Guidelines for Field Performance Tests of Energy Saving Devices and Kitchen Performance Tests

(FTs - KTs)

Dr Adam Harvey, ClimateCare, Oxford UK

Dr Amber Tomas, Statistics Analyst, Oxford, UK

Contents

Introduction

This document currently focuses one specific type of field test, the measurement of fuel saved when cooks switch from inefficient to efficient stoves. This is known as the Kitchen Performance Test (KPT), also known as the Kitchen Test (KT).

The principles of the KPT apply to similar investigations of the performance of other energy-saving devices, and technologies and practices which displace decentralized thermal energy, such as solar water heaters, heat retention cookers, building insulation, biogas digesters, and so on. This document may therefore be used as a preliminary guide to performance tests for these technologies, although it does not claim to be comprehensive or thorough in this regard in the current version.

Proponents of projects disseminating a wider range of decentralized thermal energy technologies may therefore apply the principles described here, but it is their own responsibility to apply them appropriately to achieve accurate and conservative results.

KPT Procedure

To prepare and conduct a cook-stove KPT, follow these steps:

1. Estimate the number of test subjects you will be visiting (your SAMPLE SIZE). Sample sizes need to be larger if there is a lot of variation in the amounts of fuel used and saved, which is often the case in KPTs. One way to start is to simply assume a typical variation, expressed as a Coefficient of Variation or COV. Use the tables in the section below to choose a provisional sample size (a good starting point is to choose a mid-way value in the range given in the table below)[1]. Note that this assumes simple random sampling (if you use another method of sampling you will need to increase the sample size). If you choose a COV which is smaller than the real COV, it is likely that once you have finished the tests, you will need to increase the sample size[2]. Be sure to allow for "sample size attrition", that is drop-outs; if you launch 40 tests for example, you are likely to conclude with more than 30 valid results, even if some of the test subjects make mistakes or some of the tests are incomplete in some way. If previous experience shows a drop-out or attrition rate of 10% is likely, launch 10% more tests than suggested in the tables here. The validity of this approach depends on a wider range of factors than the COV alone, and therefore a minimum sample size of 30 is recommended (after drop-outs)[3].

2. Select the kitchens sampled using a RANDOM selection method. There are different ways of doing this, and it is up to you to choose an appropriate one that will give test results which reflect the real fuel savings of the project population[4].

3. Plan your tests so that they give a reliable and conservative result. In general, there are two phases, the BEFORE phase (before the improved stove is adopted, the baseline scenario) and

[1] The difference between independent, paired, and single sampling is explained in the next step of this procedure. Another way to estimate sample size is to find a study that has already been done in similar conditions (same type of socio-economic and cultural conditions) to your project, and learn from this the MEAN value of tests, and the STANDARD DEVIATION (SD). Dividing, you get the COV (= SD/MEAN), and then you can use the table on this page.

[2] This approach is legitimate if you can justify it by showing that the supplementary households are ones which could have appeared in the random sample originally or which otherwise qualify as consistent with the aim of a sample which is representative of the project stove users and their behavior through the year

[3] Users of the Gold Standard methodology Technologies and Practices to displace Decentralized Thermal Energy Consumption should refer to the latest version of the methodology document for mandatory requirements: one of these at time of writing these guidelines is that the minimum allowed sample size is 20 (it can be assumed this is after drop-outs, and also is dependent on normality and simple random sampling).

[4] Applicable common sampling approaches are outlined in Section III, Sampling Application Guidance, of the General Guidelines for Sampling and Surveys for Small-Scale CDM Project Activities (EB 50 Report, Annex 30)

the AFTER phase (after it is adopted, the project scenario). In some cases you may not need to test fuel consumption in both phases (possibly when using default factors for baseline stoves). If you are doing both phases, consider whether it is best to run both phases in the same kitchens (PAIRED sampling), or separate phases in separate sets of kitchens having appropriately similar socio-economic and cultural conditions (INDEPENDENT sampling). Larger sample sizes are required when using two independent samples, but independent sampling may be the only option if a fixed baseline has been established or if it is necessary to conduct the baseline and project tests concurrently. In the case of paired sampling, consider how long you need to give the subjects to get used to the project stoves before launching the AFTER tests, so that your test will reflect real usage patterns in forthcoming years. Also consider whether reversing the sequence for half the subjects is wise, to avoid biases that might occur due to time passing (i.e. run half the test in sequence AFTER then BEFORE, and half the other way). If a default efficiency is used for baseline stoves, then it is possible to run a KPT on the project stove only, and combine the results with either a value of project stove efficiency or a credible value for delivered cooking energy or for average baseline fuel consumption, so calculating fuel saved. This case is called a SINGLE-SAMPLE KPT.

