Energy for Sustainable Development xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Energy for Sustainable Development

Testing institutional biomass cookstoves in rural Kenyan schools for the Millennium Villages Project

Edwin Adkins [a,*], Jiehua Chen [a], Jacob Winiecki [a], Peter Koinei [b], Vijay Modi [a]

[a] Earth Institute and the School of Engineering and Applied Science, Columbia University, New York 10027, USA
[b] UNDP-Millennium Villages Project, Kenya

ARTICLE INFO

Article history:
Received 12 March 2010
Revised 7 July 2010
Accepted 7 July 2010
Available online xxxx

Keywords:
Biomass
Cookstove
Institutional stove
Paired test
Unpaired test
Carbon emissions

ABSTRACT

Cooking tests were conducted in randomly selected school kitchens in the Sauri Millennium Villages Project site, located in Siaya District of Nyanza Province in Western Kenya. The tests compared fuel consumption measurements obtained using a traditional three-stone fire with those from newly introduced institutional stoves based on the "rocket" design. The key metric used was Specific Fuel Consumption (SFC), defined as the weight of firewood consumed in cooking a single batch of food divided by the total weight of food, measured after cooking. Tests followed the normal cooking practices in the school kitchens and included the typical range of foods prepared for midday school meals programs. The study included two types of tests: paired tests, in which most conditions were controlled between one test conducted on a three-stone fire and a matching test conducted on a "rocket" stove; and unpaired tests, in which conditions were similar, but not strictly controlled, among two large sets of relatively independent three-stone fire and rocket stove tests. Results from both paired and unpaired experiments, averaged across all types of food cooked, showed that the use of rocket stoves yielded significantly lower SFC values without prolonging cooking time when compared with three-stone fires. An analysis comparing results from paired and unpaired cooking tests suggests that, due to high variance and sources of bias in unpaired tests, experimental design should favor paired tests.

1. Introduction

An estimated 2.4 billion people globally rely on solid fuels such as coal and biomass for their energy needs, burning 2 million ton of biomass each day (Rehfuess, 2006; Rehfuess et al., 2006). This has consequences for both the environment and human health. According to the WHO, energy poverty, marked by lack of energy and lack of access to modern cooking fuels, creates obstacles to achieving the Millennium Development Goals (MDGs), the global targets for reducing extreme poverty and improving health and welfare. Indoor air pollution generated by burning solid fuels, among other adverse impacts, is estimated to be responsible for 2.7% of the global disease burden and nearly 1.6 million deaths per year, mainly pneumonia among children and chronic respiratory disease among adults (WHO, 2007).

The Millennium Village Projects (MVP), which works in 14 rural locations in 10 countries across sub-Saharan Africa, focuses on achieving the MDGs through a multi-sector program with community-based investments in agriculture, health, education, water and sanitation, infrastructure and environment (Sanchez et al., 2007). One component is a midday meal program for primary school children. Because cooking at schools is primarily accomplished using fuelwood, stoves that use less fuelwood offer the potential to reduce stress on the environment, reduce the time-labor burden of fuelwood collection, and improve air quality in indoor school kitchens.

This study compares the fuelwood use of the traditional three-stone fire and improved stoves based on the "rocket" design in field settings. Normalized fuelwood use is reported as Specific Fuel Consumption (SFC), defined for this study as the weight of firewood consumed during cooking divided by the total weight of cooked food, measured after cooking. SFC is dimensionless with data for both firewood and food measured in kilograms (kg fuelwood/kg cooked food). A smaller SFC value indicates lower fuelwood consumption. Cooking time is also measured and compared between the two stove types.

The three-stone fire, also called the "open fire", is the traditional and primary stove type used throughout sub-Saharan Africa, including the MVP project areas. It consists of three similarly sized stones on which a cooking pot is balanced over a fire. While this is the easiest way to create a simple support structure for the pot while allowing air for combustion to be drawn in, this type of fire can result in a reduction in the amount of radiant heat delivered to the pot. The heating of excess air reduces temperatures, the heated air flow has limited means to deliver its heat to the pot, and there is no control of direct heat losses from the fire. While fuelwood needed for cooking a

* Corresponding author.
E-mail address: jea98@columbia.edu (E. Adkins).

0973-0826/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.esd.2010.07.002

given amount of food in a three-stone fire can vary widely depending upon the cook, it is generally difficult to get more than 20% of the calorific value of the fuelwood used into the pot as useful heat. Moreover poor control of the air–fuel ratio and mixing can result in incomplete combustion, leading to high emissions of particulate matter as well as a large and diverse set of undesirable chemical components.

The improved biomass cookstoves introduced in the study are modeled on the "rocket" stove, a key feature of which is a vertically elongated combustion chamber which controls airflow, combustion and mixing more than an open three-stone fire. Moreover the walls of the combustion chamber are made of an insulating material with low thermal mass, and the stove design directs the heated air and gases closer to the pot, increasing heat transfer (Winiarski, 2005). Rocket stoves can range in size, materials, and portability, from small household models to large brick-lined institutional stoves. The stoves tested in this experiment are institutional models manufactured by a local contractor and selected villagers with technical support from the German Technical Cooperation (GTZ). The detailed design of the stove type tested is publicly available (Uganda Ministry of Energy and Mineral Development, 2004). A photo is provided in Fig. 1b, below, and detailed technical diagrams are provided in Fig. A1.

Jetter and Kariher (2009) compared three-stone fires with improved stove types, measuring fuelwood consumption and pollutant emissions. The results of efficiency tests using the Water Boiling Test (WBT) protocol in different laboratories consistently show that most improved stoves have better fuel efficiency than three-stone fires. However, other investigations (Bailis et al., 2007; Berrueta et al., 2008) show that laboratory results from the WBT protocol gave little indication of how a stove would perform under typical cooking conditions in field settings. Field-testing is therefore critical in estimating the achievable fuel saving of improved stoves. In light of this, our experiments were conducted in randomly selected schools in the Sauri, Kenya, MVP site, in Siaya District, Nyanza Province, and followed field experiment settings in order to provide an estimate of fuel savings in the field.

