Technical Papers supporting SAP 2009

A meta-analysis of boiler test efficiencies to compare independent and manufacturers' results

Reference no.	STP09/B05
Date last amended	25 March 2009
Date originated	06 October 2008
Author(s)	John Hayton, BRE

Summary

Independent test programmes have been carried out on domestic boilers to compare efficiencies with those claimed by boiler manufacturers. This report brings together data from nine separate datasets from such programmes, covering over 100 boilers commonly sold in the UK market, to determine whether general conclusions can be drawn about differences between efficiencies claimed by manufacturers and those obtained from independent tests on the same products. The purpose is to inform the treatment of gas and oil boilers in the Government's Standard Assessment Procedure for the Energy assessment of Dwellings (SAP).

A statistical analysis has been undertaken to determine the likelihood of any differences between the manufacturers' and independent results occurring due to chance, and hence whether any adjustment to the seasonal efficiency of boilers calculated for SAP is necessary.

The findings are that on average the manufacturers' efficiency test values are significantly (statistically) higher than the independent test values, and that the average difference increases with higher efficiencies.

It is recommended that this upward bias of manufacturers' test results relative to independent test results is taken into account in SAP by applying a reduction to the manufacturer's values before they are used in the SEDBUK formulae for SAP. The recommendation applies to all domestic boilers fuelled by natural gas, oil, and liquefied petroleum gas (LPG).

Contents

1 Introduction

A number of independent test programmes have been carried out on domestic boilers bought from merchants "off the shelf" to compare efficiencies with those claimed by manufacturers. Such test programmes have been run by Defra[1] (the Market Transformation Programme), Defra (Climate, Energy Efficiency and Households Directorate), the Energy Saving Trust (EST), and the Consumers' Association (publishers of "Which?"). This report brings together data from nine separate sets (hence the term meta-analysis) of full and part-load efficiency measurements to determine whether general conclusions can be drawn about differences between efficiencies claimed by manufacturers and those obtained from independent tests on the same products. The purpose is to inform the treatment of gas and oil boilers in the Government's Standard Assessment Procedure for the Energy assessment of Dwellings (SAP).

A statistical analysis has been undertaken to determine the likelihood of any differences between the manufacturers' and independent results occurring due to chance, and hence whether any adjustment to the seasonal efficiency of boilers calculated for SAP is necessary.

In each of the data sets, the boilers were obtained anonymously from suppliers and efficiency tests at full-load load and part-load (direct and indirect method) were carried out by accredited laboratories. The full-load and part-load efficiencies used by manufacturers to claim SEDBUK figures in the Boiler Efficiency Database[2] were examined to make the comparison.

SEDBUK (Seasonal Efficiency of Domestic Boilers in the UK) is a calculation procedure that relies on efficiency test results as input data. Distortion or bias in the test results affects the SEDBUK figure. This study looks at the differences in the test efficiencies (not the SEDBUK value) before any capping or conversion from net to gross calorific values is applied in the SEDBUK method. All efficiency values here are quoted on a net calorific value basis.

The term "offset" is used in this report and is defined as:

the independent value - manufacturer's value.

A negative offset indicates the independent value is lower than the manufacturer's value.

The datasets include data from Defra and EST as well as those kindly supplied by "Which?" from work they commissioned in 2006, 2007 and 2008. Datasets are designated A to H, and J and cover 13 oil boilers and 99 gas boilers.

[1] Now DECC (Department of Energy and Climate Change)
[2] Used in SAP assessments and visible at www.sedbuk.com

2 Full-load efficiency

106 independently tested full-load efficiencies covering the nine datasets over the period 2003 to 2008 are collated and compared with the manufacturers' values obtained from the boiler efficiency database.

Figure 1 shows the differences between the manufacturers and independent datasets in the form of a "box and whisker plot" (or boxplot for short). The purpose of the boxplot is to highlight outliers, indicate the spread in the data values and show any differences between datasets.