4. Choose an appropriate test period and an appropriate time of year (or multiple times during the year). A recommended minimum test period is 3 days. It is important to avoid times like festivals or holidays when more cooking is done than usual[5], and if you do include days of home-cooking (for example weekends) when people are not at work and eating more than usual at home, you must make sure that they are balanced by an appropriate number of working days when people eat less at home. The same applies to tests which include cooks who sell their food publicly – these tests must include days when less food is sold as well as days when food sales are high, in an appropriate ratio and erring towards a conservative result. Think of ways of designing the test so that it captures a cooking pattern representative of a whole year. For example, this may involve carrying out some of the tests in another season of the year when eating patterns or food types are different, or prescribing a representative cooking pattern during a single test (this latter approach is known as a Controlled Cooking Test, a variation on the KPT).

5. Make sure that all test subjects understand they are expected to cook normally during the tests. The aim is to capture their usual behavior in the kitchen, as if no tests were happening, to feed the usual variation of people with the usual variation of food types. You are obliged to design into the project incentives for the elimination of inefficient stoves, which must be effective as fast as possible. Nevertheless remember that your tests must measure the fuel saved by the kitchen as a whole not by one individual stove; for example it is common for a cook to use one hob sometimes, and also an extra one or two hobs, at other times. Your project stove may be a two-hob design or a one-hob design, either way there is there is the possibility that an extra non-project hob or stove is occasionally used, in areas where project stoves are still a novelty and the incentive system for elimination of non-project stoves is still on-going.

[5] The KPT should exclude large parties or infrequent cooking events, and match cooking tasks in the baseline KPT to those in the project KPT.

6. To conduct the tests, make sure the cooks use fuel only from a designated stock which you have pre-weighed. Enter key data, such as the mass of the fuel at start of tests as stocked for each subject, in an excel form such as the ones provided in KPT guidelines referenced (see the first footnote of this section). It is recommended to visit the subjects at least once a day to check that they are using only fuel from the weighed stock, and are not adding un-weighed fuel to the stock. If more fuel is needed, weigh before adding and enter the mass added in the data sheet.

7. During the tests, also find out how many people have eaten and how many meals each, so that you can enter into the data sheet the number of "person-meals" (individual meals as opposed to meals shared) cooked with the weighed fuel each day. Note that this count can include meals sold commercially as well as meals consumed in the domestic environment.

8. For practical reasons, it is often best to provide fuel for the tests (to help control the weighing and use of fuel), rather than have the subjects use fuel they are buying themselves. Nevertheless, it is important that the fuel is typical of the fuel normally used through the year, particularly in terms of moisture content. It is also important that the subjects are paying for fuel, or have an incentive to conserve it, otherwise they may use excessive amounts due to the free hand-out. Subjects can be told they will be rewarded for their effort and time at the end of the test, once it is successfully completed.

9. Run a statistical analysis on the test results, to estimate the mean fuel savings.

If using data that has been collected as part of a separate study, make sure that the sample was selected randomly. If this is not the case then the data should not be used.

Before beginning the analysis, be sure to check for "outliers", i.e. values which are very different to the majority of the sample[6]. Outliers should be examined to check for mistakes with data recording, or investigated to ascertain if there were unusual circumstances which led to that result. If so, then the observation should be removed or corrected before the analysis. The distribution of sample values should also be checked for skewness. If there are extreme outliers or skewness, or the data was not collected by a simple random sample, then methods of analysis which are more complicated than the approaches suggested here may be required.

In cases of paired and independent sampling, there are two valid options for the statistical analysis:

a. 90/30 rule. This option allows you to calculate emission reductions on the basis of the estimated MEAN (or average) fuel saved by introduction of the improved stove in a kitchen. You can only use the mean if your test results satisfy the 90/30 rule (i.e. the end-points of

[6] One way to identify potential outliers is to produce a box-plot of the data. Most statistical software enables this. Any points which are plotted individually on the box-plot are candidates for outliers and should be investigated. Equivalently, potential outliers can be identified as those points which are either greater than 1.5 times the inter-quartile range (IQR) from the third quartile, or less than 1.5 times the IQR from the first quartile. This method can be implemented in Excel (assistance with excel applications may be sought from ClimateCare).

the 90% confidence interval lie within +/- 30% of the estimated mean). If this is not the case, then you can use the test data gathered so far to estimate how much larger the sample size needs to be. The mean value will always result in a larger estimate of fuel-savings than the value obtained using the second option below, but in some cases you might choose to analyze using the second option, because it is not practical or too expensive to increase the sample size sufficiently.

b. 90% confidence rule (Lower bound of the one-sided[7] 90% confidence interval). This option allows you to obtain a result even if 90/30 precision is not achieved, although in a similar manner to the 90/30 rule, a minimum sample size of 30 is recommended. You can use this approach when the 90/30 rule forces a sample size which is difficult to implement in practice. The disadvantage is that the fuel saving result is not the mean (or average) test result, but a lower value[8]. This estimate is very conservative, and it will probably be worthwhile to augment the sample size instead in cases when augmentation is practically possible.

