Monitoring and Evaluation of Improved Cookstove Programs

for

Indoor Air Quality and Stove Performance

The Household Energy and Health Project* of

Appropriate Rural Technology Institute, India
Development Alternatives, India
Grupo Interdisciplinario de Tecnología Rural Apropiada, México
HELPS International, Guatemala
School of Public Health, University of California, Berkeley

With assistance from

Aprovecho Research Center, Cottage Grove, Oregon

Submitted to the

The Breathing Space Program of the Shell Foundation
London, UK

FINAL REPORT
Version 4.1.1
October, 2006

"You don't get what you expect, you get what you inspect."

Suggested citation: Household Energy and Health Project, 2006. ***Monitoring and Evaluation of Improved Cookstove Programs for Indoor Air Quality and Stove Performance***, Indoor Air Pollution Group, School of Public Health, University of California, Berkeley, 41 pp.

* http://ehs.sph.berkeley.edu/hem/page.asp?id=1

Abstract

Standardized monitoring and evaluation techniques for evaluating changes in indoor air quality and stove fuel performance were developed and deployed in two NGO-led programs to disseminate improved cookstoves (ICS) in India and one in Mexico. The results showed major and mostly statistically significant improvements in 48-hour indoor air pollution concentrations in those households using the stoves one year after introduction. Kitchen levels of carbon monoxide reduced from 30-70% and concentrations of small particles reduced 25-65%. Unfortunately, however, the lowered levels were all still well above the Air Quality Guidelines recommended by the World Health Organization for small particles, although approaching them for carbon monoxide.

Results for stove performance were mixed, with some stoves achieving improvement in one or another of the short-term metrics that are part of the Water Boiling Test used to evaluate stoves in laboratory (controlled) settings. The Kitchen Performance Test, which measures fuel use in households under actual use, were less easily determined because of high variation and difficult field logistics, with results that often did not reach statistical significance. Recognizing that in some cases more tests would be needed for confirmation, reduction of fuel use per person ranged from about 20-75%. From the results, it also seems clear that the Water Boiling Test is not a good predictor of actual fuel use and thus should be confined to evaluations during the design stage of stove development.

In two of the sites, the reductions in pollution roughly matched those in fuel use, although in the third, IAQ may have reduced a bit more. This indicates perhaps that for all the monitored stoves much or all of the benefits of each type came from improving the heat transfer into the pots and not from either increased combustion efficiency of the fires or stove venting (reliably working chimneys). More analyses are planned to explore these and other aspects of the stoves.

A range of recommendations are provided for future M&E efforts, with the primary one being to combine efficacy tests (small number of carefully monitored households under normal conditions) combined with larger well-designed surveys (questionnaire only) to determine actual usage and house perception. It is recommended that only those NGOs planning to develop significant long-term capability in making air pollution and stove measurement under field conditions be expected to undertake effectiveness testing, i.e, evaluate population-wide changes from real large-scale dissemination programs. The alternative it to employ professional survey and environmental consulting firms, which also has the advantage of assuring independence of the process. In either case, over the long run it is important to generate national capacities for this kind of work.

Acknowledgements

We much appreciate reviews by Nigel Bruce, Dana Charron, Kristie Jagoe, and David Pennise, and the cooperation and patience of the householders in India and Mexico who participated in the surveys. Funding for this work was provided by the Shell Foundation, London.

Introduction

In the 1980s, several tests of the performance of biomass cookstoves were developed, some of which have occasionally been deployed as part of programs to disseminate improved cookstoves (ICS). Little development or formal evaluation of these methods has occurred since, however. In recent years, there has been growing interest in disseminating ICS to not only improve fuel efficiency but also to reliably improve indoor air quality (IAQ) by use of chimneys, improved combustion, and other techniques. Although measuring IAQ in biomass-using households has been the subject of numerous research studies in the last decades, standard monitoring methods that could be deployed by non-research groups, such as NGOs, interested in evaluating ICS programs have not been developed.

This report documents efforts to help fill these gaps. Funded by the Shell Foundation (SF), four NGOs – Appropriate Rural Technology Institute (ARTI), India; Development Alternatives (DA), India; Grupo Interdisciplinario de Tecnología Rural Apropiada (GIRA), Mexico; HELPS International, Guatemala -- received funding in 2004 to disseminate ICS stoves in rural areas. As part of this effort an evaluation of the effectiveness of the ICS in improving combustion conditions and in reducing IAP was required as a measure of the impact of the programs. To facilitate this evaluation effort SF funded a team at the University of California (UC) to develop and deploy jointly with these groups a novel package of monitoring and evaluation (M&E) methods to systematically document the impact of each NGO's ICS dissemination on IAQ and stove performance (SP) as compared to traditional cooking stoves used locally. Concurrently, a team at Liverpool University developed and deployed M&E methods for health and quality of life assessment in conjunction with these NGOs, the results from which will be reported separately.

This M&E package consisted of four physical elements:

1. Written materials on principles of field sampling, the pollutants, measurement methods, sample size selection, home selection criteria, statistical analysis, and data handling.
2. Written protocols for use of the instruments, sample field monitoring questionnaires, data collection forms and spreadsheet templates for entering and evaluating the data.
3. Equipment kits for IAQ monitoring and SP testing, including computer software as well as instruments and associated supplies
4. Website for access by the NGOs to obtain updates of software, protocols, and other project materials.

Detailed descriptions of the M&E packages for IAQ and SP were provided in previous project reports to the SF and are also found on the HEH website, now maintained by UC Berkeley's Center for Entrepreneurship in International Health and Development (CEIHD, see Appendix I).

In addition to the physical elements, active support of the NGOs was conducted through several training sessions in each country by UC staff over the course of the project, which were evaluated through anonymous questionnaire surveys that were collected and analyzed independently from the personnel performing the training. The training visits generally occurred separately for IAQ monitoring and SP testing as, although personnel that conducted the M&E and the SP were separate, the intention was to focus activities around one methodology so supervisors could be familiar will all aspects of the project and not to burden the NGO too heavily at one time. Detailed descriptions of the training activities were provided in previous project reports to the Shell Foundation. Training visits for the M&E included:

1) An initial training workshop

    Workshop modules included: presentation of monitoring strategy, introduction to biomass combustion principles, introduction to sample size calculations and statistical analyses, home selection criteria, introduction to monitoring equipment, data storage and archiving, monitor positioning criteria, data collections sheets, and an introduction to software used for the monitors. The workshop culminated in demonstration of monitor use in a small number of adjacent homes using open fires, hands on training in selection of monitoring positions, and deployment of equipment including programming, downloading and storage of data.

2) Onsite training of field teams and evaluation of initial deployment in homes

    In addition to the initial training workshop, a second workshop was conducted on site for each project to train the field staff that would be conducting the M&E and would be directly responsible for data collection, storage, etc. The workshop focused on baseline calibration and cleaning of monitors, deployment positions in homes, and data archiving.

3) Data cleaning and database compilation workshop
After completion of data collection, a data cleaning and database compilation workshop was held. Although initially intended as an analysis workshop, screening of both CO and PM data files for data quality control and assurance and compilation of summary statistics into a database were required prior to analysis.
4) Data analysis workshop[1]
The data analysis workshops were conducted to introduce supervisors to statistical analysis of datasets, statistical software, and guide the analysis of results presented here.
5) Stove performance test training
Separate visits were made to each country by the UC team to conduct training for stove performance testing.

This report focuses on an assessment of effectiveness of the M&E package in this application and summarizes the results for IAQ and SP found by the stove programs by use of the package. Detailed results from each program will be published separately.

Because of unforeseen circumstances, the M&E work of HELPS in Guatemala has not yet been completed and thus no results are presented here.

M&E Strategy

To understand and evaluate the M&E methods deployed in this project, it is important to understand the philosophy behind the choices made at the start in designing the M&E strategy. M&E of the impacts of a household intervention, such as an ICS, can occur at several levels:

1. Controlled tests: In laboratory or near laboratory settings with simulated cooking, these are easiest, quickest, and cheapest to conduct, but reveal the technical performance of a stove, not necessarily what it can achieve in real household settings, particularly with devices like stoves that are sensitive to operator behavior.
2. Efficacy trials: These are conducted in real households cooking real meals, but under controlled settings in which every effort possible is made to make sure the stove is used to its best effect. They reveal what is maximally possible, but not necessarily what is actually achieved.
3. Effectiveness trials: These are conducted in the course of an actual dissemination effort with real populations cooking normally and give the best indication of real-world changes. Only these can determine actual usage under realistic promotion schemes. Evaluating such trials is more expensive, difficult, and time-consuming, partly because the nature of real households is high variability and thus sample sizes must be fairly large to obtain statistically significant results.

As the original impetus behind the SF M&E was to understand the impact on populations, it was decided that the focus would be on effectiveness, i.e., M&E as much as possible aimed to

[1] Some training also occurred in California

determine the effects in real populations. For this reason, an effectiveness measure was chosen for each outcome:

- Stove Performance: The Kitchen Performance Test (KPT) was chosen as the primary measure, which involves measurements of fuel use in real households over several days. In addition, given its common use and relative simplicity, improved protocols were also developed and training provided for the Water Boiling Test (WBT), a controlled test.
- Indoor Air Quality: Using a longitudinal "Before and After" design, a field survey of IAQ levels in actual households that were part of the stove disseminations was developed and implemented.

A number of other decisions were also made at the outset that had important implications for conducting the M&E activities. Among the most important:

- The project focused on IAQ and SP M&E activities that would result in quantitative metrics meeting normal statistical conventions for making judgments, for example, having sufficient sample size to be able to say that there was no more than 5% probability that results showing that a measured IAQ improvement were obtained by chance. This required providing training to the NGOs in basic statistical concepts and analysis both for choosing sample sizes as well conducting statistical tests with monitoring data. The reasoning for taking this, admittedly challenging, approach was that quantification of benefits that meet standard statistical criteria would be most effective in convincing donors, local authorities, and government agencies that significant benefits had accrued.
- Although personal pollution exposures are the best indicator of health impact, they are substantially more difficult to measure in household settings than kitchen concentrations and introduce additional variability and ethical considerations. Since the objectives of the NGO were to have some statistical evaluation of the performance of the stove in its most common usage, rather than introducing additional variability inherent in personal exposures due to variability in time spent in different locations, and to avoid asking the groups to undertake even larger and more difficult efforts in their first M&E program, it was decided to focus on stationary IAQ measurements in the kitchen, which is the most sensitive area for measuring changes in IAQ due to stove pollution. Although the percentage reductions in indoor levels may not be equivalent to reductions in personal exposures (due to time spent out of the kitchen), the metric provides a measure of the effectiveness of the ICS in reducing the high levels of smoke in the kitchen from open fires[2]. This is also consistent with many dozens of other studies undertaken around the world, mostly for research purposes, that have used indoor measurements due to the complexity and high participant impact of collecting personal exposure samples.
- Both to reduce variability caused by season and to measure longer term changes, not just those immediately after stove introduction, the main "before and after" measurements were undertaken a year apart. This length of time introduced difficulties, however, such as retention of households in the sample

[2] In addition, personal exposure samples can usually only be deployed for 24 hours as they impose considerable burden on participants, causing higher drop out rates. Indoor samples can sample for longer, which reduces some of the variability associated with daily variation in cooking activities.

- Although there are many pollutants in biomass smoke that probably play a role in the wide range of health effects that have been measured or suspected, it was decided to focus on the two most measured pollutants -- small particles and carbon monoxide. Each is created by somewhat different processes during combustion and has different mechanisms of action in producing health effects. Thus, they probably cover a good part of the spectrum of toxicity of the smoke, but certainly not all. Reductions in both, however, can probably be viewed as a good indication that most other pollutants have been reduced as well.
- To reduce variation between measurements in the same house, a perennial problem with household studies, a 48-h rather than the more commonly used 24-h measurement period was used, which complicates logistics somewhat.
- Because capacity building was a high priority, it was decided that all M&E activities would be undertaken by the NGO staff with equipment and methods that were sustainable, i.e., could be used by them after the end of the project with normal maintenance.

