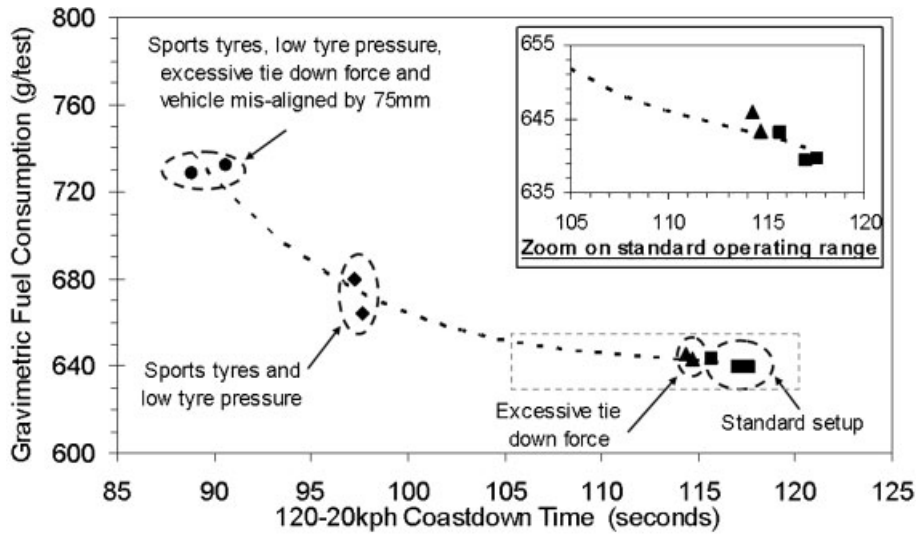


Fig. 9 Effect of coast-down time on gravimetric fuel consumption

Following this discovery, the results from the previous DoE screening design were reprocessed to incorporate the coast-down time instead of the tyre type, tyre pressure, tie-down straps, and vehicle misalignment. As the OFAT testing was best modelled through a third-order polynomial, terms up to the third order were included for the new variable coast-down time. Figure 10 shows the coefficients and their associated 95 per cent confidence bars. It can be seen that the inclusion of the coast-down time replacing the tie-down straps, vehicle alignment, tyre type, and tyre pressure does not have a significant effect on the other factors in the model. All changes in the coefficients between the models

are within the error bars from the previous model. The exception to this is the simulated vehicle mass (labelled Vehicle mass in the figure) for which the effect has trebled. Errors associated with the coast-down time are large, but still of the same order of magnitude as the errors found for the initial model.

Figure 11 shows both the OFAT results and the predicted model from the DoE over the same operating range. 95 per cent confidence intervals are also shown for the DoE results. The two offer similar trends and agree to within 0.5 per cent for coast-down times between 105 s and 120 s (normal operating region on this test-bed set-up); however, the predictions differ by about 5 per cent for a coast-

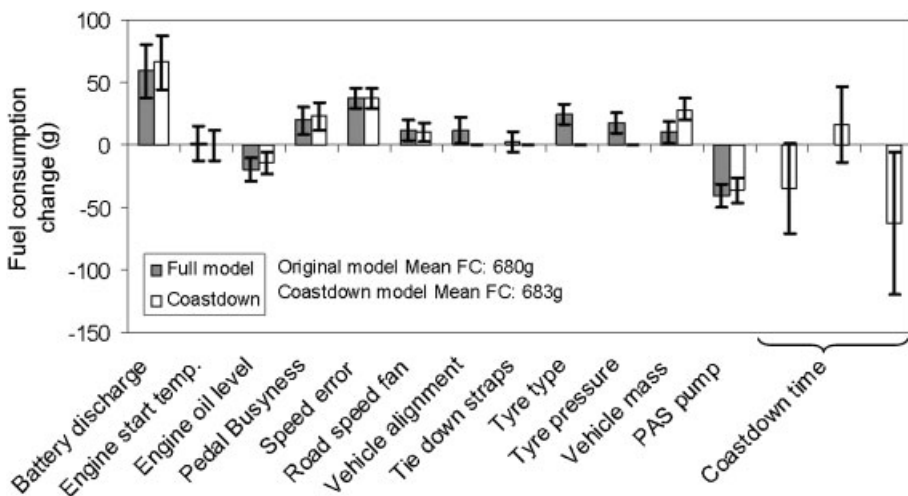


Fig. 10 Comparison of coefficients for both the full and the coast-down DoE model. The coefficients for coast-down time are shown for completeness and relate to the linear, quadratic, and cubic responses of the model