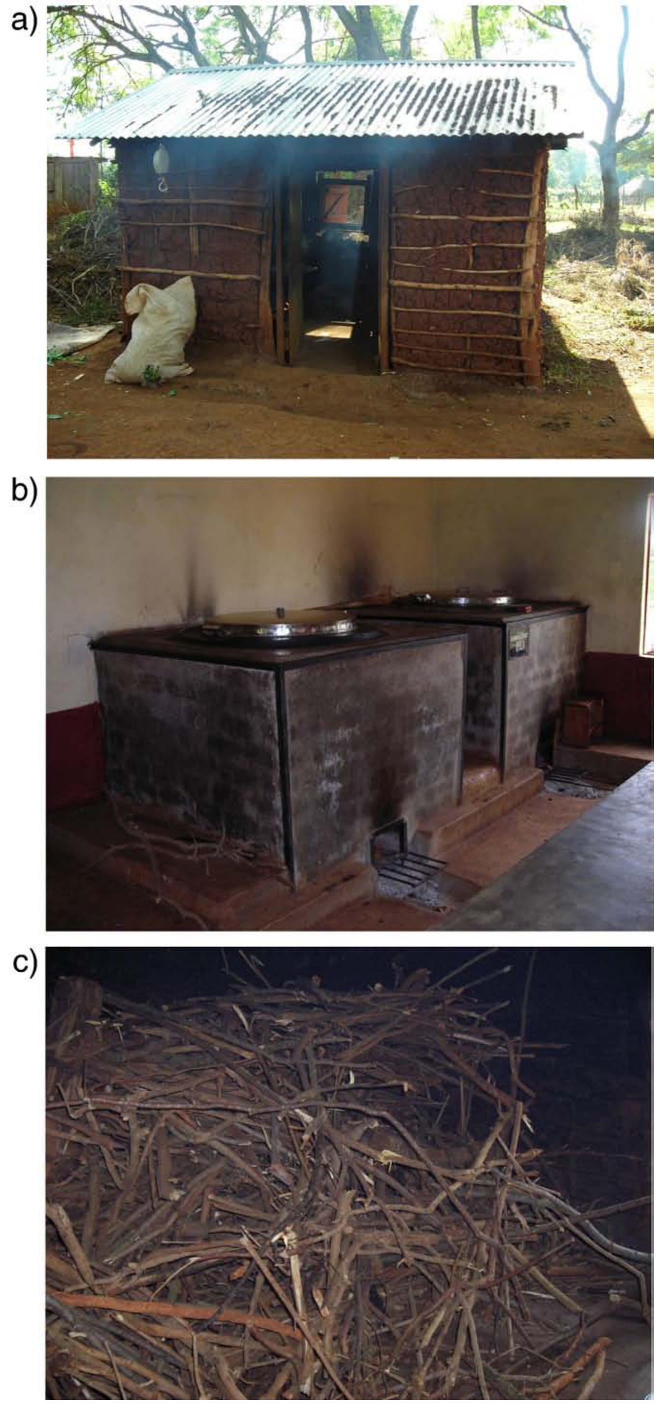

Fig. 1. Three images from Sauri village, Kenya: a) mud kitchen at Bar Tuoro Primary; b) Institutional Rocket Stove in a masonry kitchen, Nyamninia Primary; c) fuelwood in dry storage.

2. Experiments

Two kinds of cooking tests—"paired" and "unpaired"—were implemented. Paired tests were conducted using a modified version of the Controlled Cooking Test (CCT) protocol developed by the University of California-Berkeley and Shell Foundation Household Energy and Health Projects (Bailis, 2004). As in a normal CCT, the paired tests conducted in this study measured fuel use and cooking time for rocket stoves and three-stone fires in the same or very similar kitchen settings, with the same cooks, foods and food quantities, and with the same batch of fuelwood. The key modifications made in the CCT protocol for this study were the following: a) tests were conducted in rural school kitchens rather than in a laboratory; b) tests of different stove types were conducted on successive days rather than at the same time; and c) fuel was provided by the schools themselves. In contrast, for unpaired tests, the two types of stoves were tested at separate sites, with gaps in time larger than a single day, with different batches of wood, cooking varying quantities of a range of foods.

There are advantages and drawbacks to both types of tests. Paired tests require careful planning, preparation and control to ensure that all experimental factors are consistent for both testing days. While this places greater demands on school cooks and meal programs, the resulting dataset is more amenable to analysis of the impact of change in stove type. In contrast, unpaired tests are essentially conducted as normal school cooking events, requiring no additional efforts other than careful food and fuel measurements. Thus, unpaired tests require less alteration of daily cooking practices, and are easier to implement quickly and in larger numbers. In the following sections, the data from both types of experiments are analyzed separately then compared in an effort to determine which testing approach obtains the most helpful data with the least burden upon cooks and school meals programs.

2.1. Paired experiment

2.1.1. Experiment set-up

Thirty paired institutional stove tests (15 pairs) were conducted in two primary schools—Bar Tuoro Primary School and Nyamninia Primary School—during March and April 2007. Each paired test

Please cite this article as: Adkins E, et al, Testing institutional biomass cookstoves in rural Kenyan schools for the Millennium Villages Project, Energy for Sustainable Development (2010), doi:10.1016/j.esd.2010.07.002

includes measurement of fuelwood consumed and cooking time for two equal quantities of the same food, one cooked on a rocket stove and one on a three-stone fire. The experiments included four different kinds of food: nyoyo (primarily maize and beans, with small amounts of tomatoes, onions, salt, and cooking fat), ugali (boiled maize meal), beef (meat plus tomatoes, onions, and cooking fat), and kale (vegetables plus tomatoes, onions, salt, and cooking fat). Eight paired tests were conducted with nyoyo, eight with ugali, four with beef, and ten with kale. For each food, half of the tests were tested on the rocket stoves, and the other half was tested on the three-stone fires. Fig. 1 includes photos of kitchen types and fuelwood storage.