Because there are no inexpensive reliable commercial devices available, however, the choice to measure small particles required either

- Purchase of expensive commercial instruments, which could not be fit within the budget,
- Temporary employment of outside instruments and experts to do the work, which did not meet the last of the priorities listed above, or
- Use of the UCB Particle Monitor (UCB), a novel inexpensive instrument developed for research purposes under a previous SF grant, but showing high promise for use by non-research personnel, although not yet tried out in this role.

The last approach was taken, which, as described below, led to more work than envisaged because of the additional development of the UCB-PM system that was needed, particularly in software and data processing.

Methods

Here we provide only a descriptive summary of the methods used by the programs to measure IAQ and SP. A few more details are provided in Appendix II along with a summary of modifications made by individual groups. Actual protocols and data worksheets are found on the HEH website, now maintained by UC Berkeley's Center for Entrepreneurship in International Health and Development (CEIHD): http://ceihd.berkeley.edu/heh.IAPprotocols.htm

Household Selection. A Screening Questionnaire was used by field staff upon their first visit to each household to ensure that the household was suitable for and amenable to the study. If the head of the household agreed to be involved in the study, the field staff administered a Consent Form at that time. As with many rural areas where housing is not standardized, a wide range of different kitchen and stove configurations are encountered in the rural areas where the NGO stoves were deployed. The intent of the M&E was not to measure the affect of the ICS in all configuration, which would have required a much larger sample size and associated monitoring effort and cost, rather the intent was to measure the most common situation in these areas.

IAQ Monitoring. Particulate matter (PM) and carbon monoxide (CO) were measured in the same households both before and after the introduction of the ICS in a "Before-After" study design, without controls. Sample sizes were chosen to reflect the type of statistical analysis appropriate with this design, with suitable allowance for drops out, data loss, etc.

PM was measured using the UCB. This instrument is sensitive to particles of aerodynamic diameter less than approximately 2.5 microns, called fine PM or $PM_{2.5,}$ which is the size range thought to be most important for health. The monitors were produced and individually calibrated in the Indoor Air Pollution Laboratory at UC Berkeley prior to their use. The UCB recorded PM concentration every minute in its memory, which then could be downloaded into a personal computer afterwards. CO was measured with a commercial CO logger (HOBO), which also recorded concentration every minute. The HOBOs were purchased new and calibrated at the Indoor Air Pollution Lab at UC Berkeley using standard CO gas.. In the middle of the sampling campaign before the start of the "After" sampling, a co-location calibration check was conducted.

At each household, a UCB and HOBO were placed next to each other on the wall of the kitchen for 48 hours, using defined criteria (see protocols). At the end of each sampling session, a Post-Monitoring Questionnaire was administered to the main cook of the household. The questionnaire documented cooking and other activities that may have affected the kitchen IAQ during the monitoring period. The protocols and questionnaire used in this study, along with the HOBO CO Calibration Check Protocol and the Sampling Data Forms, are found on CEIHD's website. (Indoor Air Pollution Team, 2005). See Appendix I.

Stove Performance Testing. The controlled/laboratory stove test (Water Boiling Test – WBT) and field effectiveness measure (Kitchen Performance Test – KPT) are based on earlier versions developed by a US-based technical support organization in the mid-80s (VITA, 1985) and elaborated in a well-known technical report on stoves published soon after (Baldwin, 1986). Modifications to the tests derived for this project were arrived at after extensive testing conducted in partnership with the Aprovecho Research Center. In both cases, changes were made to both simplify and standardize the tests with the goal of ensuring comparability and replicability within and across different stove organizations.

Changes to the earlier tests included some minor procedural changes, the introduction of a "standard cooking pot" and the use of modern, but fairly inexpensive, measuring equipment. In addition, data analysis and calculation tools were incorporated into spreadsheet-based software to ease the analysis and standardize reporting methodologies. Detailed descriptions of the procedures, the changes that have been made relative to earlier versions of tests, and the analysis spreadsheets are available online. See Appendix II.

As the primary intention of this report is to evaluate the M&E package developed and deployed in this project, we only briefly describe here the summary IAQ and SP results achieved by the projects. These results will be presented in much more detail for each program and with more discussion in separate journal publications being prepared. As will be seen in these other publications, there are a number of complexities created due to the multiple stoves involved, changes in conditions at different measurement periods, and other factors that are not fully reflected in the summaries below. Although all the data presented here derive directly from the measurements made, therefore, some subjectivity and judgment were involved in deciding what to present as the final "bottom line"

It is important to point out that in some respects these data are of mainly historical interest at this point because the programs and their stoves have evolved over the course of the project. This is partly a direct outcome of the M&E efforts made by the NGO as they responded to the interim results seen from these M&E efforts and to field visits made by the M&E technical advisors. This illustrates the advantage of ongoing M&E for continuous improvement of ICS programs, beyond the final "bottom line" presented below.

Although there is a natural tendency to compare between stoves and between NGO projects, this can be quite misleading based on the data presented below, as stove designs, environmental conditions, fuel types, and cooking activities are quite different in each of these rural areas.

Indoor Air Quality (IAQ) Monitoring

There are considerable differences in the stoves disseminated by the three groups. The *Patsari* Stove disseminated by GIRA in the highlands of Mexico uses trunk wood and major branches. The stoves in both groups in India, ARTI and DA, were disseminated in rural areas where cow dung cakes, wood, and crop residues are used. In addition to the fuel difference, there are, of course, considerable differences in the traditional foods cooked in Michoacan, Mexico, and India, and the stove designs reflect these differences. For example on average almost two hours per day were spend making tortillas in Michoacan, necessitating a hot metal plate.

There were also considerable climatic differences between the rural areas in India. ARTI disseminated stoves in predominantly highland areas (although within their sample they had groups both in highland areas and in lower lying surrounding areas). The rural areas where stoves were disseminated by DA were predominantly flat agricultural regions.

Although CO and PM were measured in the same location, it would be expected that the reductions as a result of installation of an ICS would not exactly agree, since they are produced by different combustion processes, and behave somewhat differently (gas versus particles). The reductions for CO and PM for all groups, however, were generally in good agreement, especially for larger sample sizes.

Part of the protocol was to choose housing types that represent the most common conditions in the area, but not work in those that are infrequently found, in an attempt to reduce the variability associated with household measurements. This strategy is important as the variability is critical in determining the number of homes that are required to demonstrate statistical significance and the sample size is a significant factor in the cost, difficulty, time, intrusion, personnel, and other requirements of the field work. In spite of these efforts, the variability in all groups remained high as a result of different stove usage patterns. It is conceivable that stove usage patterns are the dominant factor in variability and should be addressed through more targeted screening questionnaires in addition to clearly distinguish differences in housing design and family size. In order to account for drop out rates, migration, differences in stove usage and measurement fallibility, one-third more houses were recruited and monitored than would be determined by sample size calculations. As a result of this intentionally conservative selection, the differences in IAQ results are statistically significant for all the projects. Even so, due to the diverse stove adoption patterns, an even more conservative sample size is recommended for future projects. This would allow more detailed analysis of the relative importance of the different household, stove, and fuel factors that affect IAQ in the households.

The IAQ summaries here refer to households in which the stoves were compared, i.e. traditional stoves in the "Before" phase and ICSs in the "After" phase and thus do not take into account those households that had stopped using the ICSs for one reason or another. They also do not indicate the presence of other combustion sources in the household environment, such as stoves for cooking animal fodder, which can influence results. Details will be provided in separate publications.

Appropriate Rural Technology Institute (ARTI): *Laxmi* stove

Here we present comparisons with three types of traditional stove. Type 1 is a two pot stove with a common combustion chamber in the center and 2 pot holes on either side in a v shape, with pot raisers to form a gap between the pot and the edge of the stove. Type 2 is a single pot stove with a combustion chamber directly below the pot hole, with pot raisers, and in some cases there may be two such single pot stoves placed side by side. Type 3 has a firemouth below the first pot, with pot raisers, with a second pot hole that is smaller in diameter and slightly raised that is attached to the combustion area via a duct. All types of traditional stove are without grate or vent. The basic structure is made of fired clay made by local potters. The outer structure and square shape is given by the housewife and is made of mud and bricks.

The improved vented *Laxmi* stove is a two pot stove with the combustion chamber directly below the first pot, and a slightly raised smaller second pot connected to the first chamber via a duct. A second duct connects this chamber to the flue, which passes through the roof and vents to the outdoors. The unvented *Laxmi* does not have this second duct leading to the flue. Both have grates. (See photo of the vented *Laxmi* below.)

ARTI's sample of 98 homes represents multiple communities in both highland and lowland areas. Since traditional stoves were different in the communities, and some communities retained a traditional stove or other cooking stoves (kerosene or charcoal) in the homes after installation of the improved stove, we present here two comparisons of changes in IAP levels as a result of installation of the improved stove, where the comparisons were not affected by the presence of other stoves. As stated above, further comparisons will be made on a community, stove and regional (highland versus lowland) basis in the journal articles.

Overall the ARTI stoves showed a significant reduction of 24% in particulate matter and 39% in carbon monoxide when all communities were pooled together (Table 1). ARTI successfully completed a large monitoring exercise, however, and has information about its ICS in both upland and lowland communities. Additional stratification in the database will appear in the published articles, but here we present evidence about the reductions achieved with the unvented *Laxmi* stove and the vented *Laxmi* stove in different communities.[3]

From these highly aggregated data, it is not possible to conclude that vented and unvented versions of this stove performed similarly because of the difference in comparison stoves, stove placement in the room, alternate fuels and stoves in use, and other issues not explored here. More analysis will be done to examine this further. That their performance may not be much different, however, is consistent with the stove performance test results which are also in the same range (Table 5). This suggests that most of the IAQ reduction may actually have been due to change in fuel use, not the performance of the chimney or improvements in combustion efficiency. This is supported by observations by the field staff that the chimneys were not working well in many of the vented stoves, but will be the subject of further analysis.

Table 1. All *Laxmi* stoves compared with all traditional stoves

		BEFORE			AFTER				Percent Change
	N	Average	Std Dev	maximum	Average	Std Dev	maximum	Wilcoxon SRT (p-value)	average
PM: (mg/m^3)	87	1.25	1.61	11.08	0.94	1.05	5.19	<0.007	24%
CO: (ppm)	98	10.82	8.71	40.0	6.65	7.10	47.3	<0.001	39%

Table 2a. Unvented *Laxmi* (Brahmnoli and Dhok Sanghvi communities) compared to traditional stoves 1 & 2

		BEFORE			AFTER				Percent Change
	N	Average	Std Dev	maximum	Average	Std Dev	maximum	Wilcoxon SRT (p-value)	average
PM: (mg/m^3)	21	0.95	1.21	5.57	0.48	0.55	1.70	<0.079	49%
CO: (ppm)	22	11.11	7.77	29.06	6.91	7.02	34.6	<0.024	38%

[3] For most purposes, a p-value of 0.05 or less is conventionally considered "statistically significant", i.e., there is less than a 5% probability that the change in IAQ levels found between improved and traditional stoves was merely the result of chance variation in the two groups.

Table 2b. Vented *Laxmi* (Sone Sanghvi community) compared to traditional stove 1

		BEFORE			AFTER				Percent Change
	N	Average	Std Dev	maximum	Average	Std Dev	maximum	Wilcoxon SRT (p-value)	average
PM: (mg/m^3)	27	1.79	2.17	11.09	0.99	1.23	5.20	<0.06	45%
CO: (ppm)	30	15.3	9.11	33.30	8.37	10.22	47.3	<0.008	45%

Development Alternatives (DA): *Sukhad*

The traditional stove in the region serviced by DA is a single pot U-shaped stove made of mud without a chimney. It can be portable or fixed. The *Sukhad* stove is a two-pot mud stove with chimney. (See photo of the *Sukhad* stove below.) It can be used with either wood or agricultural residues and is suitable for a medium-size family (5-8 members) using flat or spherical bottom vessels 19-30 cm in diameter. It also provides strong heat to the second pothole. The *Sukhad* is designed such that the pothole is raised by about 6 cm above the level of the first pothole to avoid interference between pot rims when cooking with two large pots. It can be constructed by using fabricated mould of mild steel sheet with locally available clay materials. The chimney pipe and cowl are made of asbestos cement. Optional items like pottery liners and cost iron grate can also be provided.