Each boxplot consists of:

- a shaded box with boundaries that indicate the inter-quartile range and a centre line at the median
- Whiskers (the lines that stick out) that show the expected range of the data
- Outliers beyond the whiskers are indicated by asterisks. (None are evident in figure 1.)

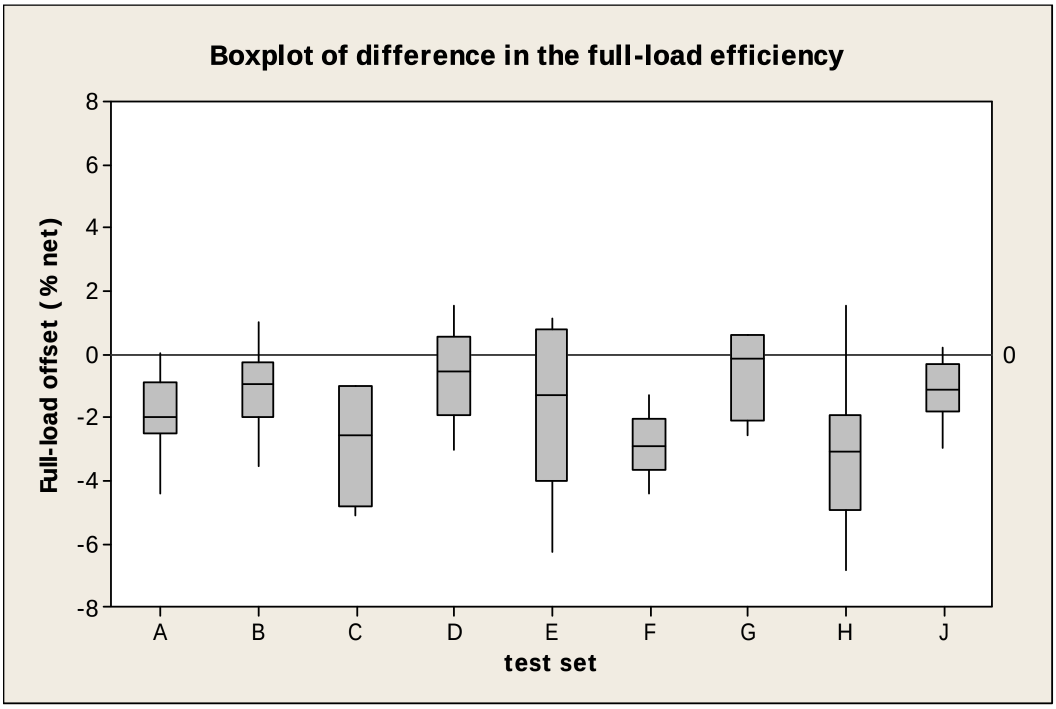

Figure 1 Boxplot of full-load efficiency offset indicating no outliers and similar datasets.

It is concluded from the boxplot of full-load offsets that (Figure 1):

a) There are no outliers

b) The median value of each dataset is negative (i.e. the independent value is lower than manufacturers value)

c) The interquartile ranges for all but three datasets span negative values

d) The expected range of each dataset overlaps with the ranges from all the other sets, and therefore it is reasonable to combine the dataset into one large set.

The offsets of the full-load efficiency are shown in the form of a histogram in Figure 2. It illustrates that the offsets are approximately normally distributed, which was confirmed by statistical test. The figure shows on average that the manufacturer's full-load efficiency is 1.7% net points higher than the independently measured value. As the data is approximately normally distributed it is concluded that, there is a 95% chance that the population mean of the difference lies between -1.4 and -2.0 net %points, and hence this is strong evidence that the manufacturer's value is higher than the independent value.