In Bar Tuoro Primary School, six paired tests were conducted during the regular school session (March 2007) with near full attendance by primary school and nursery students, plus faculty, giving a total of 300 people for each paired test. Testing data from this school was obtained in two kitchens: Three-stone fire tests took place in a small mud hut (Fig. 1a). Rocket stove tests were conducted in a larger masonry kitchen using two rocket stoves with pots 200 and 300 l in size (Fig. 1b). Each rocket stove had been in use for approximately 1 month at the time of the tests.

In Nyamninia Primary School, 24 paired tests were conducted over a 2-week period in April 2007 when school was not in session. Meals were prepared for about 150 students attending school during the holiday break using two rocket stoves of the 200- and 300-l sizes, which had been in use for approximately 4 months at the time of the tests. All testing for both the three-stone and rocket stoves at this school was conducted in a newly built, masonry kitchen.

In the paired experiments, all factors other than the stoves were controlled for as follows:

- Cooking practices: The same two cooks prepared each meal using their normal cooking methods while allowing for key measurements of fuel and food. The fire was stopped but not extinguished, then revived several times per day to permit removal and weighing of firewood and food, as well as cleaning the 200- and 300-l pots to allow multiple batches to be cooked in the same pot each day. Testing activities added to the total daily cooking time and may have influenced combustion slightly by allowing the stove to cool between food batches.
- Food: Equal portions of the each type of food were cooked on each stove for each test.
- Fuel: Preliminary measurements showed that the firewood collected tended to have high initial moisture content (30–60% using a basic, two-pin moisture meter). To reduce moisture levels, the larger pieces of wood were split, dried in the sun, stored in the kitchen overnight (Fig. 1c) and evenly divided among stoves for tests. Because each set of paired tests used the same batch of wood, subsequent moisture content measurements were not taken on test days.
- Measurement: Quantity of food and fuelwood were measured before and after cooking using a hanging, spring scale, accurate to 0.5 kg. A batch of firewood was set aside and weighed before cooking for each batch of food. The remaining wood was weighed after cooking and the amount consumed was computed by difference. The weight of coals remaining in the stove after cooking each batch of food was estimated by visual inspection. This tended to be less than 5% of the total used to cook a typical meal (less than 2 kg), and the variation in this estimate is seen as a random source of error.

2.1.2. Statistical analysis

Fig. 2 shows a scatter plot of results for the 30 tests. One point corresponds to two tests, with the SFC value for the three-stone fire charted on the *x*-axis and that for the rocket stoves on the *y*-axis for the same pair. The diagonal 45° line passing through the origin represents the set of points where the SFC values for the two stoves are equal. Points above this line indicate tests in which the SFC of the three-stone fire is less than that of the rocket stove (meaning the three-stone fire consumes less fuel per unit of food cooked), while points below the diagonal line—where most of this data resides—correspond to tests in which the fuel consumption of rocket stoves was lower. Overall, Fig. 2 shows that the rocket stove consistently consumes less fuelwood per kilogram of food cooked for all foods, with the exception of ugali, for which the results are somewhat mixed.

A paired, two-sided *t*-test supports the statistical significance of these conclusions. The mean of the differences of the SFC for the rocket stoves and the three-stone fires in each paired experiment is −0.18, and the standard error of the differences is 0.039. The 95% confidence interval of the differences is (−0.10, −0.26), and since

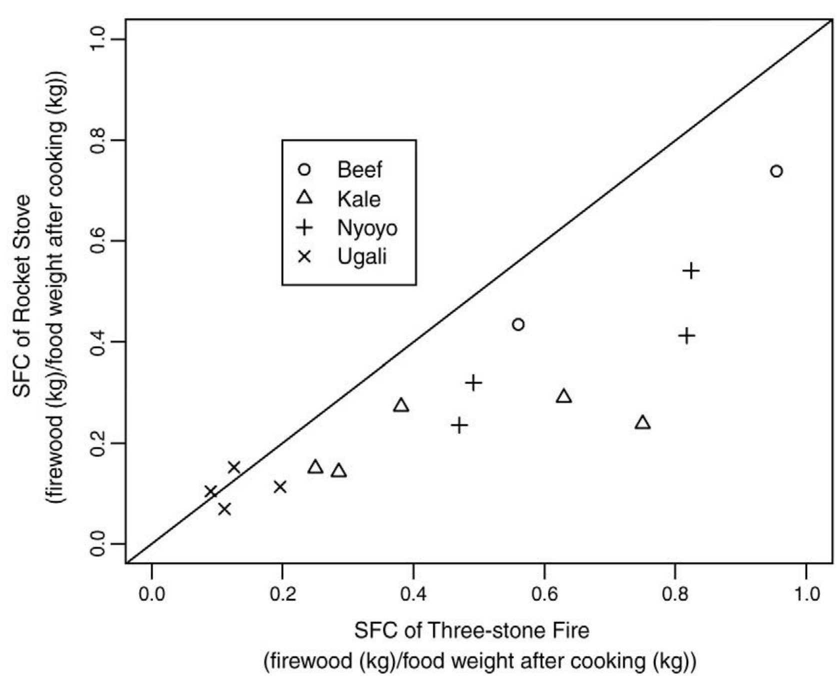

Fig. 2. Scatter plot of SFC values for 30 paired tests (15 pairs).

Table 1
Mean SFC values and comparisons for different food types and stoves in the paired cooking tests (total n = 30, 15 pairs).

	Mean SFC[a]: Three-stone	Mean SFC[a]: Rocket	Mean SFC reduction[b]	SFC % reduction[b]
Nyoyo ($n = 8$)	0.65	0.38	−0.27*	−42%*
Ugali ($n = 8$)	0.13	0.11	−0.02	−15%
Beef ($n = 4$)	0.76	0.59	−0.17*	−22%*,c
Kale ($n = 10$)	0.46	0.22	−0.24*	−48%*
All foods ($n = 30$)	0.46	0.28	−0.18*	−0.33%*

[a] SFC values are dimensionless (kg fuelwood/kg cooked food).
[b] Negative values imply rocket stoves yield fuel savings.
[c] The unusually large *t*-statistic is due to the small *n* and low variance.
* Significant at 5% level.

zero is not included in this interval, the difference of the SFC values between the two stoves is statistically significant.