DA evaluated the change in indoor air pollution due to the installation of the *Sukhad* in approximately 60 households located near Jhansi, Uttar Pradesh, India from July 2004 to September 2005, using a “before and after” study design. The data represent measurements taken in several communities during three seasons, including one year after installation (Table 3).

Table 3. Development Alternatives: *Sukhad* stove - paired before and after comparisons of 48-hour averages

a. Monsoon ~1-2 months after installation

		BEFORE			AFTER				Percent Change
	N	Average	Std Dev	maximum	Average	Std Dev	maximum	Wilcoxon SRT (p-value)	average
PM: (mg/m^3)	30	0.52	0.75	4.14	0.33	0.39	2.09	<0.02	36%
CO: (ppm)	37	7.88	6.73	29.9	5.38	3.89	18.9	<0.01	32%

b. Winter ~ 6 months after installation

		BEFORE			AFTER				Percent Change
	N	Average	Std Dev	maximum	Average	Std Dev	maximum	Wilcoxon SRT (p-value)	average
PM: (mg/m^3)	na[1]	na	na	na	na	na	na	na	na
CO: (ppm)	36	7.90	6.8	29.9	5.5	4.7	17.6	Not sig.	30%

[1] UCB are not reported during this monitoring period for data quality issues.

c. Summer (n=15) –sub-sample 1 year after installation

		BEFORE			AFTER				Percent Change
	N	Average	Std Dev	maximum	Average	Std Dev	maximum	Wilcoxon SRT (p-value)	average
PM: (mg/m^3)	15	0.65	1.01	4.14	0.36	0.47	1.93	<0.01	44%
CO: (ppm)	15	8.67	7.8	29.9	2.68	2.8	8.4	<0.001	69%

The reduction in particulate levels was generally around 40% for the *Sukhad* stove; with reductions in CO around 30%. Although a much higher reduction for CO was seen in the summer, the sample size was much smaller than in other seasons and this decreased the statistical confidence in the estimate given variability of the measurements.

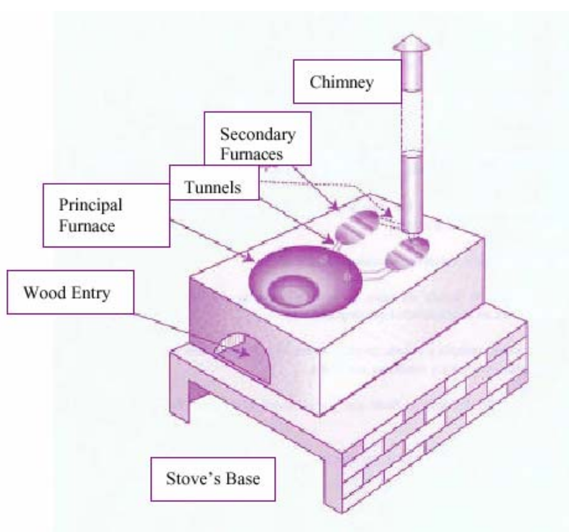
Grupo Interdisciplinario de Tecnología Rural Apropiada (GIRA): *Patsari* stove

In addition to the three-stove fire common in other parts of world, a "U type" open fire is used for cooking in Mexico, as in India. The U-type open fire ring is made out of mud or clay and is built by the users. These devices take the form of an "U" or horseshoe, and although they "enclose" the fire to a certain extent in what might be considered a combustion chamber, they do not have a chimney.

The *Patsari* departs from a modified *Lorena* that was previously disseminated in Mexico and has the following improvements: (1) optimized design of the combustion chamber and tunnels, (b) custom-designed parts for durability, including a metal chimney support and a ceramic stove entrance (see diagram below); and (3) reduction in construction time and standardized inner dimensions (Masera et al., 2005). (See diagram of the *Patsari* stove below.) The cookstove is made in approximately 2 hours with the aid of a metallic mould that ensures that critical dimensions are maintained. The stove also comes with metal *comales* -- plates to replace the pots -- that are sealed to avoid leaking of the smoke. Currently two models are disseminated, having either one and two fuel entries. The former has one combustion chamber and uses a metal *comal* of 52 cm diameter. It is ideal for cooking tortillas and is preferred by mixed fuelwood-LPG users.

The second *Patsari* model has two combustion chambers. The main one usually supports a ceramic *comal* (which is preferred by many users) for making tortillas. The smaller chamber has a metal *comal* of 35 cm diameter designed for cooking other dishes, such as beans, and other tasks, such as boiling water. Both models include tunnels that conduct the combustion gases to secondary chambers. Each chamber includes baffles to improve heat transfer between gases and the *comal.* These secondary chambers are used for ‘‘low power’’ cooking tasks, such as keeping food warm or heating water.

The body of the *Patsari* is made of a mixture of sand and mud and a small amount of cement. All the materials are available locally; the custom-made stove parts are also manufactured by local small industries. External cookstove dimensions are 80 cm by 100 cm (100 cm by 100 cm for the two-entry model), stove height is 27 cm and the distance from the base of the combustion chamber to the *comal* is 20 cm.



Because the HEH project served as a nucleus for a much larger scientific monitoring exercise, GIRA successfully accomplished the largest and most ambitious monitoring exercise in collaboration with both national and international institutions. Facilitating such a large ongoing scientific collaboration is itself a laudable output of the HEH program and helped in providing the evidence that resulted in GIRA being awarded the prestigious Ashden Award for health and welfare[4]. Of the HEH projects, GIRA was able to provide considerable scientific validation of the monitoring approaches used by HEH colleagues in India, and continues to monitor in other communities that show different stove adoption characteristics. Here, we only present the results related directly to the HEH project as the other work will be presented elsewhere (Table 4).

Table 4. GIRA: *Patsari* Stove – paired before and after comparisons of 48-hour averages

		BEFORE			AFTER				Percent Change
	N	Average	Std Dev	maximum	Average	Std Dev	Maximum	Wilcoxon SRT (p-value)	average
PM: (mg/m^3)	**33**	**1.02**	**0.79**	**4.23**	**0.34**	**0.27**	**1.16**	**<0.001**	**67%**
CO: (ppm)	**32**	**8.88**	**4.44**	**22.61**	**3.02**	**2.66**	**12.1**	**<0.001**	**66%**

For both CO and PM approximately 70% reduction in kitchen concentrations was achieved by use of the improved *Patsari* stove. Although there were a limited number of houses that incorporated a traditional stove in another room or outside (predominantly for heating water), similar percentage reductions were also observed in personal exposures (not presented here). Thus the ICS effectively decreased concentrations of PM and CO in kitchens where participants spent their time. Project results also agree with those of Zuk et al (2006) who showed that near-house and nearby ambient PM concentrations were 0.094 mg/m^3 and 0.059 mg/m^3 prior to the intervention and that there was no significant change in these concentrations after the intervention. Nearly all households (99%) reported no burning of trash, and tobacco consumption and nearby vehicle traffic were low.

Stove Performance Tests (SPT)

The main stoves disseminated by each NGO were described in the preceding section. Two types of SPTs were performed by most groups.

Water Boiling Test (WBT)[5]

The WBT has three components: a test at high-power that is conducted with both cold and hot-start conditions and a test at low-power to simulate slow cooking tasks like the cooking of beans or hard grains that is common in many developing countries. In addition, the WBT was designed to provide testers with a range of stove performance indicators because different indicators are important to different people. For example, stove technicians may be concerned about thermal

---

4 http://www.ashdenawards.org/winners/gira

5 Although HELPS International has yet been not able to complete SPTs under the Shell Foundation HEH program, a WBT was conducted on a version of their stove by Aprovecho in 2004. This report is available from Helps International.

efficiency while donors may care more about fuel consumption, and stove users might place more value on time the stove takes to boil a fixed quantity of water. The analysis software provided by the project was designed to calculate all of these indicators. Each NGO conducted tests of the local version of a "traditional stove" and their ICS using wood. These are shown below in Table 5 in the form of relative changes between the traditional stove and the ICS for each organization, some of which tested more than one traditional or ICS.

For each test, Table 5 shows improvements between the ICS and the traditional stove in black. Any ICS that performed less well than the traditional stove against which it was tested is shown in red. For example, the traditional stove boiled water 29% faster than ARTI's *Bhagya Laxmi* (unvented) stove, but the *Bhagya Laxmi* had 5% greater thermal efficiency. No ICS performed better than its traditional competitor in all categories of tests. Similarly, those ICSs that generally performed less well than the traditional competitor still did better in some aspects. Hence, according to the WBT, there were apparently no unambiguous winners or losers among the stoves being promoted by the Shell HEH grantees at the time the test were done.

Table 5: Changes in performance of improved stove relative to traditional stove for selected indicators. Improvements between the ICS and the traditional stove are shown in black; any ICS that performed less well than the traditional stove against which it was tested is shown in red. The *Bhagya-Laxmi* is the unvented stove and the *Griha-Laxmi* is the vented stove.

	GIRA		**ARTI**			**DA**			**All**
	U-stove - *Patsari*	**3SF - *Patsari***	***Laxmi***	***Bhagya Laxmi***	***Griha Laxmi***	***Sukhad***	***Laxmi***	***Anandi***	**average**
High Power-Cold Start									
Time to boil Pot # 1	253%	253%	147%	29%	-16%	61%	81%	63%	116%
Thermal efficiency	-61%	-46%	-38%	5%	10%	51%	14%	-6%	-3%
Specific fuel consumption	269%	162%	-3%	-42%	-52%	-25%	2%	18%	19%
High Power-Hot Start									
Time to boil Pot # 1	236%	213%	60%	6%	-25%	26%	35%	-5%	63%
Thermal efficiency	0%	-11%	-48%	4%	-9%	9%	74%	7%	1%
Specific fuel consumption	30%	43%	101%	-6%	4%	-13%	-35%	-14%	2%
Low Power - Simmer									
Thermal efficiency	100%	58%	47%	167%	27%	4%	4%	-10%	40%
Specific fuel consumption	-31%	-34%	24%	7%	23%	-30%	-45%	58%	-22%

DA's ICSs tended to be more efficient than their traditional stove in the high-power tests. ARTI's stoves were mixed, and GIRA's tended to perform less well. This is shown in Table 5.

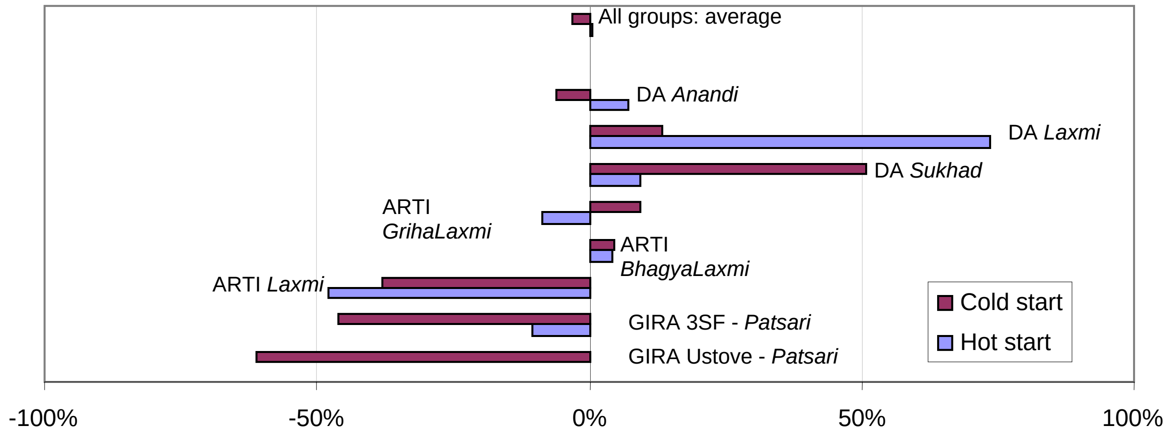


Figure 1: High-power test – percent change in efficiency between improved and traditional stoves.