The baseline and project KT data should be analyzed in combination, to estimate of the mean fuel saving. The options in 9a and 9b can then be applied. It is not valid to apply the above rules to estimate baseline and project fuel-use separately.

In cases of single samples[9], there are two valid options for the statistical analysis:

- 90/10 rule. This option allows you to calculate emission reductions on the basis of the estimated MEAN (or average) fuel use. You can only use the mean if your test results satisfy the 90/10 rule, i.e. the endpoints of the 90% confidence interval lie within +/- 10% of the estimated mean. If this is not the case, then you can use the test data gathered so far to estimate how much larger the sample size needs to be. The mean value will always result in a larger estimate than the value obtained using the second option below, but in some cases you might choose to analyze using the second option, because it is not practical or too expensive to increase the sample size sufficiently.
- 90% confidence rule (Lower bound of the one-sided[10] 90% confidence interval). This option allows you to obtain a result even if 90/10 precision is not achieved, although in a similar manner to the 90/10 rule, a minimum sample size of 30 is recommended. You can use this approach when the 90/10 rule forces a sample size which is difficult to implement in practice. The disadvantage is that the result is not the mean (or average) test result, but a

[7] The one-sided confidence interval is appropriate because it is relevant here to specify the confidence that the estimate is conservative, e.g. that the estimated fuel-savings are lower than (or to the low-side of) the true fuel-savings.
[8] Technically, it is the largest value that with a probability of 90% will be less than the true mean.
[9] In this case the sampling requirements (including precision of 90/10) for Small-Scale CDM Project Activities (EB 50 Report, Annex 30) are followed. The COV for fuel-savings is naturally higher than the COV for fuel-use, hence the precision requirement when using paired or independent samples to estimate fuel-saving has been set to 90/30, in order to achieve practically feasible sample sizes.
[10] The one-sided confidence interval is appropriate because it is relevant here to specify the confidence that the estimate is conservative, e.g. that the estimated fuel-savings are lower than (or to the low-side of) the true fuel-savings.

lower value[11]. This estimate is very conservative, and it will probably be worthwhile to augment the sample size instead in cases when augmentation is practically possible.

10. You may reward the test subjects once the tests are finished, for instance give them one or two project stoves or other compensation. Since you have already analyzed your data, you are in a good position to decide whether to extend the sample size further, or re-run tests that for some reason were invalid.

[11] Technically, it is the largest value that with a probability of 90% will be less than the true mean.

Sample sizes

Table 1: Sample sizes in cases of PAIRED samples (the same test subjects are sampled in the baseline and the project situations). Sample sizes for other values of COV can be calculated using equation (3) in Annex 1.

COV	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
90/30 precision	45	53	61	70	80	90	101	112	124

Table 2: Sample sizes in cases of INDEPENDENT samples (different test subjects are sampled in the project and baseline situations). This is the size required for each of the baseline and project samples[12]. Sample sizes for other values of COV can be calculated using equation (5) in Annex 1.

COV	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
90/30 precision	90	105	122	140	159	180	201	224	248

Table 3: Sample sizes in cases of SINGLE samples (where the tests are conducted for either baseline or project scenario but not both). Sample sizes for other values of COV can be calculated using equation (3) in Annex 1.

COV	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
90/10 precision	12	26	45	70	101	137	179	226	279

[12] In this table technically this is not the COV but the ratio of the population standard deviation for the amount of fuel used (this standard deviation is reasonably assumed here for simplicity to be equal for project and baseline populations) to the population mean fuel saving.

KPT Scope

Stove sizes.
In some instances, where this approach is justified as accurate and conservative, a project developer may assume that the size of the stove doesn't matter. For example, suppose you are selling two charcoal stoves, size 1 and size 2. A family or commercial cook cooking on Size 1 may feed less people than a family using Size 2. But they will use the same fuel per person-meal cooked. The KPT samples can in these cases include all sizes.

Commercial cooking.
In some instances, where this approach is justified as accurate and conservative, a project developer may assume that an identical quantity of fuel is used to cook a meal, regardless of whether the meal is sold or eaten by a family at home. In this case, the project stove fuel savings are a combination of savings made by cooks selling food (and often cooking for their families at the same time), and cooks providing meals at home only.