We also calculated the percentage fuelwood saving as for each pair test as:

$$\frac{\text{SFC of the rocket stoves} - \text{SFC of the three} - \text{stone stoves}}{\text{SFC of the three} - \text{stone stoves}} \qquad (1)$$

The mean fraction of fuelwood savings is −0.33, with a negative value indicating a 33% fuelwood savings for the rocket stoves relative to the three-stone fire, and the standard error is 0.063. The 95% confidence interval of the ratio of the differences is (−0.45, −0.21), which confirms that the results are statistically significant.

Table 1 lists the estimated fuel savings for different food types using the rocket stoves, among which all *t*-statistics are significant except ugali. While cooking kale achieves the highest percent fuelwood savings (at 48%), the savings seen in nyoyo cooking tests (at 42%) are comparable in magnitude and much more important for quantification of overall fuel savings in schools. Nyoyo requires a much longer cooking time because it has a high content of beans, and it is cooked in large quantities, thus it consumes much more fuel per batch. The inconclusive results for ugali, visible in the bottom left of Fig. 2, are quantitatively borne out in this table, and are important in practical terms since ugali is, along with nyoyo, one of the two most frequently cooked meals in rural Kenyan schools. These inconclusive results may be due to two factors: First, the SFC values are smaller in magnitude for cooking ugali on both stove types, making differences less easily quantifiable in a small set of measurements with high variance such as this. Second, ugali requires more frequent and vigorous stirring than other foods, which may move the pot or otherwise affect the air flow patterns within the rocket stoves, possibly diminishing a key contributor to the added efficiency of the rocket stoves. The impact of these factors could be reduced or examined with a larger dataset.

We used a paired, two-sided *t*-test to test if the cooking times for the two kinds of stoves are different. The results are presented in Table 2. None of the Mean Cooking Time Differences are statistically significant (so none have asterisks in this table).

Table 2
Estimated cooking time and comparisons for different food types and stoves in the paired cooking tests (total n = 26).[a]

	Mean cooking time (in minutes): three-stone	Mean cooking time (in minutes): rocket	Mean cooking time difference
Nyoyo ($n = 8$)	262	247.75	14.25
Ugali ($n = 8$)	56.5	80.25	−23.75
Beef ($n = 2$)	142	142	–[b]
Kale ($n = 8$)	45.5	40.75	4.75
All foods ($n = 26$)	122.92	124.38	−1.46

[a] Four data points are missing: two for cooking beef, and two for cooking kale.
[b] Only one pair cooking time is available.

Table 3
Number of unpaired tests for each food and stove type.

	Three-stone	Rocket
Nyoyo	41 (54%)	12 (41%)
Ugali	2 (2%)	2 (7%)
Vegetables	12 (16%)	8 (28%)
Tea	9 (12%)	1 (3%)
Porridge	12 (16%)	6 (21%)
Total	76 (100%)	29 (100%)

2.2. Unpaired experiment

2.2.1. Experiment set-up

Unpaired experiments were also conducted, in which fuel use and cooking time measurements were taken for cooking events in which three-stone fires and rocket stoves were used independently, with variation in factors such as the kitchen location, quantities of food, wood supply, and cooks. The three-stone fire measurements were performed in 12 primary schools in the village study area, mostly outdoors, in early 2008, while the institutional rocket stove tests were all conducted in the later part of 2008 at 10 village schools inside the newly constructed school kitchens, all of the same masonry design, including large windows and bricks which permit ample ventilation. Although the tests were unpaired, the manner of firewood provision, collection of local firewood by schoolchildren, was broadly consistent between the tests that used both stove types. While moisture content of the wood is a key factor which certainly varies with many changes in local conditions (seasonal and daily temperature, rainfall, exposure of collected wood to sunlight, splitting and storage techniques, etc.) these tests occurred over multiple sites, over duration of months that spanned changing seasons and hundreds of different collection efforts by schoolchildren. Thus, although there are many factors that cause moisture content to vary widely, the wide range of weather and geophysical conditions of fuelwood gathering and storage ensured that no single factor was likely to introduce systematic bias in fuelwood moisture content. In addition, field tests using moisture meters of recently collected wood to be used within a day or two of collection—which was the case for the unpaired tests in this study—have shown that moisture content measurements show large variance even at different points along a single stick of wood. As a result, it is reasonable to assume that the moisture content of collected wood is essentially random. Other cooking settings are similar to the paired experiments introduced in the previous section.

2.2.2. Statistical analysis

Originally, there were 79 unpaired test observations for three-stone fires and 30 observations for rocket stoves; however, some data were excluded from analysis: The sole rocket-stove test for cooking "meat" resulted in an SFC value of 4, which far exceeds other

Table 4
Mean SFC values (with t-statistics) for different food types in the unpaired experiment.

	Mean SFC: three-stone	Mean SFC: rocket	*t*-statistic	*t*-statistic for percent difference
Nyoyo	0.53	0.20	−8.68*	−16.53*
Ugali	0.59	0.15	−2.31*	−4.73*
Vegetables	0.20	0.15	−3.57*	−1.23*
Tea	0.48	0.24	NA[a]	NA[a]
Porridge	0.35	0.15	−5.46*	−10.28*
All foods	0.45	0.17	−8.71*	16.77*

* Significant at 5% level.
[a] The standard error cannot be calculated because there is only one observation.

Please cite this article as: Adkins E, et al, Testing institutional biomass cookstoves in rural Kenyan schools for the Millennium Villages Project, Energy for Sustainable Development (2010), doi:10.1016/j.esd.2010.07.002

Table 5
Estimated cooking time and comparisons for different food types and stoves in the paired cooking tests.

	Mean SFC: three-stone	Mean SFC: rocket	*t*-statistic
Nyoyo	276.98	290.92	−1.15
Ugali	126.5	160.5	−1.14
Vegetables	68.08	99.25	−1.44
Tea	52.11	16	NA[a]
Porridge	76.08	86.5	−0.69
All foods	181.68	177.25	−0.18

[a] The standard error cannot be calculated because there is only one observation.

observations and so has been designated as an outlier. Three tests for "tomato soup" were only cooked with the three-stone fires. These four exclusions reduced the dataset to 105 observations: 76 for three-stone fires and 29 for rocket stoves. The resulting dataset includes measurements by stove type for cooking five kinds of food (all as described above, unless otherwise noted): nyoyo, vegetables (as "kale," above), tea, porridge (primarily maize flour), and ugali. Table 3 lists the number of test observations for each food type followed in parentheses by the percentage of total tests for each stove type that these tests represent.