Nearly every ICS took longer to boil 5 liters of water than the traditional counterparts, which is shown in Figure 2. In some cases this was because the stove is designed for low power cooking tasks, where high power output is detrimental to the dish being cooked. In addition as noted in the feedback to the tests, the volume of water used in these tests was inappropriate for some stoves. The increased cooking time is not a surprising result as all the stoves had high thermal mass, which take longer to heat up than lighter stoves or open fires. Such stoves tend to cook more slowly than an open fire when cold. All massive stoves, however, showed some improvement in hot-start conditions, which is evident in Figure 1.

Although the overall results are ambiguous for the high power tests, the ICSs tended to perform better in the low-power or simmer tests. All but one outperformed the traditional counterpart in thermal efficiency, as is shown in Figure 3.

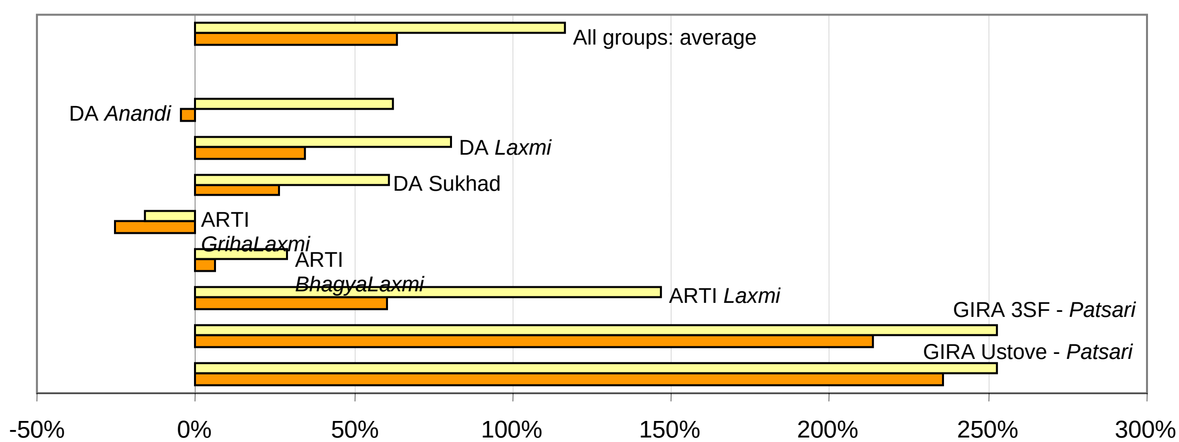


Figure 2: High-power test – percent change in time to boil between improved and traditional stoves.

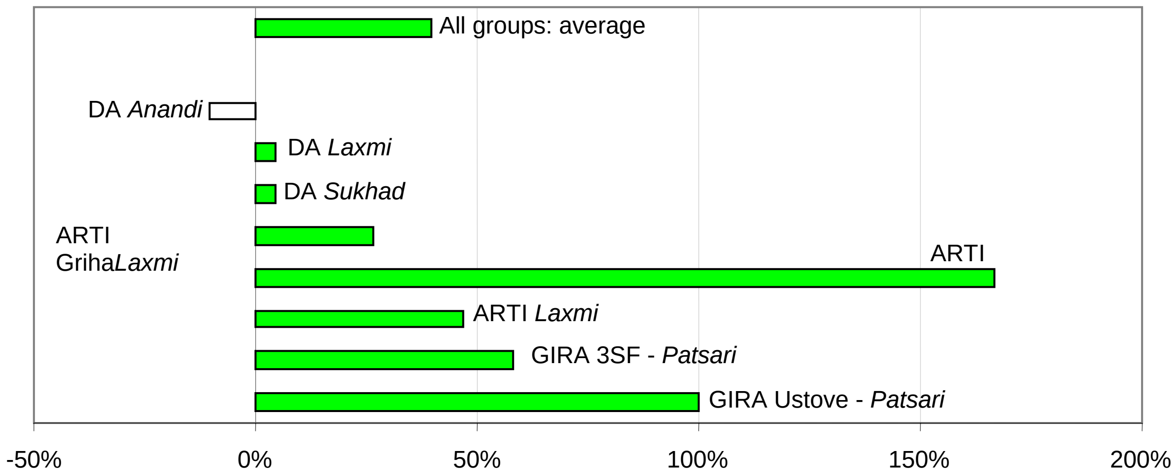


Figure 3: Low-power (simmer) test – percent change in efficiency between improved and traditional stoves

Note, only three or four tests of each stove are reported here. With small sample sizes, many of the differences shown in the table and figures above may not be statistically significant. A more detailed analysis is planned for journal publication.

Kitchen Performance Tests (KPT)

There are two main goals of the KPT: (1) to compare the performance of improved stove(s) to the common or traditional stoves or to other improved stoves as they are used in the kitchens of real families and (2) to identify qualitative aspects of stove performance through a simple survey. The latter component will not be discussed in this report. This type of testing, when conducted carefully, is the best way to understand the stove's impact on fuel use and on more general household characteristics and behaviors because it occurs in the homes of stove users (VITA, 1985).

Although based on a simpler concept than the WBT, the KPT seems to be more difficult for organizations to conduct in practice because of complicated sample selection procedures and logistical issues of working in real households. The former are particularly important because the variability in measurements of household fuel consumption tends to be higher than the variability observed in lab-based testing, hence larger sample sizes are needed to obtain statistically significant results. Several groups could not overcome this difficulty and followed different procedures, which reduces the effectiveness of the KPT. This is discussed further below.

Another difficulty with the KPT was that it was originally designed to accommodate only one fuel. While in some parts of the world households still rely on only one fuel for cooking, the organizations funded by Shell for this study work in communities where either the use of more than one fuel in a single stove, or multiple stoves, is quite common. It is intuitive that a single ICS would likely impact energy consumption across all stoves and fuels used in the household. Therefore, grantees adapted the procedure to deal with multiple stoves and fuels. Using their feedback, the procedure itself has been formally modified to account for these complications in the future. Thus, the results given here are in terms of total energy consumption and energy savings, rather than wood consumption or wood savings, which has been the quantity of interest in past stove projects. Each group's findings are discussed briefly below.

One additional aspect of monitoring for stove performance that is associated with the KPT is the follow-up (post-intervention) qualitative surveys. Like field testing, this aspect of stove monitoring is logistically difficult and may be overlooked by NGOs, particularly when resources are limited. This need is shared with that for IAQ studies.

Appropriate Rural Technology Institute (ARTI)

ARTI selected communities that were stratified by climate (high and low rainfall). In the high rainfall areas, six villages were selected: Nanegaon, Shiuli, Shivne, Khadakwadi, Yeolewadi, and Brahmnoli. In this area households use a combination of fuels: cow dung cakes (CDC) and fuelwood. People use the same variety of wood and use CDC and wood in more or less the same proportion. Many villages were selected in the high rainfall area because ARTI feared that people might drop out from the study. In addition, only a small percentage of the people were willing to buy an ICS.

In the low rainfall area only two villages were selected: Dhok Sanghvi and Sone Sanghvi. ARTI opted for smaller coverage there because they found that people were more enthusiastic and most households wanted a vented 2-pot stove. There, people use only wood as fuel; CDC is used only as a fire-starting material if at all.

The traditional cookstove (TCS) was different in the two areas. In the low rainfall area the TCS was exclusively a 1-pot unvented stove. In the High rainfall area the TCS was a 2-pot unvented stove except for some houses in Brahmnoli where a completely different type of TCS stove was being used.

People in the different climatic areas also opted for different models of ICSs. In the low rainfall area, all of the participants bought a *Laxmi* stove. In contrast, the majority of the people in the high rainfall area bought the unvented *Laxmi*-Bhagya.. Many people purchased the unvented *Laxmi* because they did not want to drill a hole in the roof and risk leakage during the rains. Every household in Nanegaon, Shivne and Shiuli adopted the unvented *Laxmi*. Participants in the other two villages (Khadakwadi and Yeolewadi) of the high rainfall area, however, adopted the vented *Laxmi* but stopped using it before the 12-month study was over.[6]

Data reported in this report only come from households in the high-rainfall area, though ARTI had trouble persuading participants to agree to the 7-day KPT and thus resorted to a modified version of the KPT. ARTI conducted 51 one-day tests spread across three communities, seven three-day tests in a fourth community (Yeolewadi), as well as four 7-day tests (one each in Nanegaon and Khadakwadi and two in Shiuli).

ARTI's results are complicated by some use of CDCs in both TCSs and ICSs. Moreover, in some homes, the ICS was not the only stove used in daily cooking activities. Among their 69 one-day

[6] The participants dropped out of the study because subsidized LPG was provided to them by another NGO working in the area.

and three-day tests, they include only 28 households that were using a traditional solid fuel stove without an improved solid fuel stove also in the house. However, 12 of these also use LPG, the consumption of which was not measured. Another 15 households listed the traditional stove as the primary stove, but also used an ICS at times. Similarly, 38 listed either an improved stove or a fossil fuel stove as the primary cooking device, but 13 of these list a traditional solid fuel stove as their secondary device (Table 6).

One community (Nanegoan) had a clear delineation in stove usage, however. The households using the traditional stove for its primary cooking fuel did not use an ICS and vice versa (the results shown on the left in Figure 4). The improved-stove users averaged 26% less fuel, but due to the large variation, this difference is only significant with 90% confidence, not quite meeting the normal 95% criterion. This difference is illustrated by a blue oval in Figure 4.

In addition, it is also possible to look at the aggregate differences among all communities. Here, several comparisons can be made from the available data. First, the difference in fuel consumption observed between households listing the main stove as traditional and households listing the main stove as either an ICS or a fossil fuel stove (shown on the right in Figure 4). This 27% difference is statistically significant and is also illustrated by a blue enclosure in Figure 4. Similarly, the total number of fossil fuel users could be divided between those who also had ICSs and those who had TCSs. The former group consumed less wood than families using TCS as their primary stoves. No other comparisons wee statistically significant.

Further, we must note that fossil fuel use was not included in the analysis. Thus, while biomass energy consumption is lower, no conclusions can be drawn about overall changes in cooking energy consumption.

