Biomass and Bioenergy 98 (2017) 321–335

Contents lists available at ScienceDirect

Biomass and Bioenergy

journal homepage: http://www.elsevier.com/locate/biombioe

Review

Laboratory protocols for testing of Improved Cooking Stoves (ICSs): A review of state-of-the-art and further developments

Francesco Lombardi [a, *], Fabio Riva [a], Giorgio Bonamini [b], Jacopo Barbieri [a], Emanuela Colombo [a]

[a] Politecnico di Milano, Department of Energy, via Lambruschini 4, Milan, Italy
[b] Politecnico di Milano, Department of Mathematics, Via Bonardi 9, Milan, Italy

ARTICLE INFO

Article history:
Received 19 July 2016
Received in revised form
1 February 2017
Accepted 6 February 2017

Keywords:
Improved Cooking Stove
Cookstove
Laboratory protocols
Cooking system
Testing procedure
User habits

ABSTRACT

Around 2.7 billion people rely on biomass fuelled inefficient devices for cooking and heating. Improved Cooking Stoves are promoted as a means to mitigate the economic, environmental and social implications of this practice. However, their diffusion is hindered by a number of factors, including in particular the lack of agreement on performance evaluation methodologies. Laboratory protocols are designed to give useful indications to cookstoves developers, in order to improve their performance under controlled conditions, while field protocols provide the assessment of real performance of a cookstove in a given context. However, due to high time and finance requirements of the latter, lab results are often used also for stoves selection, also because of a general misunderstanding regarding their correct utilisation. In this work, we provide a review of all lab protocols officially published to date, comparing conceptual and technical aspects. We find that no protocol takes into account all the relevant factors at once. As a result, lab tests carry little information about real field performance, and can be misleading regarding stoves optimisation. Therefore, the analysis reveals the need to define better standards, regarding: (i) repeatability, metrics and statistical analysis of results; (ii) burn sequences calibrated from time to time according to the specific user.

Contents

* Corresponding author. Via Lambruschini 4, 21056 Milan, Italy.
E-mail addresses: francesco.lombardi@polimi.it (F. Lombardi), fabio.riva@polimi.it (F. Riva), giorgio.bonamini@polimi.it (G. Bonamini), jacopo.barbieri@polimi.it (J. Barbieri), emanuela.colombo@polimi.it (E. Colombo).

http://dx.doi.org/10.1016/j.biombioe.2017.02.005
0961-9534/

1. Introduction

The lack of access to clean cooking facilities in developing countries involves 2.7 billion people [1] who rely on inefficient devices fuelled by traditional biomass sources (including wood, animal dung, crop waste and charcoal) for cooking and heating purposes. This entails serious health implications – 4.3 million premature deaths per year due to indoor air pollution [2] – and emission of climate-forcing pollutants [3–6]. In addition, depending on the context, the traditional use of biomass may contribute to the stress on forest resources, although a direct link between deforestation or forest degradation and the use of biomass is not fully assessed in the literature [7–15]. *Improved Cooking Stoves* (ICSs) represent the most commonly promoted technological solution to address this issue. As indicated by Kshirsagar and Kalamkar [16], an ICS is «a cookstove designed using certain scientific principles, to assist better combustion and heat transfer, for improving emissions and efficiency performance» as compared to *traditional cooking devices*, which are identified as very cheap or costless stove models traditionally in use within a certain population [16]. The literature identifies also a third category of stoves, namely *Advanced Biomass Stoves*, which includes manufactured cookstoves characterised by most recent research innovations and features – e.g. a blower injecting air above the fire to improve the combustion efficiency [16,17]. Nevertheless, the success of ICSs is hindered by a number of factors, such as stove stacking (the combined use of traditional and improved technologies for different tasks) [18], competing uses of biomass with other activities (lighting, heating, brick-making), nutritional habits (*viz.* food choices) [19] and lack of agreement about methodologies for performance evaluation [20,21]. This paper focuses on the latter aspect; in fact, a correct evaluation of performance represents a key factor for informing decision-makers and ensuring the selection of technologies that are appropriate for the target context of use [17,22]. In addition, reliable data on ICSs performance are critically needed also as an input for global climate prediction models [23,24]. Performance of ICSs can be evaluated by means of two different approaches: (i) laboratory-based and (ii) field-based testing protocols.

Laboratory-based protocols are traditionally meant for design optimisation, and represent a tool for stove developers to assess changes in performance due to different designs and features. Accordingly, they are performed in a controlled laboratory setting, which should allow avoiding any variability related to external conditions and user's behaviour that would strongly influence the performance of the "cooking system" – the combination of technology, fuel, pot and burn sequence – in a real context of use [6,20,25–31]. Typical parameters evaluated by laboratory protocols include efficiency, specific consumption, CO emissions, PM emissions.