Different fuels.
If the project disseminates stoves using different fuels (e.g, charcoal and wood), it is not possible to conduct a single KPT on both fuels. Each fuel becomes what the methodology calls a "cluster" and separate KPT Are needed.

Data recording and analysis format

Annex 2 presents a simplified template for recording data from a KPT. In practice it is recommended to gather more data of various types, for example the kind of data collected during Baseline Surveys, and data which helps verify that that the notes above with regard to scope of KPTs and procedure, can be addressed in a test report and verified. Using Annex 2 as a basis, a project developer should develop a excel sheet which addresses points such as type of food, number of hobs, person-meals, date and season, location, working day or holiday, and so on, also adding data important to a report, such as name of test supervisor, weighing equipment used and its calibration status, and so on.

Example 1 here gives an example of how the data from a paired sample is analyzed, to obtain a mean value which is valid on the basis of 90/30 precision or as a lower bound of the 90% confidence interval. In this example, the sample size is adequate for adoption of the mean value.

Example 2 is also from a paired sample and shows data from the same number of subjects, but with a different COV. The result is that the sample size is inadequate for 90/30 precision, and the calculation shows how many more test subjects are needed, or what the result is if no more tests are conducted and the lower bound of the 90% confidence interval is adopted instead.

Example 3 shows data from two independent samples, i.e. non-paired data. In this example 90/30 precision is not achieved, so either the sample sizes need to be increased or else the lower limit of the 90% confidence interval can be used.

Example 4 shows data from two independent samples with the same number of subjects as Example 3, but in this case 90/30 precision is achieved and so the mean fuel saving can be used.

Examples 5 and 6 shows data from a single test providing a single set of data. For example, the test data may be from a Field Test sampling Project Fuel consumption, in a case where a default factor is being used to define a baseline; or it may be from a Field test sampling of baseline fuel consumption, where no test data for the project situation exists because the project is a zero-emission technology.

Example 1: Sample size is within required precision	SUBJECT I.D.	PROJECT FUEL	BASELINE FUEL	SAVED FUEL
	HH1	0.94	1.03	0.09
	HH2	1.05	1.39	0.34
	HH3	1.14	1.87	0.73
	HH4	0.94	1.14	0.19
	HH5	0.53	1.18	0.65
	HH6	0.82	0.94	0.12
	HH7	0.94	1.14	0.20
	HH8	0.64	1.16	0.52
	HH9	1.55	2.19	0.64
	HH10	1.14	1.52	0.38
	HH11	1.41	1.80	0.39
	HH12	1.05	1.89	0.83
	HH13	1.59	2.28	0.69
	HH14	0.78	1.33	0.55
	HH15	0.81	1.11	0.30
	HH16	0.69	1.17	0.49
	HH17	0.63	0.91	0.28
	HH18	1.01	2.04	1.04
	HH19	0.89	1.24	0.36
	HH20	0.92	1.17	0.25
	HH21	1.09	1.50	0.41
	HH22	1.08	1.62	0.54
	HH23	0.44	1.39	0.95
	HH24	0.84	1.03	0.19
	HH25	1.05	1.39	0.34
	HH26	1.14	2.02	0.88
	HH27	0.94	1.14	0.19
	HH28	0.43	1.38	0.95
	HH29	0.72	1.34	0.62
	HH30	0.94	1.24	0.30
	HH31	0.64	1.16	0.52
	HH32	1.55	2.46	0.91
	HH33	1.14	1.32	0.18
	HH34	1.41	1.80	0.39
	HH35	1.05	1.89	0.83
	HH36	1.59	1.81	0.22
	HH37	1.08	1.28	0.21
	HH38	0.81	1.36	0.55
	HH39	0.74	1.59	0.85
	HH40	0.67	1.14	0.46
			Mean fuel saving (Kg per stove per day)	**0.49**
			Standard Deviation (Kg per stove per day)	0.27
			COV	0.55
			What is the precison attained?	15%
			Does the sample size satisfy the 90/30 rule?	**YES**
			What fuel saving may be claimed?	**Use mean**
Note: All the calculations here assume			**Implied sample size for 90/30 precision?**	10
Simple Random Sampling has been			**How many additional samples are required?**	0
			90% Lower Limit (Kg per stove per day)	**0.43**