For these experiments, an unpaired, two-sided *t*-test is required. The difference between the mean SFC for rocket stoves and the mean for three-stone fires for all tests is −0.27, with a standard error of 0.031 for this mean difference. The 95% confidence interval of the differences is (−0.33, −0.20), indicating that the SFC reduction is statistically significant. Table 4 contains the *t*-statistics for the SFC differences for different food types, and all of the unpaired *t*-statistics are significant at the 5% significant level.

For the unpaired experiment, we cannot calculate the percentage fuelwood savings directly using the ratio of differences as defined by Formula (1). Instead, an approximate result can be obtained using the Delta method (see Casella and Berger, 2002 for details) which approximates the distribution of the mean percentage fuelwood savings as its asymptotic distribution, the distribution of a random variable when the number of data points approaches infinity. By the Delta method, the mean of the percentage fuelwood savings is −0.62 (or 62%), approximately double the value obtained from the paired tests, with a standard error equal to 0.036. Again the SFC reduction is statistically significant.

Also by an unpaired, two-sided *t*-test, the difference between the mean cooking times for the two kinds of stoves is 4.4 min, and the standard error is 23.5 min. The 95% confidence interval is (−41.7, 50.5) min, and the difference is not statistically significant, as seen in the results from the paired experiment (Table 2).

As was done previously for paired (Table 2), a two-sided *t*-test was conducted using the unpaired data to test if the cooking times for the two kinds of stoves are different. As before, none of the Mean Cooking Time Differences are statistically significant (see Table 5).

Table 6
Estimated mean SFC differences (with *t*-statistics) between stove types for paired and unpaired tests for nyoyo, ugali and all foods combined.

	Unpaired experiment		Paired experiment	
	Mean	*t*-statistic	Mean	*t*-statistic
Nyoyo	−0.33	−8.68* (53)	−0.27	−5.55* (8)
Ugali	−0.43	−2.31* (4)	−0.02	−0.83 (8)
All foods	−0.27	−8.71* (105)	−0.18	−4.61* (30)

Numbers in the parentheses are the total number of tests implemented for both types of stoves.
* Designates that the *t*-statistic is significant at the 5% significance level.

3. Discussion

In this section, we will compare the results from the paired and unpaired experiments.

3.1. Statistical power comparison

Fig. 3 summarizes SFC differences between three-stone fire and rocket stoves for both paired and unpaired experiments, with dots representing estimated means, and lines extending outward representing 95% confidence intervals. The paired experiment data produced confidence intervals of equal or narrower width than the unpaired data with less than one third as many data points. Therefore the variance of the SFC data obtained in the paired experiment is smaller than that of the unpaired experiment, which implies that in order to have the same *t*-statistic, the unpaired experiment requires fewer tests than the unpaired experiment.

It is important to address the possibility that differing overall SFC values and variances may be due to differences in the foods tested for paired and unpaired tests, as shown in Tables 1 and 4. To control for this, we compare SFC data for only nyoyo and ugali, the two foods common to both paired and unpaired datasets. These results are listed in Table 6.

Again, results for both nyoyo and ugali show that the smaller variance of the paired test results have smaller variance than the unpaired experiment.

3.2. Matching

As shown in Table 6, averaged across all foods, the unpaired experiment shows larger estimated SFC reduction than the paired experiment estimates (−0.27 vs. −0.18). This section considers possible reasons.

Unlike the paired tests in which cooking conditions are strictly controlled, in the unpaired tests the quantity of food cooked and pot sizes were allowed to vary according to factors such as daily school attendance and the convenience or availability of cooking equipment. Before the installation of the institutional rocket stoves, most schools cooked with multiple three-stone fires using pots predominantly

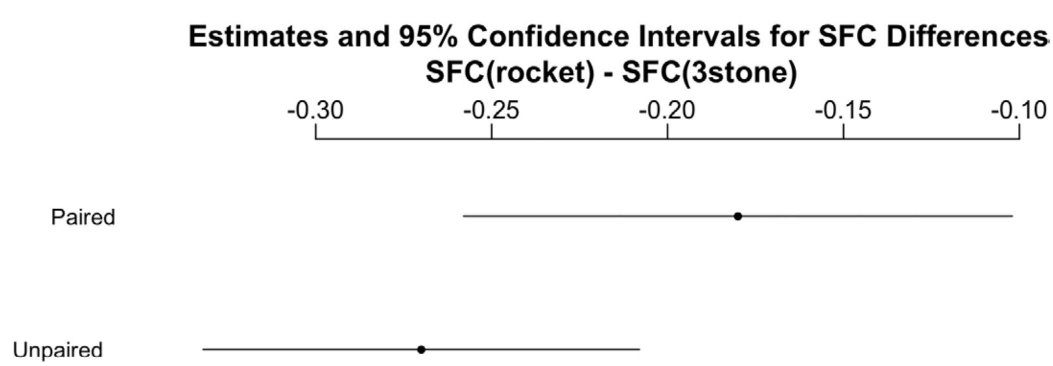

Fig. 3. Estimated means and 95% Confidence Intervals for SFC values from paired and unpaired tests.