On the other hand, field-based protocols typically consist in prolonged surveys to measure fuel savings among households using an ICS, as compared to a baseline. Hence, field-based protocols can provide reliable insights about the true energy performance of a stove when it is adopted in a real context of use, especially when the assessment on the field is performed by employing the so-called Stove Use Monitors (SUMs), which are low-cost temperature and/or emissions data loggers. The latter can be installed on ICSs, in order to return reliable estimates of their pattern of utilisation [32]. In some cases, field campaigns may include also an assessment of the emission performance of stoves in a real context of use. We provide a few examples: Bailis et al. [33] and Roden et al. [29] assessed PM, CO and CO_2 emissions from cookstoves in Kenya and Honduras, respectively. Johnson et al. [6] measured emissions in the field in Mexico, including also CH_4 and TNMHC (total non-methane hydrocarbons). In general, those studies highlight how emissions measured in the field differ from laboratory-based estimates. Field tests are indeed supposed to follow lab testing for an effective evaluation of stove energy and emissions-related performance among the target population and in a specific context. Moreover, they are usually considered as the most reliable reference by the Clean Development Mechanism (CDM), and certification bodies such as Gold Standard, in the voluntary carbon credit market for carbon offsets programs [24], in order to assess the real savings in terms of greenhouse gases (GHGs) emissions in the framework of climate change mitigation initiatives. However, large sample sizes are often needed to obtain statistically significant results. According to a recent study by L'Orange et al. [34], assuming that a 95% level of significance is required, 160 test replicates per cookstove would be needed to differentiate the performance of Tier 4 and Tier 3 stoves – in terms of CO indoor concentration, as defined by the International Workshop Agreement (IWA 11:2012 [35]) – in the field, while 180 tests would be required to determine CO indoor concentration «sample means within 5% of the population mean» [34]. The entity of the sample sizes requires therefore large financial resources – as compared to a few thousand dollars needed for a certified laboratory evaluation [36] – for sampling equipment, personnel support and logistical expenses over prolonged observation periods [4,24,28,37–39]. Consequently, it is not always feasible to use them as the standard method for field performance evaluation. An alternative solution may be represented by the Uncontrolled Cooking Test (UCT), assessing stove performance in the field when cooking any meal according to local practices and conditions, with the stove being operated by target users. The test is characterised by a lower variability than traditional field tests, potentially allowing for the use of fewer resources to draw statistically significant conclusions. Still, the UCT does not completely solve the problems related to the intrusiveness and the logistical effort required by field studies, and is not currently recognised by carbon offsets programs. Finally, both traditional surveys and uncontrolled cooking tests are not functional for design optimisation, as several complex sources of variability arise in a field context – such as those due to how the users operate the stove, or those related to the presence of unpredictable airflows in a kitchen or in an outdoor setting [34] – preventing a precise assessment of the system performance.

A first attempt to solve the dichotomy between lab and field is represented by the Controlled Cooking Test (CCT), which is still a lab protocol, but prescribes to cook a typical target meal of the target population on the stove, to assess the potential adaptability of the stove to users' cooking habits. This test is supposed to provide a validation of lab stove performance when reproducing a real cooking task. However, according to different studies, the indications provided by a single-meal cooking test cannot be sufficiently representative of the variety of tasks occurring in a field context [24,27,40].

Hence, on the one hand, results from laboratory-based protocols are not reliable in order to understand the performance of a stove in the field [6,28,29,41]; on the other hand, field protocols are too costly to be widely adopted. As a result of this framework, lab protocols – «cheaper and easier to implement» [4,24] as compared to all other methods – are often misused as if their results were reliable indicators of performance in a real context of use. This common practice leads to: (i) errors in technology selection, since poor field performance or inadequacy to the user 's needs entail low adoption rates [21,41,42]; (ii) errors in climate impact estimates, as emission factors measured in the lab significantly differ from those assessed in the field [24,29,43,44]. It seems therefore critical to provide a comprehensive review of all existing lab protocols and a comparison of both conceptual and technical aspects. The present work contributes to highlight strengths and weaknesses of each approach and to identify the key common issues to be addressed in order to develop a reliable, affordable, and effective standard, capable of providing information that is useful for both design optimisation and evaluation of the real performance of cookstove technologies.