Example 2: Sample size is not within required precision	SUBJECT I.D.	PROJECT FUEL	BASELINE FUEL	FUEL SAVED
	HH1	0.75	1.03	0.28
	HH2	1.26	1.39	0.13
	HH3	0.95	2.12	1.18
	HH4	1.09	1.14	0.04
	HH5	1.14	1.38	0.24
	HH6	1.56	2.02	0.46
	HH7	0.04	1.14	1.10
	HH8	0.22	1.16	0.93
	HH9	1.32	2.66	1.34
	HH10	0.79	1.32	0.53
	HH11	0.99	1.80	0.81
	HH12	2.08	1.89	-0.19
	HH13	1.07	1.68	0.60
	HH14	0.38	1.33	0.95
	HH15	1.40	1.11	-0.29
	HH16	1.24	1.17	-0.07
	HH17	0.23	0.91	0.68
	HH18	2.27	2.04	-0.22
	HH19	0.46	1.24	0.78
	HH20	0.22	1.17	0.95
	HH21	1.46	1.20	-0.26
	HH22	1.83	1.62	-0.20
	HH23	0.87	1.39	0.53
	HH24	0.17	1.03	0.85
	HH25	0.79	1.39	0.60
	HH26	2.12	2.12	0.01
	HH27	0.99	1.14	0.14
	HH28	0.67	1.38	0.71
	HH29	0.91	2.02	1.12
	HH30	1.34	1.14	-0.20
	HH31	1.19	1.16	-0.03
	HH32	2.94	2.66	-0.28
	HH33	1.29	1.32	0.03
	HH34	1.85	1.80	-0.05
	HH35	1.86	1.89	0.03
	HH36	1.16	1.68	0.51
	HH37	0.09	1.28	1.19
	HH38	0.35	1.36	1.01
	HH39	1.22	1.59	0.38
	HH40	1.01	1.14	0.12
			Mean fuel saving (Kg per stove per day)	**0.41**
			Standard Deviation (Kg per stove per day)	0.49
			COV	1.20
			What is the precison attained?	32%
			Does the sample size satisfy the 90/30 rule?	**NO**
			What fuel saving may be claimed?	**Increase sample size or use 90% lower limit**
Note: All the calculations here assume Simple Random Sampling has been used			**Implied sample size for 90/30 precision?**	45
			How many additional samples are required?	5
			90% Lower Limit (Kg per stove per day)	**0.31**

Example 3: Two independent samples with failed precision	SUBJECT I.D.	PROJECT FUEL		SUBJECT I.D.	BASELINE FUEL	
	HH41	0.76		HH1	1.03	
	HH42	0.93		HH2	1.39	
	HH43	0.69		HH3	1.87	
	HH44	1.21		HH4	1.14	
	HH45	1.66		HH5	1.18	
	HH46	0.32		HH6	0.94	
	HH47	0.66		HH7	1.14	
	HH48	0.57		HH8	1.16	
	HH49	1.43		HH9	2.19	
	HH50	0.19		HH10	1.52	
	HH51	1.44		HH11	1.80	
	HH52	1.34		HH12	1.89	
	HH53	0.61		HH13	2.28	
	HH54	0.68		HH14	1.33	
	HH55	1.57		HH15	1.11	
	HH56	1.27		HH16	1.17	
	HH57	1.59		HH17	0.91	
	HH58	0.28		HH18	2.04	
	HH59	0.69		HH19	1.24	
	HH60	0.68		HH20	1.17	
	HH61	0.30		HH21	1.50	
	HH62	0.82		HH22	1.62	
	HH63	1.33		HH23	1.39	
	HH64	1.18		HH24	1.03	
	HH65	1.50		HH25	1.39	
	HH66	0.26		HH26	2.02	
	HH67	1.41		HH27	1.14	
	HH68	1.51		HH28	1.38	
	HH69	1.29		HH29	1.34	
	HH70	1.23		HH30	1.24	
	HH71	1.52		HH31	1.16	
	HH72	1.20		HH32	2.46	
				HH33	1.32	
				HH34	1.80	
				HH35	1.89	
				HH36	1.81	
				HH37	1.28	
				HH38	1.36	
				HH39	1.59	
				HH40	1.14	
Mean fuel use (Kg per stove per day)		1.00			1.46	
Standard Deviation (Kg per stove per day)		0.46			0.39	
COV		0.46			0.27	
sample size		32			40	
	Mean fuel saving (Kg per stove per day)					**0.45**
Note: All the calculations here assume Simple Random Sampling has been used	Standard Error (Kg per stove per day)					**0.10**
	What is the precison attained?					37%
	Does the sample size satisfy the 90/30 rule?					**NO**
*This is the number of additional samples required for each group, assuming equal additions.	**What fuel saving may be claimed?**					**Increase sample size or use 90% lower limit**
	How many additional samples are required*					79
	90% Lower Limit (Kg per stove per day)					**0.32**