Table 6: ARTI's observed differences in energy consumption (MJ/person-day) in households using TCS and ICS in all high rainfall communities

	n	Average	St Dev	% change with TS	p-value
Main stove is TCS	31	27.7	11.3	-27%	0.002
Main stove is ICS or FF	38	20.1	7.9	-21%	0.010
FF users with ICS	34	21.8	5.7	-9%	0.425
FF users w/o ICS	28	25.1	13.6	-15%	0.072

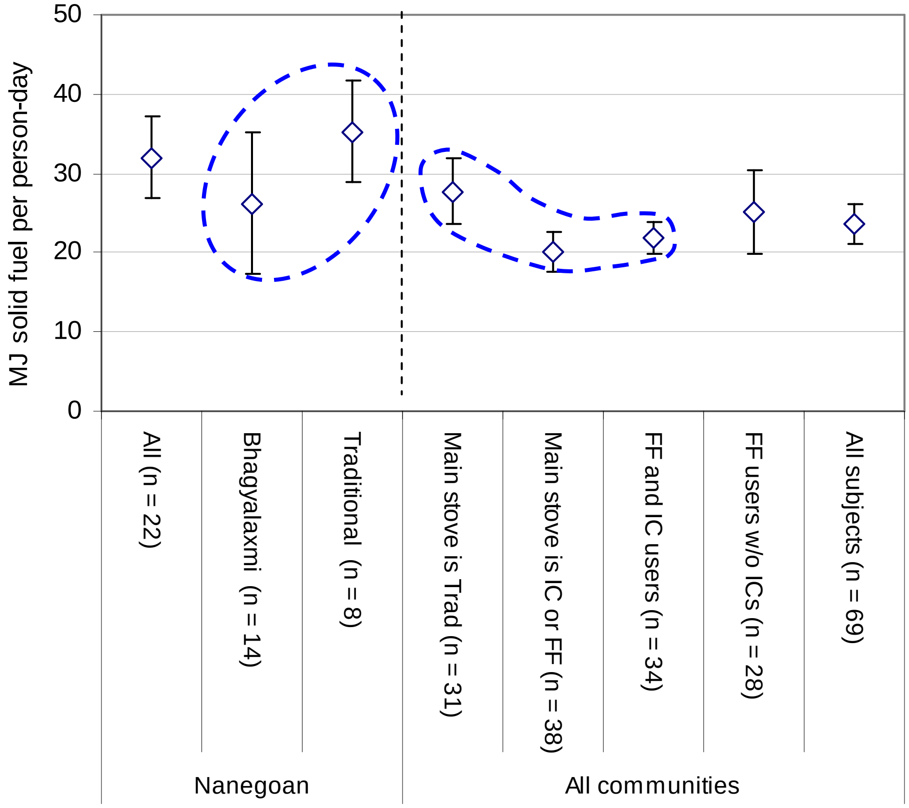


Figure 4: Energy consumption measured in ARTI communities (MJ/person-day) based on 1-3 days of measurements [mean ± 95% CI]. [7] **The blue enclosures depict differences in fuel consumption between households using improved or fossil fuel stoves and households using traditional stoves that are significant with greater than 90% confidence.**

Development Alternatives (DA)

DA conducted KPTs in three communities during three different seasons. They found an improvement trend in every test and some apparent differences by village and season, but statistical significance has not yet been calculated based on the number of days of testing in each household. Recognizing these limitations, their tests showed, on average, a 21% improvement in fuel consumption. Their results are depicted in Figure 5. Further analysis will appear in the journal publications.

[7] In this graph, ICS = Improved cookstove and FF = fossil fuel (LPG or kerosene).

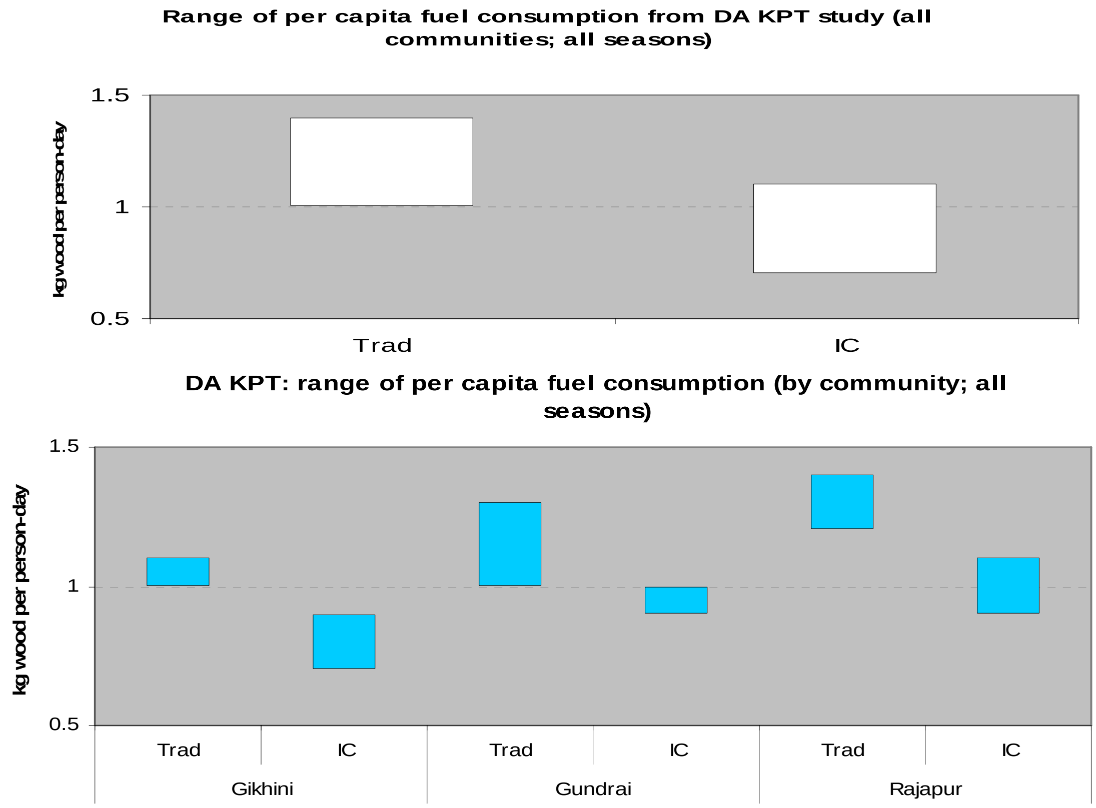


Figure 5: Results of DA's KPTs.

Grupo Interdisciplinario de Tecnología Rural Apropiada (GIRA)

GIRA followed the procedure for the KPT closely. The only adjustment made was to divide their sample into two groups in order to differentiate households using LPG and households not using LPG. They initially selected 23 households using LPG and wood together and 20 households using only wood for pre-intervention testing. However, it appears that they had a substantial number of dropouts as they only they tested 8 and 6 households respectively with ICSs. Nevertheless, they found sufficient differences in households using ICSs such that their results remain statistically significant in spite of the attrition. What is not possible to evaluate without more detailed household surveys, however, is how atypical these loyal ICS users are compared to the population as a whole. Their results are summarized in Table 7.

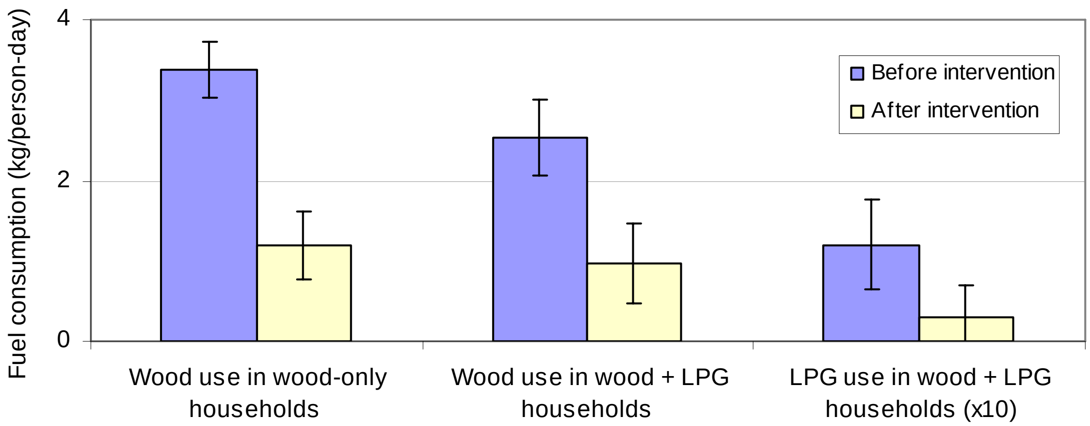
Table 7: Total energy consumption (MJ/person-day) in households before/after receiving the *Patsari* stove*

	Pre-adoption – traditional stoves				**Post-adoption – improved stoves**					
	n	**Average**	**St Dev**		**N**	**Average**	**St Dev**		**% diff**	**p-value**
Households using only firewood	23	67.6	16.2		8	23.8	10.6		65%	< 0.001
Households using firewood and LPG	20	56.7	21.9		6	20.7	12.5		63%	< 0.001

*Analysis with only the stoves that were in both groups does not appreciably change the results.

The differences between traditional stoves and *Patsari* stoves are significant with > 99% probability. Interestingly, GIRA not only found that overall energy consumption went down with the adoption of the *Patsari* stove, but also that both wood and LPG consumption went down by 75% in households using both fuels. This result is also significant with > 99% probability and is illustrated in Figure 7. According to discussions with GIRA, LPG consumption went down because the *Patsari* stove could be used to perform cooking tasks for which cooks previously relied on LPG stoves such as boiling water, which would seem to be a major benefit.

Figure 7: GIRA's daily fuel consumption in "wood-only" and "wood + LPG" households (note LPG is multiplied by a factor of 10 in order to be visible on the same scale). Error bars show 95% confidence intervals.



GIRA's good effort in the KPT is particularly notable because their stove performed quite poorly in the high-power phases of the WBT. This demonstrates that for the stoves and households in this study **there is a not a direct relationship between the WBT and fuel consumption in real household conditions**. This must be kept in mind when funding and evaluating stove projects. Money and training should be made available to conduct these difficult but valuable field tests, which are often overlooked in favor of simpler lab testing.

Extending the HEH project protocols, GIRA also conducted Controlled Cooking Tests (CCT), which are intermediate in difficulty and representativeness between the WBT and KPT. A CCT compares fuel use in a locally appropriate cooking task with a standard amount of food and thus the results cannot be compared across regions. Tortilla making was chosen because it is the most intensive energy task within local households. The single-entry *Patsari* scored the best (0.63 ± 0.03 kg fuelwood/ kg tortilla). Savings from *Patsari* stoves reach 67 % compared to the U-type fire and 59 % compared to the three-stone fire, and were statistically significant at 95 % confidence level (Masera, et al., 2005).

Comparison of IAQ Results with WHO Guidelines

The World Health Organization (WHO) sets air pollution guidelines to offer guidance in reducing health impacts of air pollution (both indoor and outdoor) based on current scientific evidence. The WHO recently set new Air Quality Guidelines (AQG) for PM2.5, ozone, nitrogen dioxide, and sulfur dioxide, along with interim targets which are intended as incremental steps in a progressive reduction of air pollution in more polluted areas (WHO, 2005). The guideline for carbon monoxide was set in 2000 (WHO, 2000).

The results of the IAP monitoring in the project households are compared to the World Health Organization's AQG and interim target-1 in Tables 8-10 below (WHO, 2005). Note that the CO concentrations reported above in parts per million (ppm) were converted to mg/m^3 to match the units used by WHO.[8]

Table 8. ARTI (all *Laxmi* stoves): Comparison of kitchen concentrations to WHO guidelines.

	Before [traditional stove] (48-hr ave)	**After [improved stove] (48-hr ave)**	**WHO interim target-1**	**WHO Air Quality Guideline**
$PM_{2.5}$	1.25 mg/m^3	0.94 mg/m^3	0.075 mg/m^3 (24-hr mean)[1]	0.025 mg/m^3 (24-hr ave)[1]
CO	12.4 mg/m^3	7.6 mg/m^3	NA	10 mg/m^3 (8-hr ave)[2]

[1] WHO, 2005. [2] WHO, 2000.

It can be seen that although both the unvented and vented *Laxmi* stoves together achieved statistically significant reductions in both particle levels, the lowered particle concentrations were still far above WHO guidelines, even the most liberal Interim Target-1. Although the 48-hr ICS CO levels were lower than the 8-h WHO AQG, they were undoubtedly near zero at night when no burning occurred and thus probably exceeded the AQG during periods of the day when cooking occurred. Detailed analysis of IAQ levels by time of day and other parameters will be done for the journal publications.

Table 9. DA: Comparison of kitchen concentrations to WHO guidelines.

	Before [traditional stove] (48-hr ave)	**After [improved stove] (48-hr ave)**	**WHO interim target-1**	**WHO Air Quality Guideline**
$PM_{2.5}$	0.52 mg/m^3	0.33 mg/m^3	0.075 mg/m^3 (24-hr mean)[1]	0.025 mg/m^3 (24-hr ave)[1]
CO	9.02 mg/m^3	6.17 mg/m^3	NA	10 mg/m^3 (8-hr ave)[2]

[1] WHO, 2005. [2] WHO, 2000.

The 48-hour average particle concentrations measured after one month of installation, and after 1 year of installation were approximately 4.5 times higher than the 24 hour average Interim Target-I established by the WHO. Although the stoves are continually improving and several iterations have passed since initiation of the M&E, it is hard to envisage how the WHO Interim targets could be achieved in indoor environments in these settings.

[8]By multiplying by the gram molecular weight of CO, 28, and dividing by the conversion factor of 24.45.