Please cite this article as: Adkins E, et al, Testing institutional biomass cookstoves in rural Kenyan schools for the Millennium Villages Project, Energy for Sustainable Development (2010), doi:10.1016/j.esd.2010.07.002

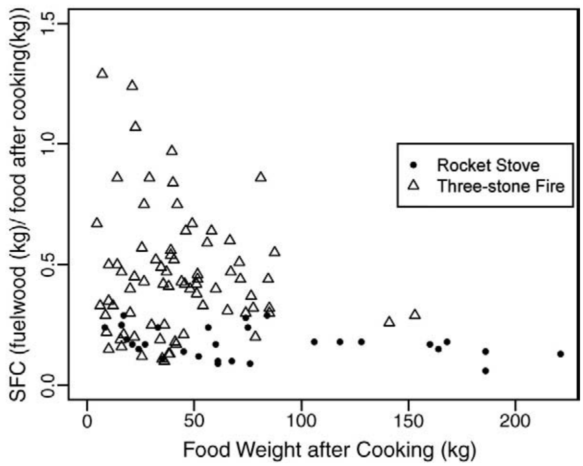

Fig. 4. Specific Fuel Consumption vs. Weight of food after cooking (Unpaired Tests, All Foods).

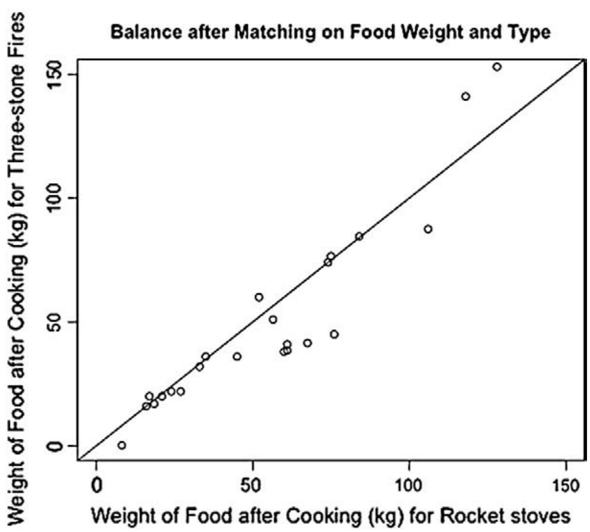

Fig. 5. Unpaired data after matching by both food weight and type.

sized between 20 and 70 l. However, the institutional rocket stoves were installed with pots generally of 100 l or more, allowing cooks to prepare fewer batches of food, using larger pots and greater food quantities per stove test. As shown in Fig. 4, which plots the food weight versus SFC for the unpaired experiment, 8 out of 29 rocket stove tests (28%) cooked food in quantities greater than 100 kg; however for unpaired three-stone fire tests this number was only 2 out of 76 (3%). Moreover, the regression between weight of food cooked and SFC show significantly negative association. Therefore the tendency toward larger batches in the rocket stove tests creates a potential systematic bias in the estimate of SFC reduction.

Also, as shown in Table 1, the number of tests conducted for each food differs by stove. The following sensitivity analysis considers the impact that both quantity and type of food cooked in each test may have upon the estimates of SFC reduction for the unpaired tests. There are of course other uncontrolled factors which may also introduce bias in the unpaired tests, including weather conditions and individual cooks' practices. However, as measurements for these were not taken in this experiment, they are not considered in detail here.

The unpaired data can be analyzed in a manner that mimics the paired tests by employing the nonbipartite matching algorithm introduced by Lu et al. (2001). This method matches unpaired measurements from rocket stove and three-stone fire tests in which similar amounts of the same food are cooked, then treats these matched measurements as paired tests. The algorithm balances these pairs of measurements by selecting values such that the sum of squares of differences between weights of food cooked in each pair is minimized. This method created 23 matched pairs of rocket stove and three-stone fire measurements. As was shown in Fig. 4, there are 6 rocket stove tests in which more than 150 kg food was cooked, and these cannot be matched with any three-stone fire test. Fig. 5 plots weights of food after cooking for rocket stoves and three-stone fires for each matched pair, most of which fall close to the diagonal line, indicating that the weights are similar.

For statistical analysis, the SFC values from these pairs are then treated as results from a paired experiment. The mean of the SFC differences is −0.11, and the standard error is 0.03. The mean of the ratio of the SFC differences defined in Formula (1) for those 23 matched pairs is −0.28, with a standard error of 0.071, indicating a statistically significant fuelwood savings of 28% using the rocket stoves.

Fig. 6 shows how the estimated SFC differences and fuelwood savings from the matched data reported immediately above differ from both the paired and unpaired (and "unmatched") test results initially reported in Section 2.2.2. The matched unpaired data show smaller SFC differences and fuel savings relative the unmatched data, and are closer to the results of the paired tests.

These comparisons of matched with unmatched data imply that the uncontrolled conditions in the unpaired experiment do bias results. Because there are still other unmeasured factors that could produce bias—such as weather conditions and techniques used by individual cooks—the matching analysis presented here cannot

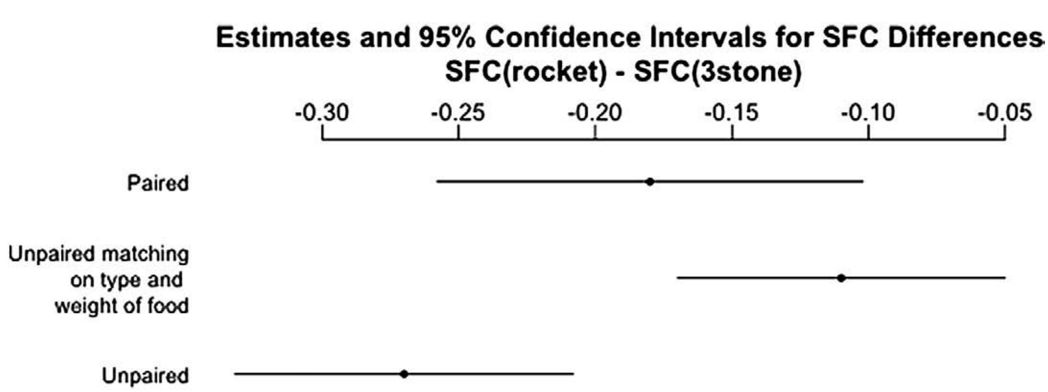

Fig. 6. Mean estimates and 95% Confidence Intervals for all paired, unpaired, and matched datasets.

Please cite this article as: Adkins E, et al, Testing institutional biomass cookstoves in rural Kenyan schools for the Millennium Villages Project, Energy for Sustainable Development (2010), doi:10.1016/j.esd.2010.07.002

remedy the inherent biases in the unpaired test dataset, but it can indicate the greater reliability of paired test results.