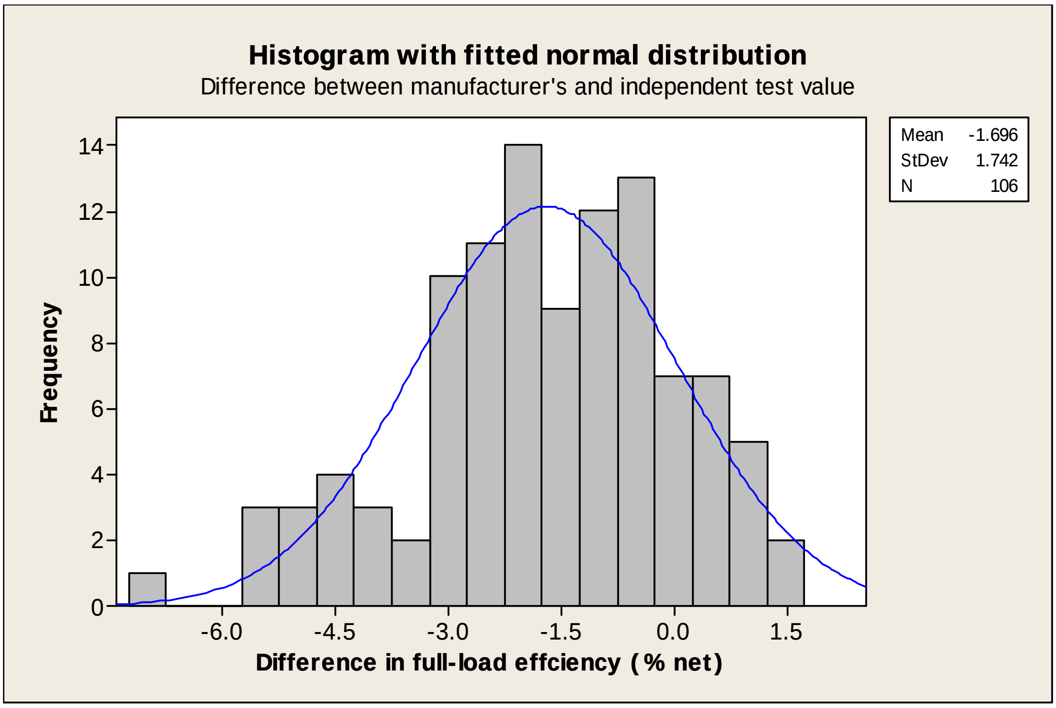

Figure 2 showing strong negative bias

The offset at full-load efficiency is plotted against the manufacturer's value to see if there is systematic bias (Figure 3). The scatter plot shows the negative offset increases on average with increasing efficiency.