Table 10. GIRA: Comparison of kitchen concentrations to WHO guidelines.

	Before [traditional stove] (48-hr ave)	**After [improved stove] (48-hr ave)**	**WHO interim target-1**	**WHO Air Quality Guideline**
$PM_{2.5}$	1.02 mg/m^3	0.34 mg/m^3	0.075 mg/m^3 (24-hr mean)[1]	0.025 mg/m^3 (24-hr ave)[1]
CO	10.2 mg/m^3	3.5 mg/m^3	NA	10 mg/m^3 (8-hr ave)[2]

[1] WHO, 2005. [2] WHO, 2000

.

Mean 48-h baseline open fire measurements (0.95 mg/m^3) fall in the range of other Mexican studies (Saatkamp 2000, Riojas-Rodriguez 2001) but post-intervention measurements still remain 5 times higher than Mexican Ambient Standard of 0.065 mg/m^3 (SSA, 2003). These fall within the range of comparable studies worldwide (Saksena et al 2003). Similar to the other HEH projects, the 48-h average kitchen concentrations were 4.5 times the WHO interim target. Although not part of the HEH project, average 48-h outdoor concentrations in the yard near the houses (0.094 mg/m^3) also exceeded the WHO interim standard, probably from pollution derived from the households.

Since CO concentrations measured by GIRA with the improved *Patsari* stove were considerably reduced, the WHO 8-hour guideline may not have been exceeded by the average CO concentrations measured by this ICS program. As with the CO results from the other groups, however, further analysis will be needed to verify this judgment.

In summary, the WHO AQGs for PM were still substantially exceeded by all stoves, but CO levels lowered to near the AQG for some.

Summary and Recommendations for Future Stove Dissemination Projects

We do not attempt an overall assessment of the achievements of the individual stoves here, which will require more detailed analyses and additional information that were not gathered as direct components of IAQ and SP monitoring. These details are planned for publication in a special issue of the journal, *Energy for Sustainable Development* in 2007, assuming normal acceptance after peer review. From these limited data, however, we do offer one observation.

In two of the sites, the reductions in pollution roughly matched those in fuel use, although in the third, IAQ may have reduced a bit more. This might indicate that for some stoves much or all of the benefits of each type came from improving the heat transfer into the pots and not from increases in combustion efficiency of the fires nor in stove venting (reliably working chimneys). Before making such a conclusion, however, further analysis is needed.

Based on the projects' experience in developing and implementing the M&E techniques described in this report, a series of recommendations for conducting M&E for IAQ and SP in future stove dissemination programs is presented below.

IAQ Monitoring Recommendations[9]

As noted in the introduction, unlike stove performance, M&E methods for establishing changes in IAQ have not been developed previously for ICS programs. Related methods used in research applications required significant modification, simplification, and standardization in this first effort and a number of lessons have been learned in the process, which are:

a. The project achieved its main objective in providing an estimate of the IAQ changes due to the ICSs for the three programs in a manner consistent with standard quantitative methods. In this the NGOs should be commended.
b. The groups, however, exhibited important differences in the extent of their achievement. GIRA, which is associated with both national and international research groups, were able to accomplish most. They conducted significant monitoring beyond the scope of the HEH program itself, including assessment of personal exposures, measurements in other household location, as well as measurements of volatile organic chemicals, and will undoubtedly continue with this kind of work. DA and ARTI conducted their HEH work well, but found it more challenging and resource-intensive than they expected, partly because of the initial difficulties with the UCB monitor. HELPS, in contrast, was not able to mount either systematic IAQ or SP testing. This, however, is mainly because they do not operate as a grass-roots dissemination organization with their own field staff and close connections to local communities. Their highly successful role is as an intermediate stove manufacturer and provider of ICSs to grass-roots organizations. Thus, in retrospect, it was probably unrealistic to expect them to conduct the hands-on field work required in the HEH protocols.
c. The decision to train NGO staff in simplified statistical theory and analysis methods may have been too ambitious in retrospect. Although intelligent, motivated, and often well educated, NGO staff were more confident when they had direct assistance from UC staff. An even more “cook book” set of methods may have been more appropriate, although this approach risks misapplication when complex and unexpected field conditions are encountered (which is frequently the case).
d. Household field studies to derive quantitative conclusions on elusive metrics such as air pollution are difficult to undertake, even by trained research groups. Thus it is not surprising that NGO staff found it so as well. Part of the problem was a general underestimate of time and other resource requirements necessary for this kind of work, thus sometimes making the staff’s work more difficult.
e. Drop out of households between sampling rounds plagued all the groups requiring care in choosing sample sizes with sufficient margins of safety and extra effort in motivating participation. Such losses occurred not only because households refused to continue, were constructing separate kitchens, had moved away, or otherwise became unavailable, but also because of significant continuation in the use of traditional stoves and transitional adoption patterns. Although this would be expected from social theories of technology dissemination within populations,

[9] See Appendix III for a summary of the team responses to evaluations of the IAQ methods.

this complicates collecting and analyzing the IAQ data in a consistent manner. Clearly, however, if improved adoption rates of the ICS are not seen over time, this would be a disappointing outcome of so much effort and resources to disseminate stoves.

f. Multiple fuel and stove use in some areas complicated the collection and interpretation of changes in fuel use and air pollution. In such areas, future studies may have to do more careful stratification of households to obtain the statistical power to make judgments.
g. Although in some ways the best for showing actual changes, one drawback of the Before-After study design is that it is difficult to implement in a market-based dissemination effort, in which it cannot be easily predicted who will buy and the motivation to participate is lower than with other types of dissemination. In such cases, a cross-sectional design may be more feasible, if less satisfying statistically and requiring greater sample numbers.
h. Referring back to the choice of three principal measurement types, even though the best indicators of actual changes in the population, the "effectiveness" field-based measures deployed in the HEH project may not be suitable for many NGOs in the future because of the difficulties noted above. There will be exceptions, of course, such as illustrated by GIRA.
i. Using lab tests for IAQ assessments has not been validated and would seem to suffer from all those drawbacks related to SPTs (see below) plus others. In particular, there seems no way to realistically simulate either the leakage of smoke into a household from normal use of a chimney stove or the re-entry of outside smoke from the chimney back into the house. Research support for validating the relationships among IAQ measurements in the different settings (lab, efficacy, and effectiveness) are needed.[10]
j. There is, however, an intermediate "efficacy" test that could be deployed that might be an appropriate alternative in many circumstances. This is the "test house" approach in which one or at most a few houses rented for the purpose are monitored as they change from no cooking, to traditional stove use, to ICS use, but using real cooks and cooking. It essentially reveals what the maximum improvement might be, but does not directly reveal what will be achieved on a population basis. It has also been recently used in a study in China after training by the UC Berkeley group, although results are not yet available. More description of this method is found in Naeher et al. (2000).
k. As planned, the two types of IAQ instruments (UCB for particles, HOBO for CO) were easily deployed in the field by the NGOs after training. They also held up well under field conditions. Since this involved a lot of development, as these are the first projects globally where such an approach has been tried, the initial timelines of this M&E assessment were underestimated. With subsequent refinement of the techniques, subsequent projects implemented by CEIHD with

[10] As part of expanded research efforts triggered by the HEH project, GIRA has conducted preliminary exploration of how well both laboratory WBTs and household WBTs perform in predicting emission., To date, analysis indicates little relationship. Emissions data from real households, however, tend to agree with the KPT data in terms of wood savings reported.

somewhat simpler protocols have completed the assessment with a timeline similar to that initially estimated (Pennise et al., 2006).

l. Because both instruments produce many data (minute averages over 48 hours), however, and require some manipulation of their outputs to produce reliable concentration measurements, data handling, cleaning, processing, and interpretation took much longer than anticipated, necessitating extra training sessions and more extensive involvement by the UC team than planned. These requirements were reduced late in the project to some extent by further development of the software to facilitate batch processing of files, but could not be entirely automated by the very nature of such continuous monitoring.
m. Although there are no commercial alternatives that meet the cost criteria of the UCB, the UCB is a prototype monitor, which underwent significant development during the course of these HEH projects. While the intent was for NGO to input summary values into a database, in practice our direct participation was required in additional checking of data files, analysis, data processing, and data cleaning. Future efforts are unlikely to be hampered by these developmental concerns, but it did introduce significant delay into the timeline of the M&E.
n. Future studies may wish to deploy devices designed to give summary values only, thus reducing the need for data handling and processing. Although providing only a small fraction of the information available from continuous monitoring (nothing about peaks or temporal distributions, for example), simple means may be sufficient for many purposes of interest to NGOs. At present, unfortunately, although there are simple devices for giving means for CO concentrations, no easily deployed devices exist for particles.

Given the experience of this project, therefore, we have the following major recommendations for future large-scale NGO-driven ICS dissemination efforts,

1. Consideration be given to relying principally on efficacy measures in the form of test house studies in typical households. These studies can be done using outside professional monitoring teams from universities, research institute, or environmental consulting firms, if the NGO is not interested in developing this capacity on its own. In this way, more expensive and reliable equipment can be used, because highly trained people will be deploying it for short periods only, after which it can be used elsewhere.

2. Field surveys of pilot interventions are necessary, however, but can focus on assessing[11]
    a. the degree to which the ICSs are actually put to use
    b. the extent to which traditional stoves remain in use (with or without the ICS)
    c. the durability of the ICS in real kitchens
    d. householder suggestions for improvements to the design or construction of the stove to keep maintain good IAQ, which may only be revealed after extended use (perhaps best revealed by focus group discussions)

[11] We do not elaborate further here partly because of space but also because the report from the teams working with the Liverpool University group will go into this further with examples from their investigations.

3. Large-scale effectiveness monitoring efforts only be undertaken with appropriate resources, training, and partners and after results from the above procedures indicate that serious population benefits may be occurring.

SP Test Recommendations[12]

Monitoring and evaluation has been a weak point in stove programs for some time. This effort represents one of the first efforts to incorporate a rigorous M&E methodology into an improved stove program – particularly one spanning continents. Although there have been some rough patches, the efforts of the grantees should be commended. Whether from the donor or commercial sectors, to attract serious funding and other resources, interventions designed to address the widespread lack of access to clean and affordable cooking fuels and proper ventilation will never progress far without reliable monitoring tools. However, lessons can be drawn from the experiences of Shell HEH grantees. These are listed below:

a. Although a useful tool for stove development, the standardized WBT is not an appropriate test for validating fuel savings of stoves in use. Nor is it appropriate as an absolute measure for organizations to rate or value different types stoves. It should be emphasized primarily as a tool for the early stages of stove design rather than a post-design or intervention test. [13] It may also be useful as a quick indicator of whether stoves recently disseminated in households were built as designed. We draw this conclusion because the WBT is not a good predictor fuel consumption in real households for any of the stoves in this study. Consequently, based on our evidence, claims about fuel savings based on controlled (lab) measurements should not be accepted without field verification. Such verification should take the form of both qualitative surveys to assess user satisfaction and rates of use, [14] and quantitative measurements to account for actual fuel consumption.
b. The KPT, on the other hand, although giving more reliable estimates of actual fuel use in households, was difficult for the NGOs to perform because of the logistics requirements of field work, the difficulties of retaining cooperation by the householders, and the complexities of multiple fuels and stoves in many households. It remains the best test, but because of these difficulties may not be feasible in many circumstances.
c. The controlled cooking test (CCT), an intermediate "efficacy" measure with elements of both lab and field conditions might thus be considered when designing SPT protocols. It measures the best possible performance with real cooking in a real house, but controlled such that no other stoves and fuels are used, no fuel is used outside the strict cooking period, and in other ways so that

---

[12] See Appendix IV for a summary of the team responses to evaluations of the SP testing methods.

[13] One of the difficulties that the Shell HEH organizations had with the M&E protocols is that they were introduced after the programs were well underway. The groups had already settled on their designs and were conducting WBTs as an afterthought.