Example 4: Two independent samples with precision achieved	SUBJECT I.D.	PROJECT FUEL		SUBJECT I.D.	BASELINE FUEL	
	HH41	1.00		HH1	1.49	
	HH42	1.18		HH2	1.54	
	HH43	0.87		HH3	1.66	
	HH44	1.12		HH4	1.36	
	HH45	1.13		HH5	1.59	
	HH46	0.84		HH6	1.58	
	HH47	1.07		HH7	1.70	
	HH48	0.80		HH8	1.29	
	HH49	0.87		HH9	1.69	
	HH50	0.85		HH10	1.38	
	HH51	0.70		HH11	1.26	
	HH52	0.77		HH12	1.62	
	HH53	1.14		HH13	1.59	
	HH54	1.09		HH14	1.43	
	HH55	1.17		HH15	1.52	
	HH56	0.72		HH16	1.54	
	HH57	0.97		HH17	1.42	
	HH58	1.17		HH18	1.72	
	HH59	0.90		HH19	1.64	
	HH60	0.74		HH20	1.21	
	HH61	0.97		HH21	1.61	
	HH62	1.16		HH22	1.52	
	HH63	0.91		HH23	1.50	
	HH64	1.07		HH24	1.31	
	HH65	0.89		HH25	1.43	
	HH66	1.07		HH26	1.56	
	HH67	0.93		HH27	1.36	
	HH68	1.07		HH28	1.46	
	HH69	1.06		HH29	1.60	
	HH70	1.07		HH30	1.72	
	HH71	0.98		HH31	1.38	
	HH72	1.05		HH32	1.40	
				HH33	1.22	
				HH34	1.59	
				HH35	1.50	
				HH36	1.23	
				HH37	1.21	
				HH38	1.45	
				HH39	1.52	
				HH40	1.48	
Mean fuel use (Kg per stove per day)		0.98			1.48	
Standard Deviation (Kg per stove per day)		0.14			0.15	
COV		0.15			0.10	
sample size		32			40	
					Mean fuel saving (Kg per stove per day)	**0.50**
Note: All the calculations here assume					Standard Error (Kg per stove per day)	**0.03**
Simple Random Sampling has been used					**What is the precison attained?**	11%
					Does the sample size satisfy the 90/30 rule?	**YES**
*This is the number of additional samples required for each group, assuming equal additions.					**What fuel saving may be claimed?**	**Use mean**
					Additional samples reqd for 90/30 precision*	0
					90% Lower Limit (Kg per stove per day)	**0.46**

Example 5: Single sample. Sample size is within required precision	SUBJECT I.D.		BASELINE FUEL	PROJECT FUEL
	HH1		Use other data	0.94
	HH2		sources, e.g.	1.05
	HH3		default efficiency	1.14
	HH4		and	0.94
	HH5		data on delivered	0.53
	HH6		energy	0.82
	HH7			0.94
	HH8			0.64
	HH9			1.55
	HH10			1.14
	HH11			1.41
	HH12			1.05
	HH13			1.59
	HH14			0.78
	HH15			0.81
	HH16			0.69
	HH17			0.63
	HH18			1.01
	HH19			0.89
	HH20			0.92
	HH21			1.09
	HH22			1.08
	HH23			0.44
	HH24			0.84
	HH25			1.05
	HH26			1.14
	HH27			0.94
	HH28			0.43
	HH29			0.72
	HH30			0.94
	HH31			0.64
	HH32			1.55
	HH33			1.14
	HH34			1.41
	HH35			1.05
	HH36			1.59
	HH37			1.08
	HH38			0.81
	HH39			0.74
	HH40			0.67
	Mean fuel use (Kg per stove per day)			**0.97**
	Standard Deviation (Kg per stove per day)			0.30
			COV	0.31
	What is the precison attained?			8%
	Does the sample size satisfy the 90/10 rule?			**YES**
Note: All the calculations here assume Simple Random Sampling has been used	**What fuel saving may be claimed?**			**Use mean**
	Implied sample size for 90/10 precision?			27
	How many additional samples are required?			0
	90% Lower Limit (Kg per stove per day)			**0.91**