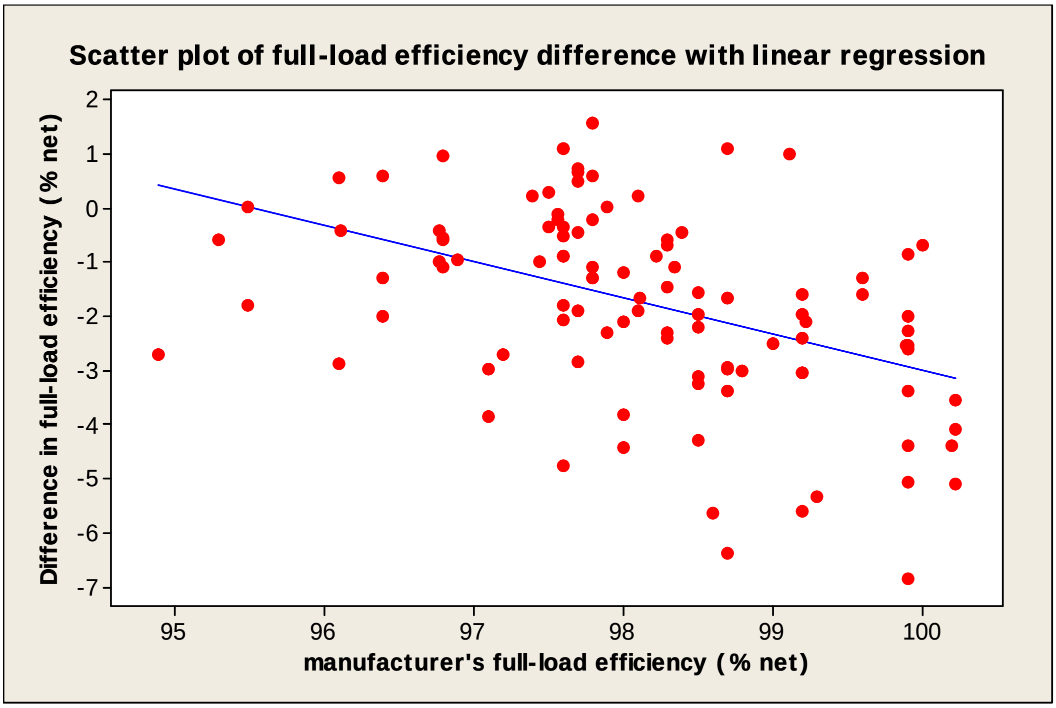

Figure 3 showing increasing negative bias of full-load efficiencies

Despite the large scatter in Figure 3, the increasing negative offset with increasing full-load efficiency is very unlikely to occur by chance alone. For example (see Table 1), the p-value is the probability due to chance that coefficient is zero and hence there is no linear relationship. Alternatively it can be said, with 95% confidence that the value of the coefficient lies between -0.41 and -0.93 and hence is very unlikely to be zero.

Table 1 Student t-test data for the full-load offset.

Parameter test	Coefficient (% net)	Standard error of coefficient (% net)	p-value
Constant	64.28	12.84	<.001
Manufacturer's full-load efficiency	-0.6729	0.1308	<.001

The linear regression prediction for a full-load efficiency (FLE) <u>above 95.5%</u> is estimated by:

Full-load offset (net) = -0.673 x (FLE(net) – 95.5%) eqn1

For full-load efficiencies below 95.5% net the offset is zero.

The highest manufacturer's quoted value in the sample is 101.2%; which would receive an offset of -3.8 net %points.

It is recommended that the offset of equation 1 is applied in SAP to account for bias in manufacturers' test results relative to independent test results.

3 Part-load efficiency

3.1 Part-load test methods

There are two principal ways of testing for efficiency at part-load in the laboratory: the direct and indirect method. The indirect method involves a steady test at the lowest firing rate and a standby test. The part-load efficiency is then estimated from the efficiency at the lowest rate minus a constant based on the standby heat loss test result.

The direct method determines the efficiency at part-load by setting the boiler to fire at 30% of the full-load over a number of ten minute cycles.

3.1.1 Comparison of part-load offsets

Table 2 compares the direct and indirect part-load efficiency offset and shows, on average, that manufacturer's value is 2.26 and 2.91 net %points lower than the direct and indirect measured independent value respectively.

Table 2 Comparison of part-load efficiency offsets (% net)

Difference between manufacturer's and independent value	Sample number	Mean	95% confidence interval of mean	Standard deviation
Direct method	97	-2.26	±0.568	2.854
Indirect method	54	-2.91	±0.800	3.012
Overall	151	-2.46	±0.465	2.913
Overall (minus outliers)	142	-2.00	±0.372	2.259

Is there a statistically significant difference between the direct and indirect efficiency offsets? To answer this question a two sample t-test that tests if means are equal is undertaken; assuming equality of standard deviations (additional tests confirm the latter assumption)[3]. A p-value of 0.274 is produced which is greater than 0.05 and so there is statistically no evidence to reject the proposition that the population means, direct offsets and indirect offsets are equal, hence it is concluded that they are equal and can be combined into one larger part-load offset sample.

[3] The F-test for equality of standard deviations shows no evidence to reject the proposition that the variances are equal and hence can be treated as equal, but is strictly only valid for a normal distribution. An additional test; a Levene's test is valid for any continuous variable (not only those normally distributed) and indicates that there is no evidence to reject that the variances are equal.