4. Conclusions

Results from the paired and unpaired experiments, as well as the "matched" unpaired data, all support the claim that the institutional rocket stoves tested in the Sauri village area show significant fuelwood savings relative to the three-stone fire for cooking the most commonly prepared foods in this field setting. The most conservative estimate based only on results from the paired experiment, is that, on average, without prolonging cooking time, for every 1 kg food cooked the rocket stove saves 0.18 kg fuelwood, a 33% reduction in fuelwood consumption compared to the three-stone fire.

The potential for aggregate fuelwood savings for the study area can be estimated based on the calculated fuel saving from these results and the assumption that all 28 primary schools throughout the Millennium Villages Project area perform all cooking using improved institutional rocket stoves in place of the three-stone fire. The MVP

Capacity: 10 liters; Fuel: Firewood	Designed by MEMD/EAP Drawn by: Leonard Mugerwa Date: April, 2007

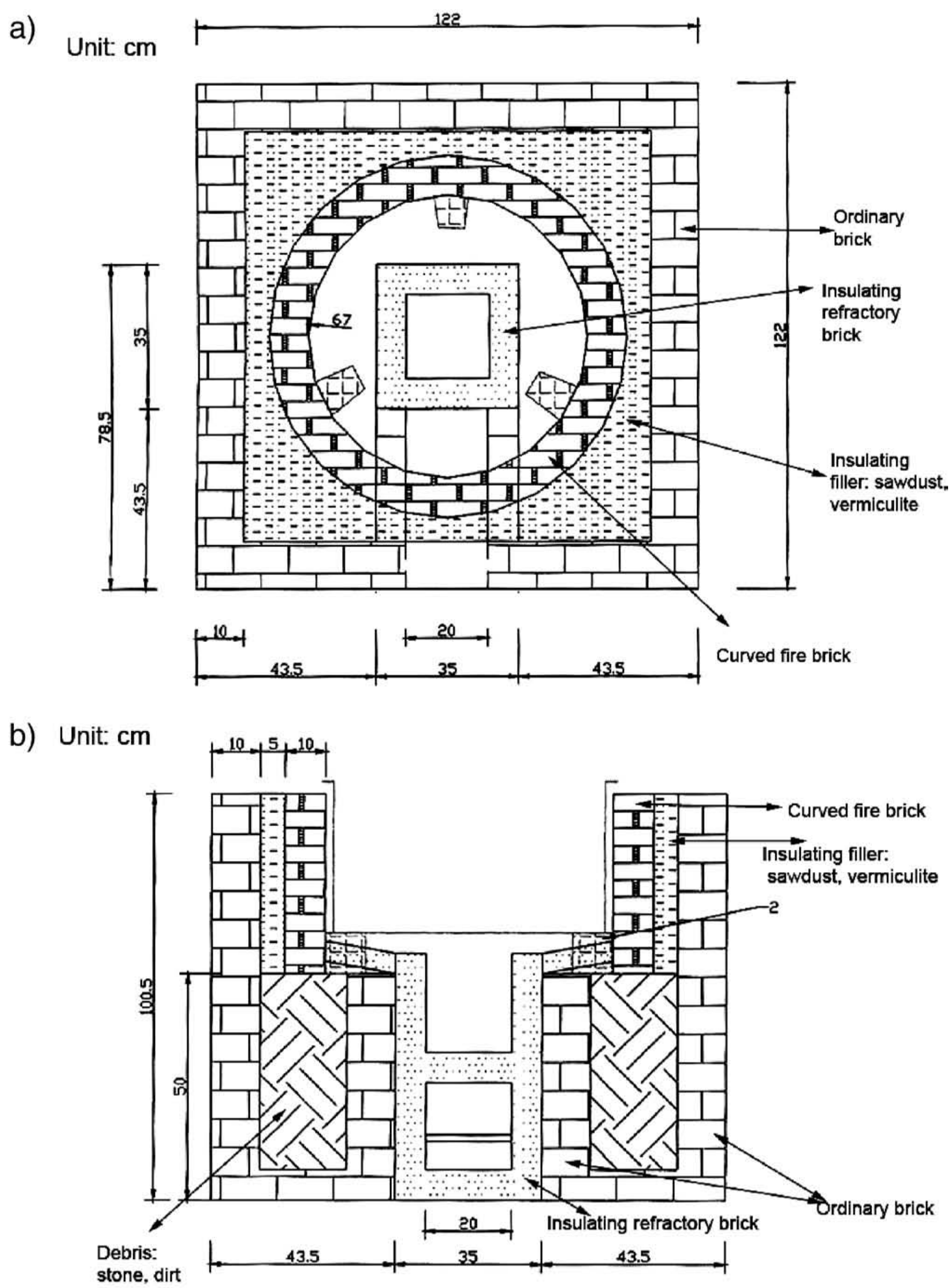

Fig. A1. a. Institutional rocket stove (top view, fuel opening faces downward). b. Institutional rocket stove (front view, fuel opening opening faces forward). c. Institutional rocket stove (side view, fuel opening opening faces to the left).

Please cite this article as: Adkins E, et al, Testing institutional biomass cookstoves in rural Kenyan schools for the Millennium Villages Project, Energy for Sustainable Development (2010), doi:10.1016/j.esd.2010.07.002