Example 6: Single sample. Sample size is not within required precision	SUBJECT I.D.		BASELINE FUEL	PROJECT FUEL
	HH1		Use other data	0.64
	HH2		sources, e.g.	0.02
	HH3		default efficiency	2.26
	HH4		and	0.93
	HH5		data on delivered	0.62
	HH6		energy	1.81
	HH7			0.77
	HH8			2.02
	HH9			2.54
	HH10			2.45
	HH11			1.32
	HH12			3.01
	HH13			0.92
	HH14			1.35
	HH15			0.98
	HH16			1.65
	HH17			1.89
	HH18			0.43
	HH19			2.52
	HH20			3.70
	HH21			0.53
	HH22			2.53
	HH23			0.95
	HH24			0.68
	HH25			1.24
	HH26			3.66
	HH27			3.43
	HH28			3.18
	HH29			1.71
	HH30			0.42
	HH31			0.89
	HH32			0.94
	HH33			3.13
	HH34			0.34
	HH35			1.80
	HH36			1.95
	HH37			2.33
	HH38			0.24
	HH39			0.08
	HH40			2.97
	Mean fuel use (Kg per stove per day)			**1.62**
	Standard Deviation (Kg per stove per day)			1.06
			COV	0.66
		What is the precison attained?		17%
	Does the sample size satisfy the 90/10 rule?			**NO**
Note: All the calculations here assume Simple Random Sampling has been used		**What fuel saving may be claimed?**		**Increase sampl**
	Implied sample size for 90/10 precision?			120
	How many additional samples are required?			80
	90% Lower Limit (Kg per stove per day)			**1.40**

Annex 1 Technical notes

Calculations

Let $\bar{y}$ denote the estimate of mean fuel savings calculated from the sample, and SE_y be the standard error of this estimate. Then for a sample of size n, the estimated precision is given by the formula

$$precision = 1.67 \times \frac{SE_y}{\bar{y}} \times 100. \qquad (1)$$

Here 1.67 is used as an approximation to the critical value $t_{0.95,n-1}$, which will vary between 1.75 and 1.64 as the sample size n increases from 15 to very large. This approximation is conservative for samples larger than 70, but may slightly underestimate the required sample size for smaller samples.

If the required precision is not met and it is decided to use the lower bound method, then the lower bound is calculated by

$$\bar{y} - 1.3 \times SE_y. \qquad (2)$$

The value of 1.3 is used to approximate the critical value $t_{0.9,n-1}$.

Sample Size

The following formulae show the calculations used to estimate minimum sample sizes in the special case of Simple Random Sampling.

Suppose paired sampling is used. If s_y is the standard deviation of the sample fuel savings, then SE_y is equal to $s_y/\sqrt{n}$, where n is the sample size. Equation (1) can then be rearranged to show that the minimum required sample size to achieve "90/x" precision with a paired sample is

$$n \geq \left(\frac{s_y}{\bar{y}} \times \frac{1.67}{x/100}\right)^2. \qquad (3)$$

Equation (3) can also be used with a single sample (one-sided KT), in which case s_y denotes the standard deviation of the sampled fuel-use values.

If the project and baseline samples are independent, then the standard error of the estimate is

$$SE_y = \sqrt{\frac{s_b^2}{n_b} + \frac{s_p^2}{n_p}}, \qquad (4)$$

where n_i denotes the size and s_i denotes the standard deviation of the ith sample, for i equal to $p$roject or $b$aseline. Assuming that both samples are the same size, $\tilde{n}$ say, substituting (4) into Equation (1) and rearranging shows that the minimum required sample size to achieve "90/x" precision with two independent samples is approximately equal to

$$\tilde{n} \geq \left(\frac{\sqrt{s_b^2 + s_p^2}}{\bar{y}_b - \bar{y}_p} \times \frac{1.67}{x/100} \right)^2 . \qquad (5)$$

The total required sample size in this case is thus $2\tilde{n}$. The sample size calculations in Table 2 above further assume for simplicity that $s_b^2 = s_p^2$, and substitutes the given values into (5). The Excel calculations for additional sample size assume an equal number of samples will be added to both project and baseline samples, and uses Equation (4) substituted into Equation (1) to find the smallest value for which 90/30 precision is met. Note that the assumptions of equal sample size or standard deviation are made only to simplify the presentation of sample size calculations. In general when taking independent samples, if the baseline population standard deviation is greater than the project population standard deviation it will be more efficient to take a larger baseline sample than project sample, and vice versa if the project population standard deviation is greater than the baseline population standard deviation. The Excel spreadsheet also provides a calculator which gives the precision for any values of s_p, s_b, n_p, n_b and $\bar{y}_b - \bar{y}_p$.

The **coefficient of variation** (COV) is defined as the ratio of the population standard deviation to the population mean. It is a measure of how much variability there is in the population relative to the size of the mean that is being estimated. If Simple Random Sampling (SRS) is used then the ratio $s_y/\bar{y}$ provides an estimate of the COV.

For the case of paired sampling or a one-sided KT, if SRS is used the estimated COV relates directly to the sample size required for a given level of precision. The sample size entries in Table 1 were obtained by substituting in the COV values for $s_y/\bar{y}$ in Equation (3). For other sampling designs the relationship between COV and sample size is much more complicated.