3.1.2 Part-load offset.

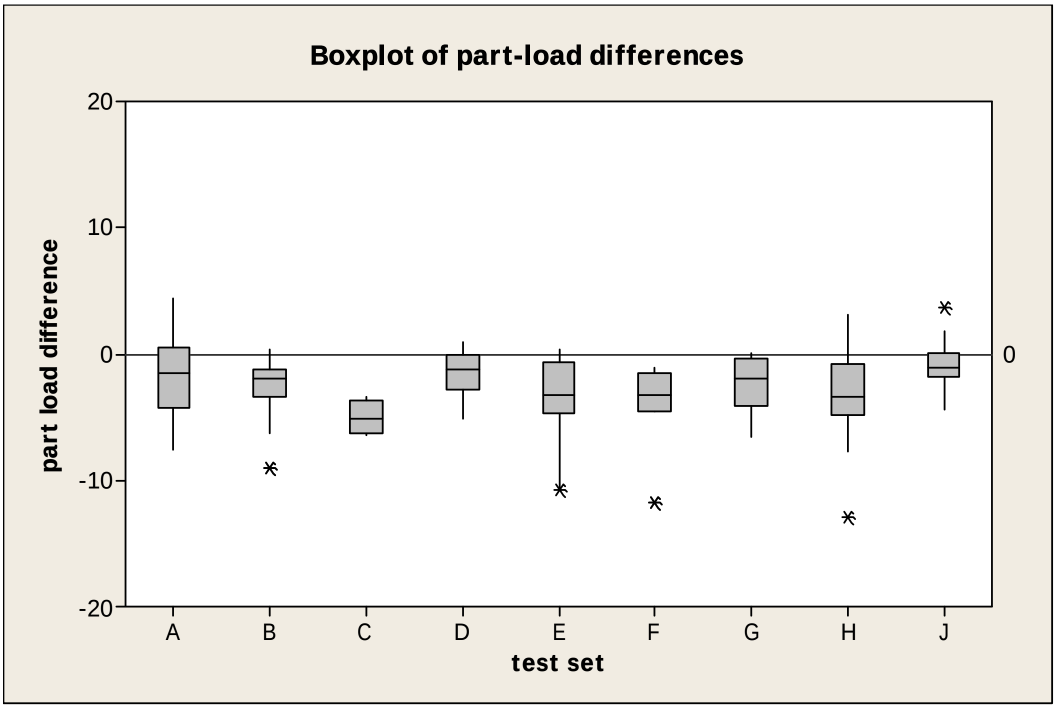

Figure 4: Part-load offset boxplot indicating five outliers but similar datasets.

It is concluded that from part-load boxplot (Figure 4) that:

a) There are five outliers (shown by the asterisks)

b) The interquartile range of all except dataset A spans negative values (i.e. the independent value is lower than the manufacturer's value).

c) Each median (centre lines) is negative.

d) The range of each dataset overlaps with all other datasets, and can therefore be treated as one large dataset

The 144 offsets of the part-load efficiency are shown in the form of a histogram in Figure 5 excluding the outliers noted above. It illustrates that the offsets are approximately normally distributed, which was confirmed by statistical test.

The histogram in Figure 5 showing the sample mean part-load offset is -2 net %points for 144 comparisons. That is, on average, the manufacturer's value is 2 %points higher than the independently measured part-load value. With 95% confidence, the population mean part-load offset lies between -1.6 and -2.4 %points.