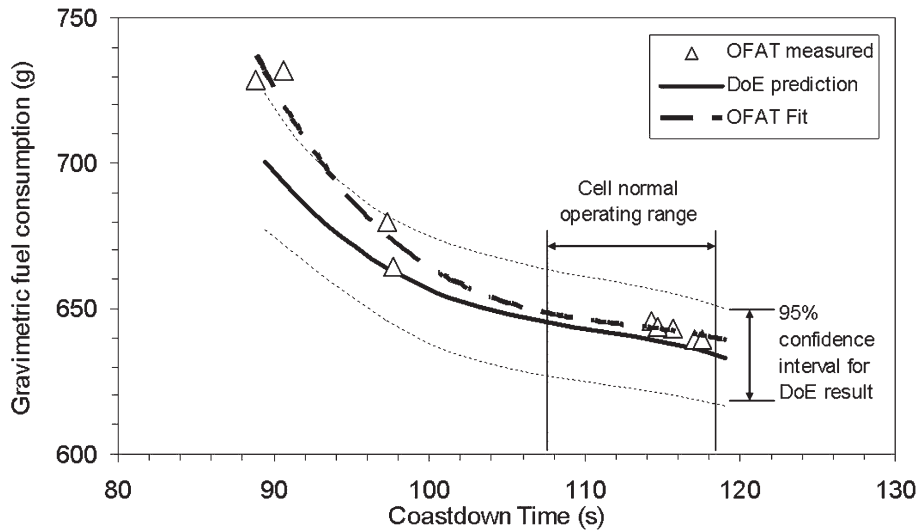


Fig. 11 DoE and OFAT model results for the coast-down time

down time of 90 s. This difference between the DoE model and the OFAT testing was thought to be due to the lack of test points in the initial test design for coast-down times less than 100 s (of the 32 tests run, 27 had coast-down times greater than 100 s and only five less than 100 s).

The new DoE model predicted a change of 58 g (7 per cent) for a reduction in coast-down time from 115 s to 90 s whereas the OFAT testing had measured 89 g (12 per cent).

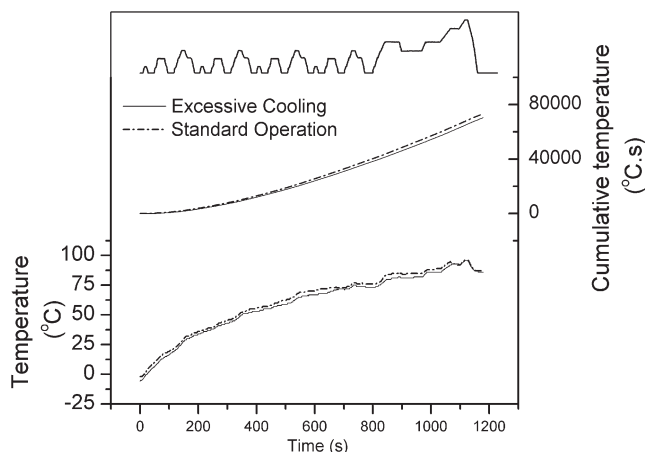
The results have shown that the coast-down time may be used as a good measure for quantifying the factors influencing the rolling resistance. Factors not easily measured such as the tyre type or effects of tie-down straps may be quantified and controlled. Based on the tolerance levels established earlier in this paper, the suitable tolerance for coast-down time was  $\pm 2.5$  s, when operating around the standard coast-down speed of 115 s. Obviously, while control over testing may be simplified through the measure of the coast-down time, all diagnostics will inevitably be solved through the four combined factors (the tyre pressure, tyre type, tie-down strap set-up, and vehicle alignment).