[14] Stove dissemination is not the same as acceptance and use. Many projects are evaluated on the number of stoves handed out or sold. However, there may be little correlation between stove dissemination and the rate of actual use. Obviously, fuel savings, which depend on user behaviour as well as acceptance of the stove, are still less correlated with dissemination rates.

variability and needed sample size is lower than with KPTs. It also avoids the difficulties created by multiple fuels and stoves, although of course this reduces its ability to actually estimate changes in fuel use within populations. In addition, it does not allow comparisons among populations with different cooking habits. As noted, GIRA conducted this test and found it more useful than the WBT in understanding the performance of their stove, which served their purpose although perhaps not that of a donor interested in widescale comparisons or reliable estimates of population fuel savings. Table 11 compares the three tests.

d. Both the KPT and CCT are difficult to do in field conditions – particularly because stove dissemination among ICS groups is coming to be largely market-based, meaning that the population of ICS users from which to draw a sample depends on household purchasing decisions.

Table 11: Outcomes expected from each type of stove performance test discussed in this report

Theoretical issues → Applied issues

Lab → Field		Understand basic principles of stove operation	Test variations in stove design	Test if stove is appropriate for local cuisine	Test if stove is suitable for local users	Test for fuel savings among users
Lab	WBT	++	++	–	–	–
	CCT	±	++	++	++	±
Field	KPT	– –	–	+	+	++

In summary, the WBT is meant as a tool for stove developers to understand the basic principles of stove function and the effect of changing certain parameters of stove design. It serves as a useful way to refine the design of the stove prior to dissemination. However, over the years, the WBT has sometimes become the benchmark test for all stove activity at the expense of field tests. An exception was the Chinese National ICS Program (NISP), which performed such tests with a sample of in-household stoves in each county as part of the verification phase of dissemination (Smith et al., 1993; Sinton et al., 2004). Many claims of actual fuel savings from stove projects are derived from lab-based WBTs and are quite likely inaccurate. There is unfortunately little evidence of a simple relationship between efficiency or fuel consumption in lab-based tests and fuel consumption in real stove use. The experiences of the Shell Foundation's HEH grantees demonstrates this well.

Although tempting to use stove performance tests as an absolute metric of the value of an ICS, stove adoption and use are dependent on many other factors including climate, cultural norms, specific cooking needs etc. Since success of an ICS program is defined by the numbers of stoves in actual usage in communities, rather than simply the number of stoves that are disseminated and built, both performance and acceptance by local communities should be incorporated in evaluation. Such an approach is currently under development for the Shell Foundation's Chinese stove competition, although the weighting of the different components many require adjustment over time.

For large-scale ICS-driven disseminations, therefore, we have the following major recommendations;

1. Consideration be given to relying principally on efficacy measures in the form of controlled cooking tests in typical households. These studies can be done by NGOs after modest training, but short-term assistance in data analysis and interpretation may be needed from outside professional consultants from universities, research institutes, or environmental consulting firms, if the NGO is not interested in developing this capacity on its own.

2. As with IAQ monitoring, however, field surveys of pilot interventions will also be necessary, however, but can focus on assessing[15]
    a. the degree to which the ICSs are actually put to use in the population
    b. the extent to which traditional stoves remain in use (with or without the ICS)
    c. the durability of the ICS in real kitchens
    d. householder suggestions for improvements to the design or construction of the stove to keep maintain good performance, which may only be revealed after extended use (perhaps best revealed by focus group discussions)

3. Large-scale effectiveness monitoring efforts using KPTs only be undertaken with appropriate resources and training and after results from the above procedures indicate that serious population benefits may be occurring. For this reason, stove organizations need technical and financial support to do reliable and rigorous field monitoring. Donors must understand this and be willing to support these activities. Budgets for stove projects must reflect this need by including funds for in-group training, hiring an outside consultant, or subcontracting some or all M&E tasks to a third party firm. Simple WBTs or their equivalents are insufficient for this purpose.

Final summary

In summary, Table 12 compares the three levels of tests -- lab, efficacy, and effectiveness -- for both IAQ and SP measurements, with brief mention of the advantages and disadvantages of each. Here, we recommend that future M&E efforts focus on the middle type, efficacy measures.

Finally, related to both IAQ monitoring and SP tests, credibility of results will always be enhanced by engaging independent organizations to conduct the M&E in order to reduce the chance for bias and the temptation to report partial results that do not fully reflect what was found.

---

[15] We do not elaborate further here partly because of space but also because the report from the teams in conjunction with the Liverpool University group will go into this further with examples from their investigations.

Table 12. Comparison of testing methods

	Lab	*Efficacy*	Effectiveness
SP example	Water Boiling Test	*Controlled Cooking Test*	Kitchen Performance Test
IAQ example	--none developed--	*Test House*	Paired Before-After field measurements*
Utility	Good for design phase stove development	*Reveal best achievable in real households*	Achievable in real household dissemination
Control	All factors except stove type	*Household, fuel, food, cook, season, compliance*	None, or reduce variability minimally by eliminating outliers in population
Advantages	Quick, comparable internationally, relatively simple if equipment available	*Relatively quick, can be done by short-term team and thus use expensive equipment*	Only way to provide indication of changes achieved by an intervention in real populations. Allows measurement of effect over time as well.
Disadvantages	Difficult to translate to other settings, weak connection to real performance in populations	*Cannot be translated between populations, does not measure effect of compliance*	Expensive, time-consuming, and requires relatively sophisticated field research skills

* Other field study designs are also possible, such as cross-sectional, prospective cohort with control group, etc. All share the same characteristics described here to a considerable degree.

Appendices

I. Website Links

SP Testing: The Center for Entrepreneurship in International Health and Development (CEIHD) has posted all of the protocols developed for monitoring stove performance. They can be downloaded at:

http://ceihd.berkeley.edu/heh.stove_perf_eval.htm

Aprovecho Research Center has assisted in the development of the stove monitoring protocols and tested hundreds of stoves with the WBT. More information is available at:

http://www.aprovecho.net/

For a review of stove testing procedures around the world, see the International Review of Household Stove Standards for Performance and Emissions, Testing Methodologies, and Enterprise Innovation (Kaisel, 2005)

http://ceihd.berkeley.edu/Docs/Standards%20review%20final%20Dec%2020.pdf

For a general discussion of stove issues, including monitoring and evaluation of stove performance in the lab and the field, see the stoves discussion list and archive, at:

http://www.bioenergylists.org/

IAQ Monitoring: CEIHD has posted all of the protocols, forms, data analysis spreadsheets, etc., developed for IAQ monitoring in the HEH program

http://ceihd.berkeley.edu/heh.IAPprotocols.htm

II. Standard Methods Summary

Indoor Air Quality (IAQ) Monitoring and Stove Performance (SP) Tests were undertaken by four NGOs – Appropriate Rural Technology Institute, India; Development Alternatives, India; Grupo Interdisciplinario de Tecnología Rural Apropiada, México; HELPS International, Guatemala – with training and support from the University of California Berkeley, for the purpose of documenting the impact of each NGO’s improved cookstove (ICS) on indoor air quality and stove performance in a subsample of households as compared to the traditional cooking stoves they replaced.

Standard IAQ Protocols

Household Selection: A Screening Questionnaire was used by field staff upon their first visit to each household to ensure that the household was suitable for and amenable to the study. If the head of the household agreed to be involved in the study, the field staff administered a Consent Form at that time. The combined Screening Questionnaire/Consent Form is included on the website. As with many rural areas where housing is not standardized, a wide range of different kitchen, and stove configurations are encountered in the rural areas where the NGO were deployed. The intent of the M&E was not to measure the affect of the ICS in all configuration, which would have required a much larger sample size and associated monitoring effort and cost, rather the intent was to measure the most common situation in these areas.

Measurements: Fine particulate matter (PM) and carbon monoxide (CO), the two most important and most studied pollutants associated with biomass combustion smoke, were measured in study households both before and after the introduction of the improved stove (a "Before-After" study design, without controls). Paired PM and CO data were collected in each household, before and after dissemination of the ICS. Sample sizes were chosen to reflect the type of statistical analysis appropriate with this design, with suitable allowance for drop outs, data loss, etc.

PM was measured using the UCB-PE, a version of the UCB Particle Monitor designed for NGOs and containing only a photoelectric detector, not the dual detectors used in the original research version. The UCB-PE was set to record PM concentrations every minute (in units of milligrams PM per cubic meter air, mg/m^3) in its memory (datalogger). It is sensitive to particles of aerodynamic diameter less than approximately 2.5 microns – called fine PM or $PM_{2.5}$ – although not strictly being confined to such sizes (Litton et al., 2004; Edwards et al., 2006). This is the size range thought to be most important for health. The devices were launched using the UCB Monitor Manager© software and most data analysis was done using the UCB Data Browser© software, the latter being upgraded substantially over the course of the project. The monitors were produced and individually calibrated in the Indoor Air Pollution Laboratory at UC Berkeley prior to their use.

CO was measured with the HOBO CO logger (model #H11-001, Onset Computer Company, Bourne, MA, USA: http://onsetcomp.com/), which recorded the CO concentration every minute (in units of parts CO per million parts of air, ppm). The loggers used in this study were purchased new and calibrated at the Indoor Air Pollution Lab at UC-Berkeley using CO standard gas cylinders of 5 and 60 ppm. In the middle of the sampling campaign, before the start of the "After" sampling, a co-location calibration check was run, where the six routinely used loggers were compared to the seventh ("gold standard") logger which was only used for such co-location calibration checks. Changes in sensitivity of the HOBOs found in this way were used to adjust the values during data analysis.

The UCB PM and HOBO CO loggers were placed on the wall of the kitchen for 48 hours (± 2hrs) according to the following criteria:

1. Approximately 100 cm from the edge of the combustion zone (this distance away from the stove approximates the edge of the active cooking area)
2. At a height of 145 cm above the floor (this height relates to the approximate breathing height of a standing woman)
3. At least 150 cm away (horizontally) from doors and windows, where possible

Both devices were co-located (placed next to each other) and placed in a relatively safe location to minimize the risk of interrupting normal household activities or being disturbed or damaged.

Post-Monitoring IAQ Questionnaire. At the end of each 48-hour IAP sampling session, a short Post-Monitoring Questionnaire was administered to the main cook of the household. The questionnaire was designed to document the cooking and other activities that may have affected the indoor air quality in the kitchen during the monitoring period.

The protocols for household selection, the three IAQ measurement methods used, the HOBO CO Calibration Check Protocol, and the Sampling Data Forms can be found at CEIHD's website: http://ceihd.berkeley.edu/heh.IAPprotocols.htm (Indoor Air Pollution Team, 2005).

Standard SP Test Protocols

The controlled stove test (Water Boiling Test – WBT) and field effectiveness measure (Kitchen Performance Test – KPT) are based on earlier version developed by US-based technical support organizations in the mid-80s (VITA, 1985) and elaborated in a well-known technical report on stoves published soon after (Baldwin, 1986). Modifications to the tests derived for this project were arrived at after extensive testing conducted in partnership with the Aprovecho Research Center**,** which has since adopted the standardized WBT in order to analyze hundreds of stoves.[16] In both cases, attempts were initially made to both simplify and standardize the tests with the goal of ensuring comparability and replicability within and across different stove organizations.

The selection of households should be randomized from a suitable typical population. If necessary, communities from which the sample is drawn should also be stratified by factors that may systematically bias fuel consumption such as fuel scarcity, level of poverty, etc. Pure random sampling is difficult however, because households that are selected must also be willing to cooperate with the study, which can be quite intrusive. An additional element of selection bias must be introduced because households must be willing to adopt the improved stove – in some cases this also means that they must be willing to purchase it. Thus, in many instances it may be necessary to stray from a random sample in order to ensure that participants are cooperative and willing to adopt the stove n question.

[16] For more information on Aprovecho's stove work, contact Dean Still at stoves@aprovecho.net.

SP Test Modification: Changes to the earlier tests included some minor procedural changes, the introduction of a "standard cooking pot" and the use of modern, but fairly inexpensive, measuring equipment. In addition, data analysis and calculation tools were incorporated into spreadsheet-based software to ease the analysis and standardize reporting methodologies. Detailed descriptions of the procedures, the changes that have been made relative to earlier versions of tests, and the analysis spreadsheets are available online at http://ceihd.berkeley.edu/heh.stove_perf_eval.htm.