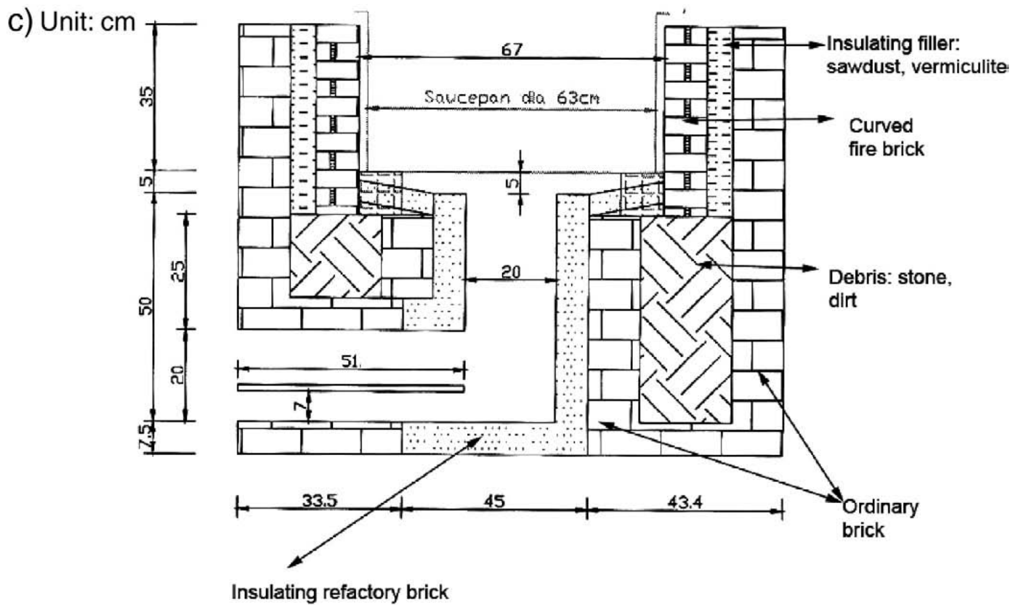

Fig. A1 (*continued*).

school meals program serves 16,800 meals on each of the 195 days of the school year. Each student is served approximately 400 g for each meal, resulting in a total of 1,310,000 kg of cooked food per year. The difference in SFC between the rocket and three stone fire, 0.18 kg fuel/kg cooked food, amounts to a yearly fuelwood savings—assuming 100% usage rate for all stoves—of approximately 236 tons/year. Taking into account occasional use of three-stone fire—such as for cooking of unusual food items, or during periods when stoves are being repaired—suggests a more likely usage rate of perhaps 80%, which reduces total estimated fuelwood savings to around 189 tons.

These savings could form a basis for a plan to support an institutional cookstove subsidy program with carbon credits. Using the conversion factor of 1.747 tons of CO_2 emitted per ton of fuelwood burned (IPCC, 2006), it can be predicted that around 9 tons of CO_2 emissions may be avoided per year for each school that adopts and regularly uses this type of institutional rocket stove for most cooking. Carbon credits trading at US$15/ton would then provide US$135 per school per year. Assuming that fixed, brick rocket stoves of the type used in this study have a usable life of 5–8 years (manufacturers typically claim 10–15 years), this leads to a total income from carbon credits per school of between US$675–1100 over the life of the stoves. Assuming that stoves cost $1,500 each, and two are purchased per school, the total cost for stoves per school would be around US$3,000. A fuelwood consumption monitoring program necessary for obtaining carbon credits would likely cost perhaps US$50 per school per year, bringing the total lifetime stove costs to roughly $3300–3400 per school. By this very rough estimate, carbon credits might be expected to cover roughly 1/4 to 1/3 of overall stove costs, assuming that a program could be implemented at a scale sufficient to lower monitoring costs per stove to the level given above.

In terms of testing methodology, this study suggests that results from paired cooking tests (a modified version of the Controlled Cooking Test) are more reliable than unpaired tests. Since the paired test procedure controls the cooking conditions for both stoves, it leads to smaller variance in measurements and thus requires substantially fewer tests to achieve the necessary statistical power. In addition, it creates a clearer causality argument since many factors that are unknown or difficult to measure can be excluded from consideration. Thus, for both quantitative and qualitative reasons, it is highly recommended that paired experiments are used in stove testing programs implemented in the field.

Acknowledgments

The authors would like to thank all the cooks, administrators, other staff and faculty from the schools who participated in and supported this work. We would also like to express our gratitude to the Sauri, Kenya, Millennium Villages Project staff who assisted in this trial, with special thanks to Patrick Mutuo for committing his team's time and resources to this program. We are grateful for logistical assistance provided by ICRAF, Kenya. Finally, we wish to acknowledge generous funding support from the Bill & Melinda Gates Foundation.

References

Bailis R. Stove performance testing protocols. Controlled Cooking Test (CCT) Version 2.0. Shell Foundation. University of California Berkeley; 2004.

Bailis R, Berrueta V, Chengappa C, Dutta K, Edwards R, Masera O, et al. Performance testing for monitoring improved biomass stove interventions: experiences of the household energy and health project. Energy for Sustainable Development 2007;11(2):57–70.

Berrueta VM, Edwards RD, Masera OR. Energy performance of wood-burning cookstoves in Michoacán, Mexico. Renewable Energy 2008;33(5):859–70.

Casella G, Berger RL. Statistical inference, 2nd ed; 2002.

IPCC. 2006 IPCC guidelines for national greenhouse gas inventories. In: Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K, editors. Prepared by the National Greenhouse Gas Inventories Programme. Japan: IGES; 2006.

Jetter J, Kariher P. Solid-fuel household cook stoves: characterization of performance and emissions. Biomass Bioenergy 2009;33(2):294–305.

Lu B, Zanutto E, Hornik R, Rosenbaum PR. Matching with doses in an observational study of a media campaign against drug abuse. J Am Stat Assoc 2001;96:1245–53.

Rehfuess E, Mehta S, Prüss-Üstun A. Assessing household solid fuel use: multiple implications for the millennium development goals. Environ Health Perspect 2006;3:373–87.

Rehfuess E. Fuel for life: household energy and health. Geneva, Switzerland: World Health Organization; 2006.

Sanchez PA, Palm CA, Sachs JD, Denning GL, Flor R, Harawa J, et al. The African millennium villages. Proc Natl Acad Sci 2007;Vol. 104(No. 43):16775–80.

Uganda Ministry of Energy Mineral Development. How to build the improved institutional stove: a construction manual. Energy Advisory Project with support from the German Technical Cooperation (GTZ); 2004.

Winiarski L. Design principles for wood burning cook stoves. Aprovecho Research Center, Partnership for Clean Indoor Air, Shell Foundation; 2005.

World Health Organization (WHO). Indoor air pollution: national burden of disease estimates. Geneva, Switzerland: WHO Press; 2007.