The tests indicated that a difference of 10 s in the coast-down time could result from a 0.5 bar (7 lbf/in<sup>2</sup>) change in tyre pressure. This highlights the need to implement a procedure for controlling the tyre pressure. It is also important to remember the variation in the tyre pressure as a result of tyre temperature, meaning that the tyre pressures should be checked at the same stage for every test, and preferably in the cold condition before testing.

### 5.3 Excessive cooling

No obvious correlation was observed between the final engine temperature and fuel consumption. This was thought to be because of the layout of the NEDC, which finishes with a high power cruise. During this time, the cooling fan speed is saturated, meaning that in all tests the cooling powers are the same, which results in similar engine temperatures. A better way to assess the effect of cooling was to look at the differences in temperature over the whole test cycle. This was achieved by assessing the integral of the engine temperature over the drive cycle. As can be seen in Fig. 12, for two tests where the final temperatures are identical, the cumulative temperature (integral) over the drive cycle is lower in the case of excessive cooling. Although the integral of the engine temperature has little physical meaning, it serves as a good way of quantifying and visualizing small temperature differences between the two test settings.

Figure 13 shows the gravimetric fuel consumption as a function of the cumulative engine temperature for various test set-ups. As would be expected, lower operating temperatures are correlated to increased fuel consumption, probably owing to increased oil viscosity and effects on the injection timing. There seems to be a strong relationship between the cumulative engine temperature and fuel consumption with a linear relationship, achieving a fit with an  $R^2$  value of 0.85. However, there is also a large spread of results for each cooling mode, which shows that the control over engine cooling is quite crude. The results show that careful control over the cooling



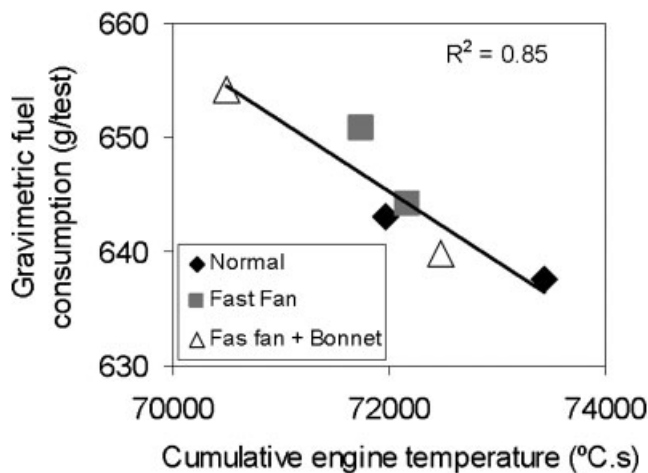
**Fig. 12** Engine temperature and cumulative temperature over the NEDC for two tests in standard conditions and with excessive cooling. The NEDC is shown at the top for reference

method should be installed but may not be sufficient to control fully the effects on fuel consumption measurement variability.

## 6 CONCLUSIONS

Nine factors in the chassis dynamometer set-up have been found to have a significant effect on the fuel consumption measurement: the battery state of charge, engine oil level, pedal busyness, speed error, road fan speed, vehicle alignment, tyre type, tyre pressure, and simulated vehicle mass. Four of these factors were then combined into a coast-down time.

All these factors were found to have effects of the same order of magnitude as the effect of removing



**Fig. 13** Effect of the cumulative engine temperature on the gravimetric fuel consumption over the NEDC

the PAS pump (6 per cent) and larger than a typical change in oil properties (0.9 per cent), two examples of typical chassis dynamometer testing objective. The response model had an  $R^2$  value of 0.95 and a PRESS  $R^2$  value of 0.84, quantifying the level of fit of the model. Models were consistent for different fuel consumption measurement techniques. The results included an 8.7 per cent increase in fuel consumption following a 90 min discharge using the headlamps, a 5.5 per cent increase when the rig was running 3 km/h faster on cruises, and a 2.6 per cent increase when tyres were deflated by 0.5 bar. Interestingly, the engine start temperature had no significant effect on fuel consumption.

The results were then combined with statistical tolerancing methods to suggest margins for the various set-up parameters. Suggested tolerances included setting tyre pressure to within 0.1 bar, simulated vehicle mass to within 50 kg, and controlling the roadspeed fan speed to within 20 per cent of calibrated value.