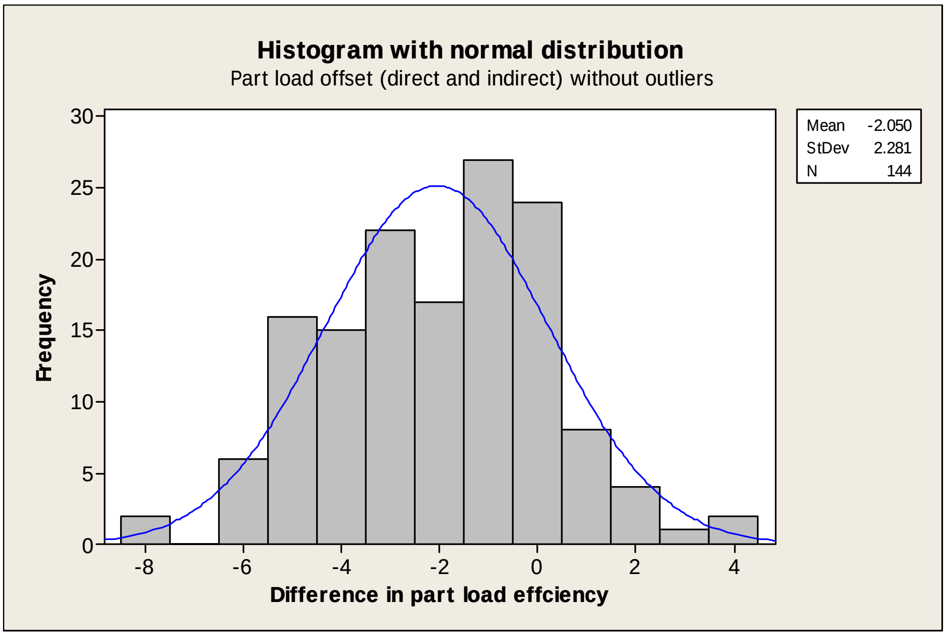

Figure 5 Part-load offset histogram showing a negative offset

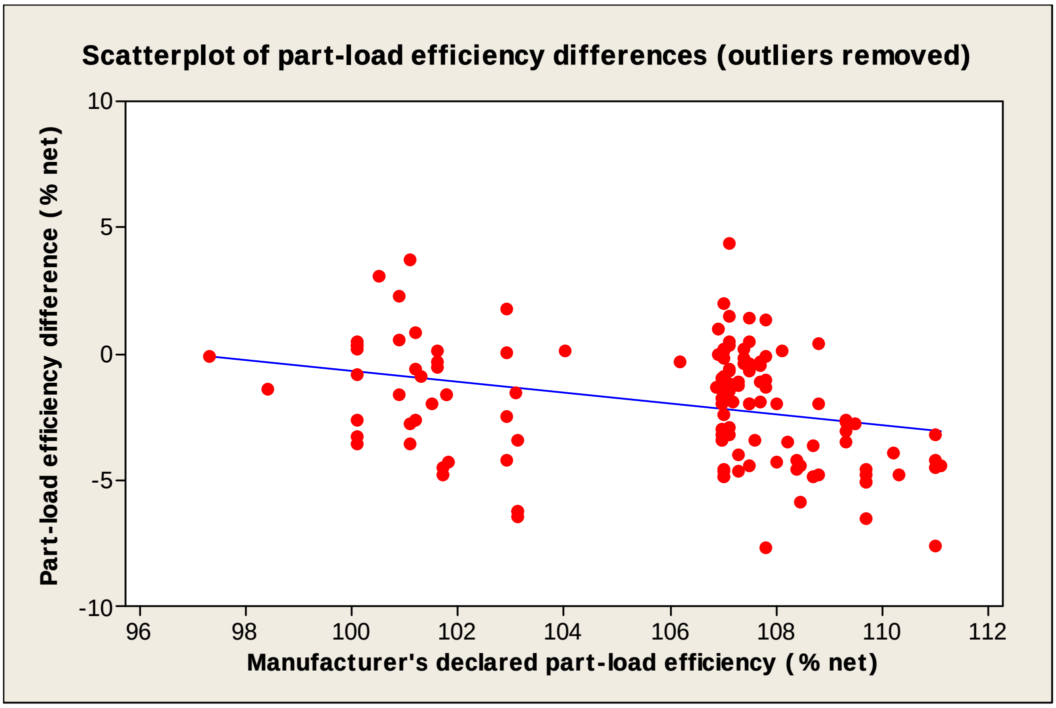

Figure 6 Part-load efficiency offset showing increasing negative bias with increasing efficiency

A scatter plot is drawn to test the relationship between offset and the manufacturer's value (see Figure 6). Figure 6 shows the increasing upward bias of the manufacturer's part-load efficiency with a large scatter. Despite the large scatter, the increasing negative part-load offset with increasing part-load efficiency is very unlikely to occur by chance alone. This increasing bias is statistically significant to 99.9%. That is, there is a 0.001% chance that this relationship occurs by chance alone.