Modifications of protocols in practice

Each group found it necessary to adapt the standard protocols to fit their own circumstances.

ARTI

Village selection: The village must not be very far from the ARTI office and must have reasonable access by road. The village head must give permission and most of the houses in these villages should have the typical housing pattern common in that area.

Household Selection: In the villages selected, a short household survey was conducted using a survey sheet which we (ARTI) developed. The criteria for household selection included family size, presence of children, kitchen architecture, economic status, housing pattern, willingness to buy the ICS, and willingness to participate. These information were entered into an Access database from which a selection was made so that all the houses in a particular area (high or low rainfall) had the same structure and building material, family size was more or less same (nuclear families were selected), presence of young children below the age of 5 years, willingness to buy the stove and otherwise participate, and currently using the same traditional stove and fuel.

The households selected were informed and their consents were taken. WBT tests were conducted in some of these households in each village. Wherever they agreed to make their cookstoves available for 3 days, KPTs were also conducted in these homes. Fewer agreed to a 7-day test.

DA

The project was implemented in three clusters of Bundelkhand region near Jhansi (Radhapur, Niwari and Thona). Households that had elected to buy an improved stove were selected and monitored for baseline data and during the monsoon, winter and summer seasons. The selection of the household was based on an extensive survey conducted initially using a pre-monitoring questionnaire based on criteria of kitchen type, kitchen size and family size. The households were selected from the survey after prior approval.

GIRA

Selection of villages and households was done considering the sample already chosen for a separate health study being conducted locally in which young women and at least one child less than 3 years old were selected for a total of 600 households in 6 villages. For the KPT tests, an initial sub-sample was selected from the 600 households. For the second and third measurements, some of the initial families were not willing or not able to participate in the tests, and the sample was completed with new families.

KPT: The wood used for the test came exclusively from the households' own stocks. For these reasons, in some cases the test did not last for the whole seven days. However, not supplying the families with additional wood allowed measuring the fuel consumption in actual local circumstances. Measurement of LPG consumption was also included, which confirmed that households that use only fuelwood for cooking save more fuel than those using both LPG and fuelwood. The reason is that LPG is expensive and having access to a clean and efficient woodstoves, households started to switch from LPG to fuelwood for some cooking tasks previously conducted on the LPG stove.

WBT. The *Patsari* cookstove, with a 52 cm wide pan adapted for tortilla making, was not designed to handle the 5 liter vessel to boil water indicated by the standard protocol. The problem was that heat was lost by the surface of the *comal* not in contact with the vessel. For these reasons measurements were done at $90^{o}C$ instead of the $94^{o}C$ that is the temperature at which water boils in the Mexican highlands.

Controlled Cooking Tests were also performed for the main cooking task within Mexican rural households: tortilla making. Results from these are available in Masera et al (2005).

III. Feedback from HEH Grantees – IAQ monitoring

Though each NGO faced individual challenges, common difficulties amongst the organizations included:

- Communicating the project's intent to household members, including clearly transmitting evaluation goals and objectives;
- Building the confidence of team members in their monitoring and evaluation ability;
- Building confidence of community members in field team members;
- Maintaining motivation in the field team to promote accurate data collection; and
- Keeping household intrusions to a minimum.

More specific to IAP monitoring, the groups collectively reported a number of challenges. Some challenges were better anticipated – including inaccuracies in household surveys, household drop-outs, and the standard difficulties associated with field logistics and time management– than others, including handling the UCB particle monitor instruments and managing collected data. One NGO characterized the entire process as very difficult for an NGO with no prior IAP monitoring experience. For example, it was challenging to maintain consistent position of the monitors in the kitchen and to manage time efficiently to ensure monitoring periods consisted of 24 hours. Specifically, the UCB provided unique challenges including:

- Equipment malfunctions;
- Management of UCB data quality;
- Lack of detailed protocols for data interpretation;

- Low measurement correlation between UCB and other instruments in co-location experiments.

Throughout the monitoring process, the NGOs shared their challenges with the UC Berkeley team. In response, the UCB team took a number of steps to simplify the air pollution monitoring process (refer to Shell Foundation Progress Report 2005). Additionally, all NGOs provided important suggestions for improving the quality of the monitoring; these included scheduling regular conference calls to discuss problems; providing a detailed UCB manual and finalized versions of all protocols; and hiring an outside contractor to do all the air monitoring to ensure quality results. Despite the challenges of air quality monitoring, most groups have adequate confidence in their data to distribute posters and pamphlets and/or hold focus groups to share their findings with community members.

Suggestions for Future IAP Trainings

In a written survey[17], the NGOs suggested that the future IAP training be restructured to introduce procedures for data sampling, data management and data analysis in one workshop so that the sampling and data management teams could better assess the quality of data obtained. Given this framework, teams could identify problematic homes immediately and re-monitor HH as necessary. Data analysis could also proceed simultaneously with monitoring, thus saving time. They also suggested providing a "suitable period" of training in the handling, calibration, installation and analysis of data from the equipment. A final suggestion was to work towards building capacity of all the field team members to carry out interviews, questionnaires, selection of houses, and handling of all the equipment in order to avoid delays in case of disease, etc., of some member of the field team.

IV. Feedback from HEH Grantees – SP Testing

As part of the process of reviewing the effectiveness of the SPT protocols, a survey was sent to each Shell HEH organization in May 2006.[18] The responses to these surveys revealed several weaknesses in the testing procedures for both the WBT and the KPT. These included:

- Quantity of water in WBT was too large (wasted water, fuel, and time)
- "Hot-start test" from WBT not necessary (wasted fuel and time)
- Provisions needed for non-woody fuels in WBT
- Standard pot not appropriate
- Provisions needed for multiple stoves and fuels in KPT
- KPT is too lengthy (raised by one out of four respondents)

With the exception of the final issue, these issues have all been addressed in the revisions of the WBT and KPT protocols, which were completed in consultation with Aprovecho in May and June of 2006. The final issue is a difficult one – as was mentioned above, assessing real world

[17] A copy of the survey is available on request.
[18] A copy of the survey is available on request.

fuel consumption takes a lot of time and effort. Nevertheless, it is critical. Therefore no changes have been proposed to the KPT procedure regarding sample size or recommended duration of testing.

Additional issues were raised in the surveys that go beyond the procedural such as the utility of standardized SPTs for stove development organizations. While most groups felt that some kind of monitoring and testing is essential the overall sentiment was that this degree of formalized and standardized testing is perhaps too much of a burden on the organizations. One group in particular questioned the value of a standardized and replicable WBT. They raised two issues: first, the WBT procedure only loosely resembles a cooking process and has very little predictive power concerning actual fuel consumption in the field; second, cooking is highly local – relying on a test that is identical for all stoves throughout the world ignores the local specificity of stoves that stove developers must rely on so heavily to create an acceptable stove.

Partly to deal with this last concern, one group (GIRA) undertook on its own initiative a third variety of test, the controlled cooking test (CCT), an “efficacy” measure, which was also promoted by Vita and Baldwin in the 1980s. They found it more useful than the WBT in understanding the performance of their stove and not as difficult to do as the KPT. They have results posted on their website.[19] It, however, cannot be used to compare results across regions because it measures the amount of fuel used to cook a local meal. Nevertheless, if a locally specific CCT were conducted by various stove developers in different regions of the world, a metric such as *percentage reduction in fuel consumption* or *time savings*, both relative to the traditional stove, could be used to compare stove performance.

In summary, other than the one respondent who felt that it was too long, the KPT received less criticism than the WBT. Groups felt that the information collected from households in the field was highly valuable for their organizations and generally worth the effort.

[19] See http://www.gira.org.mx/index.php?option=com_content&task=view&id=75&Itemid=71 (in Spanish).

References

Baldwin S. F. (1986). *Biomass Stoves: Engineering Design, Development, and Dissemination.* Princeton, NJ, Center for Energy and Environmental Studies**:** 287 pp.

Edwards R, Smith KR, Kirby B, Allen T, Litton CD, Hering S. (2006) An Inexpensive dual-chamber particle monitor: Laboratory Characterization, *Journal of the Air and Waste Management Association*, 56: 789-799. http://ehs.sph.berkeley.edu/krsmith/publications/2005%20pubs/Edwards%20et%20al%20%20final.pdf)

Indoor Air Pollution Team, (2006) Indoor Air Pollution Monitoring Protocols, Version 05.1, School of Public Health, University of California, Berkeley, 1992-2005, http://ceihd.berkeley.edu/heh.IAPprotocols.htm

Kaisel D. (2005) International Review of Household Stove Standards for Performance and Emissions, Testing Methodologies, and Enterprise Innovation, Center for Entrepreneurship in International Health and Development, University of California, Berkeley. http://ceihd.berkeley.edu/Docs/Standards%20review%20final%20Dec%2020.pdf

Litton CD, Smith KR, Edwards R, Allen T (2004). Combined optical and ionization measurement techniques for inexpensive characterization of micrometer and submicrometer aerosols, Aerosol Science and Technology, 38:1054–1062. (http://ehs.sph.berkeley.edu/hem/hem/documents/Litton%20et%20al%20Final.pdf)

Masera O., R. Diaz and V. Berrueta (2005) From cookstoves to cooking systems: The integrated program on sustainable household energy use in Mexico. Energy for Sustainable Development 9(1): 25-36, (2005))

Naeher LP, BP Leaderer, KR Smith, (2000) Particulate matter and carbon monoxide in highland Guatemala: Indoor and outdoor levels from traditional and improved wood stoves and gas stoves, Indoor Air, 10:200-205. http://ehs.sph.berkeley.edu/krsmith/publications/Naeher%20et%20al.%202000%20IA.pdf

Pennise D, EnterpriseWorks-Ghana, (April 2006) Indoor air pollution monitoring in Ghana: Third progress report; project reference number: 21341. Submitted to the Shell Foundation.

Riojas-Rodriguez H., P. Romano-Riquer, C. Santos-Burgoa and K. R. Smith (2001). "Household firewood use and the health of children and women of Indian communities in Chiapas, Mexico." Int J Occup Environ Health 7(1): 44-53.

Saksena S, Thompson L, Smith KR, (2003) Database of Household Air Pollution Studies in Developing Countries, Protection of the Human Environment, WHO, Geneva. http://ehs.sph.berkeley.edu/hem/page.asp?id=33

Saatkamp BD, Masera OR, Kammen DM. (2000) Energy and health transitions in development: fuel use, stove technology, and morbidity in Jarácuaro, México. Energy for Sustainable Development, 8(2): 7-16.

Sinton JE, Smith KR, Peabody JW, Liu Y, Zhang X, Edwards R, Gan Q, (2004) An Assessment of Programs to Promote Improved Household Stoves in China, Energy for Sustainable Development 8(3): 33-52.

Smith KR, Gu SH, Huang K, Qiu DX, (1993) "100 Million Improved Stoves in China: How Was It Done?" World Development 21(6): 941-961.

Salud S.D. (2003). Proyecto de Modificación a la NOM-025-SSA1-1993. Mexico City, Mexico.

VITA (1985). Testing the Efficiency of Wood-Burning Cookstoves: Provisional International Standards. Arlington, VA, Volunteers in Technical Assistance**.**.

World Health Organization (2005), Air Quality Guidelines Global Update - 2005, WHO Regional Office for Europe, Copenhagen http://www.euro.who.int/Document/E87950.pdf

Zuk M., Rojas L., Blanco S., Serrano P, Cruz J., Angeles F., Tzintzun G., Armendariz C., Edwards R.D., Johnson M., Riojas-Rodriguez H., Omar Masera O., 2006. The Impact of Improved Wood Burning Stoves on Fine Particulate Matter Concentrations in Rural Mexican Homes, Journal of Exposure Science and Environmental Epidemiology. advance online publication, May 24, 2006; doi:10.1038/sj.jes.7500499