Additional testing produced using an OFAT approach proved consistent with the DoE model, thus increasing the confidence in these results. Further investigation into the battery condition has shown that the use of voltage is not an accurate measure of state of charge. A better way of assessing excessive engine loading is to monitor the energy flow into the battery using current clamps. This approach has highlighted irreversibility and that the effects of a battery discharge can persist over multiple test cycles. The combination of these findings leads to the strong suggestion that a rigorous battery management regime should be implemented to ensure that the battery remains fully charged for all tests.

Analysis of the rolling resistance on the basis of coast-down time showed a clear trend, and highlighted the need for good management in the tyre pressures to achieve repeatable testing. The coast-down time may be used as a control measure of the impact of factors influencing the rolling resistance.

The cumulative engine temperature proved a good basis for assessing the fuel consumption; however, this factor was not easy to control, as much variation was found for identical testing set-ups.

© Authors 2009

## REFERENCES

- Houben, M.** Engine test conditions to verify fuel and oil compositions improvements on emissions and fuel consumption. In Proceedings of the

- Fourth International Symposium on the *Performance evaluation of automotive fuels and lubricants*, 1993, SAE paper CEC EF 04 (Coordinating European Council, Earl Shilton, Leicester).
- 2 **Bartz, W. J.** Gear oil influences on efficiency of gear and fuel economy of cars. *Proc. IMechE, Part D: J. Automobile Engineering*, 2000, **214**(2), 189–196.
  - 3 **Bannister, C. D., Hawley, J. G., Brace, C. J., Pegg, I., Dumenil, J. C., and Brown, A.** Determining the effect of lubricating oil properties on diesel engine fuel economy. In *Proceedings of the Conference on Internal combustion engines: performance, fuel economy and emissions*, London, 11–12 December 2007, pp. 247–258 (Chandos, Oxford).
  - 4 **Akehurst, S., Hawley, J. G., Pegg, I., and Piddock, M.** Front end auxiliary drive (FEAD) configurations focusing on CO<sub>2</sub> benefits. SAE paper 2004-01-0596, 2004.
  - 5 BSI ISO 10521-2:2006 *Road vehicles. Road load. Part 2: reproduction on chassis dynamometer*, 2006 (British Standards Institution, London).
  - 6 BSI ISO 10521-1:2006 *Road vehicles. Road load. Part 1: determination under reference atmospheric conditions*, 2006 (British Standards Institution, London).
  - 7 **Hawley, J. G., Bannister, C. D., Brace, C. J., Cox, A., Ketcher, D., and Stark, R.** Vehicle modal emissions measurement – techniques and issues. *Proc. IMechE, Part D: J. Automobile Engineering*, 2004, **218**(8), 859–873. DOI: 10.1243/0954407041581057.
  - 8 **Bannister, C. D., Hawley, J. G., Brace, C. J., Cox, A., Ketcher, D., and Stark, R.** Further investigations on time alignment. SAE paper 2004-01-1441, 2004.
  - 9 **Box, G. E. P., Hunter, J. S., and Hunter, W. G.** *Statistics for experimenters. Design, innovation and discovery*, 2005 (John Wiley, Hoboken, New Jersey).
  - 10 **McPherson, G.** *Statistics in scientific investigations. Its basics, application and interpretation*, 2000 (Springer-Verlag, New York).
  - 11 **Mack, C.** *Essentials of statistics for scientists and technologists*, 1966 (Heinemann Educational Books, London).
  - 12 ASTM D4624-93 (1998). *Standard test method for measuring apparent viscosity by capillary viscometer at high temperature and high-shear rates*, 1993 (ASTM International, West Conshohocken, Pennsylvania).
  - 13 **Taylor, W. A.** Process tolerancing: a solution to the dilemma of worst-case versus statistical tolerancing, 1995, available from <http://www.variation.com/techlib/ta-2full.html> (last accessed 1 December 2008).

## APPENDIX

### Notation

DoE	design of experiments
ECU	electronic control unit
HThS	high-temperature high-shear
NEDC	New European Drive Cycle
OFAT	one factor at a time
PAS	power-assisted steering
PRESS	predicted residual sum of squares
$R^2$	coefficient of determination
SD	standard deviation