The linear regression equation is:

For a manufacturer's part-load efficiency (PLE) above 96.6%:

Part-load offset (net) = -0.213 x (PLE(net) - 96.6% net) .... eqn 2

Values below 96.6% have an offset of zero.

The highest value quoted in the sample is 111% net. This would incur an offset of -3 %points.

It is recommended the offset in equation 2 is applied to the part-load (net) before applying any capping in the SEDBUK formulae used for SAP.

4 Discussion of results

There is strong evidence to show that the full- and part-load efficiencies claimed by manufacturers are higher than those from independent test programmes (sample sizes of 144 and 106 comparisons for the part-load and full-load data respectively). Furthermore, the differences increase, on average, at higher efficiencies. It can therefore be said that manufacturer results exhibit upward bias relative to independent test results.

If the purpose of tests was solely to compare different boilers the upward bias would make little difference. However, SAP is used to estimate energy consumption and carbon emissions in all types of dwellings, and must ensure that different heating technologies are treated equitably. An over-estimation of the efficiency of one type would confer a disadvantage on the others. Moreover any distortion may misinform the UK Government's energy efficiency policies.

The bulk of evidence analysed concerns gas boilers, and insufficient data was available to reach separate conclusions regarding oil and LPG boilers. However, as laboratory tests are carried out on the same rigs and the measurement procedure is essentially the same, there is no reason to suppose that bias would be different between gas, LPG or oil. The procedure used for measuring the amount of fuel differs but in other respects the test procedures and conditions are largely to be the same. The standards for measuring the amount of different fuels specify the same overall accuracy, and it is reasonable to conclude that the same bias will occurs for oil and LPG boilers.

Although the bulk of evidence concerns condensing boilers, the same test rigs and measurements are used for non-condensing boilers too. Similarly there is no reason to suppose any differences between tests on the different boiler types (e.g. combi boilers, storage combi boilers and regular boilers).

The test procedure for part-load efficiency can differ between on/off and modulating boilers. If boilers can modulate at 30% of the full output, the part-load test is carried out under steady conditions. Boilers that cannot modulate down to 30% are tested either under cyclic conditions (direct method) or by a combination of a steady test and a standby heat loss test (indirect method), so it could be argued that bias may vary with modulation capability. However, it was concluded that there was no significant difference in bias between direct and indirect part-load methods. This suggests the main contribution to bias is related to the measurement procedures or the boilers tested rather than varying firing rates during the test, so modulation is not expected to alter the bias.

5 Conclusions and recommendations

An analysis of 106 full-load and 144 part-load independent test results on domestic boilers commonly sold in the UK market has been compared to manufacturer's values submitted to the Boiler Efficiency Database. The main conclusions are:

1. The results exhibit an upward bias of manufacturer's test efficiencies relative to independently tested efficiencies;
2. The part-load test results confirm that here is no statistical difference between the differences based on the independent direct and indirect part-load measurements;
3. There is strong evidence that the upward bias increases with claimed efficiency.

Therefore it is proposed that efficiency test values received from manufacturers should be reduced before application to the SEDBUK formulae in SAP. This can be done either by applying a constant offset or an offset that varies with efficiency. In view of conclusion no. 3 the latter is recommended as it is considered fairer. It is recommended that:

1. For a manufacturer's full-load efficiency of 95.5% net or below no offset is applied.
2. For a manufacturer's full-load efficiency (FLE) above 95.5%, the offset applied is:
   Full-load offset (net) = -0.673 x (FLE(net) – 95.5%) eqn1
3. For a manufacturer's part-load efficiency of 96.6% net or below no offset is applied.
4. For a manufacturer's part-load efficiency (PLE) above 96.6%, the offset applied is:
   Part-load offset (net) = -0.213 x (PLE(net) - 96.6% net) .... eqn 2
5. These offsets are applied to all domestic boilers fuelled by gas, oil and LPG as the tests are all carried out using essentially the same equipment, methods and precision.