If a sampling method other than SRS is used, such as cluster sampling or stratified sampling, the standard error of the estimate for a given sample size is likely to be larger than if using SRS. Thus a larger sample size than that given in the tables will be required to achieve the desired level of precision. Data from a pilot study or previous work may be used to estimate $SE_y/\bar{y}$ directly, and this can be substituted into (1) to obtain the minimum required sample size[13]. Although for an alternative sampling design such as cluster sampling to attain the same level of precision as a SRS a larger sample is generally required, these methods of sampling can be much cheaper to carry out in practice.

Small Sample Sizes

The calculations for precision and for the lower 90% confidence limit assume that the sampling distribution of the estimator is normal. If the sample size is at least 30, then the central limit theorem

[13] Also see Section III, Sampling Application Guidance, of the General Guidelines for Sampling and Surveys for Small-Scale CDM Project Activities (EB 50 Report, Annex 30)

implies that this assumption is likely to be valid. For smaller sample sizes, it is necessary to check this assumption more carefully.

The following refers to the case of SRS and a paired or one-sided KT design; modifications are available for other sampling designs, but if the design is too complicated a larger sample may be required. If a histogram of the sample values looks approximately bell-shaped and symmetrical, then the precision and lower bound can be calculated as described above. If this is not the case, another way to check the suitability of the normality assumption is to use so-called "bootstrap" methods[14] to approximate the sampling distribution directly. To obtain a bootstrap sample, take at least 1000 random samples of size n with replacement from the observed values, and for each sample record the mean. If a histogram of these 1000+ estimates looks approximately normally distributed with mean $\bar{y}$ and standard deviation $s/\sqrt{n}$, then the methods described above for calculating precision and the lower confidence limit can be used. If not, it may be necessary to increase the size of the sample.

[14] See Efron, B. and Tibshirani, R. (1993) *An Introduction to the Bootstrap*. New York: Chapman & Hall.

Annex 2: Data recording

A basic template for recording data during a KPT, for use on excel sheets, is suggested here. In practice it is recommended to include within an excel sheet cells which record more information, for example the kind of data collected during Baseline Surveys, and data which helps verify that that the notes above with regard to KPT scope and procedure, can be addressed in a test report and verified. A project developer should therefore use this example as a basis to develop a excel sheet which reflects the notes made above, addressing points such as type of food, number of hobs, date and season, location, working day or holiday, and so on, also adding data important to a report, such as name of test supervisor, weighing equipment used and its calibration status, and so on.

		BEFORE PHASE										
Name of subject	Details of subject	DAY1			DAY2			DAY3			ALL DAYs	
		KG of fuel in store at START DAY1	KG Fuel added to store	Number of person-meals DAY1	KG of fuel in store at DAY2	KG Fuel added to store	Number of person-meals DAY2	KG of fuel in store at END DAY3	KG Fuel added to store	Number of person-meals DAY3	KG of fuel used	Number of person-meals
HH1											Calculate	
HH2												
HH3												
etc												

AFTER PHASE									RESULTS	
DAY1			DAY2			DAY3			ALL DAYs	
KG of fuel in store at START DAY1	KG Fuel added to store	Number of person-mealsDAY 1	KG of fuel in store at DAY2	KG Fuel added to store	Number of person-mealsDAY 2	KG of fuel in store at END DAY3	KG Fuel added to store	Number of person-meals cooked DAY3	KG of fuel used	Number of person-meals

Annex 3 History of this protocol

This protocol was prepared in the context of a carbon-reduction-measurement methodologies prepared for various cook-stove projects supported by ClimateCare since 2004. The first methodology was developed in 2006 and approved by the standards body VCS for use in Cambodia in 2007, and then in 2007 ClimateCare developed the voluntary carbon market methodology "Improved Cook-stoves and Kitchen regimes" for approval by the Gold Standard Foundation. In 2007 and 2008 the protocol was developed with assistance from Dr Tim Heaton of the Oxford University Statistics department, and applied to measurements of improved cook-stove performance in Ghanaian homes. Supervision reports from Tim Heaton were submitted in May 2008 as part of a carbon accreditation process. In 2008-9 Dr Amber Tomas provided advice for further development of the protocol in particular incorporating the sampling guidelines from the CDM as approved by EB50 in October 2008, and ClimateCare applied the protocol to various projects, including dissemination of heat retention energy-saving devices in South Africa. In 2009 ClimateCare recommended revisions to the CDM small-scale cook-stove methodology AMSIIG applying the principles of the protocol (such as tiered analysis combining precision and lower bound approaches). The precision levels of 90/10 and 90/30 in this document were decided in early 2011 in discussion with the Gold Standard Foundation and the various contributors to the new GS VER methodology 'Technologies and Practices to Displace Decentralized Thermal Energy Consumption'.