The review makes extensive use of both scientific and grey literature concerning testing protocols and ICSs performance evaluation. As a matter of fact, it has been essential to consider a wide range of literature resources, in order to recover information that is not specified in the protocols documents and to draw a comprehensive outlook of the topic. The types of documents considered include: research papers, international standards, conference proceedings, as well as reports and webpages from research centres, international organisations, cooperation agencies and NGOs. The study is organised in two main sections: section 2 is dedicated to the presentation of the historical evolution of laboratory-based protocols; section 3 presents a comprehensive review of all the six lab protocols officially published to date, and recognised by the “Global Alliance for Clean Cookstoves” (GACC), which are mainly discussed in terms of strengths and weaknesses. The analysis also includes conceptual considerations about two further protocols that are currently under development by the ISO Technical Committee 285. Section 4 reports the key findings of the review and the open questions to be addressed by further studies for the development of a new standard. Appendices A, B and C provide the readers with technical details about all the reviewed test protocols, such as procedure, equipment details, performance indicators and metrics formulation.

2. History and evolution of lab protocols

2.1. The first testing procedures

The history of laboratory-based protocols, reassumed in the scheme of Fig. 1, starts in 1980, when the “Intermediate Technology Development Group” (ITDG) – now known as “Practical Action” – made a first attempt to define a procedure for testing a cookstove in

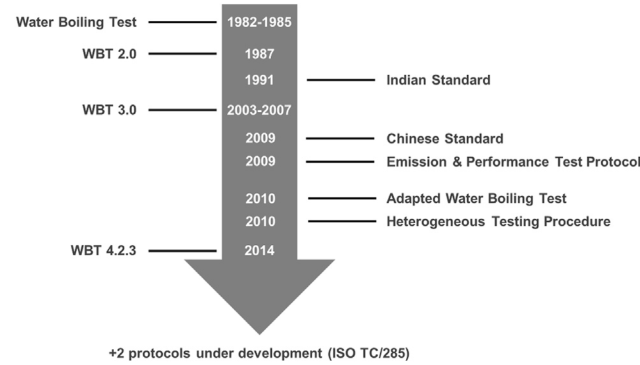

Fig. 1. Summary of protocols evolution over time.

laboratory [45]. Between 1982 and 1985, the “Volunteers in Technical Assistance” (VITA) developed the ideas from ITDG and from the “Eindhoven Woodburning Stove Group” [46] into the first version of the Water Boiling Test (WBT), aimed at measuring how much wood is used to boil water under fixed conditions. In this first version, WBT consisted of two phases, a high power and a low power phase, during which water was brought rapidly to the boil and then simmered for 30 min. Emission testing was not included yet. A first revision and discussion of the VITA's WBT was made by Baldwin's technical report on stoves [47], one of the most widely-cited references for stove developers (it will be referred to as WBT 2.0).

2.2. The Indian and Chinese programmes

The first years of the 1980s (1982–1983) were also characterised by the launch of the first two large scale dissemination programmes of ICSs in India – the “National Programme on Improved Chulhas” (NPIC) – and in China – the “Chinese National Improved Stoves Programme” (CNISP) [16,48,49]. The two countries created their own methods for testing cookstoves, although they were formalised and published only some years later. The first version of the Indian Standard on Solid Biomass Chulha-Specification (often referred to with the acronym BIS, *viz.* Bureau of Indian Standards) is dated 1991 [50]. A revised version, dated 2013, has been cited by Sutar et al. [51], although it is currently available as draft version only [52]. The first version of the Chinese Standard (CS) that can be found in the literature is instead much more recent (2008) [53].

2.3. The WBT 3.0

In 2003, the “Shell Household Energy and Health” project commissioned the University of California-Berkeley to revise the VITA's protocol, which was performed by Dr. Kirk Smith and Rob Bailis in collaboration with researchers from the “Aprovecho Research Centre”. WBT version 3.0 [54] was completed between 2003 and 2007. The most significant variations in the WBT were the introduction of a further phase, namely “Hot-Start”, and the standardisation of pot sizes and water amounts due to consultations with field organisations and analysis of common cooking practices [55]. After the release of these updated version different research teams started to discuss and critique both the testing procedure and some of the calculations [55]. In the same period Berrueta et al. [28] published the first studies showing «little observable association» between WBT results and field performance; thus, different attempts at improving testing protocols were made worldwide.

2.4. The WBT 4.1.2 and the spread of alternative protocols

The group "Engineers in Technical and Humanitarian Opportunities of Service" (ETHOS) promoted a further revision of WBT, coordinated by Dr. Tami Bond of the University of Illinois along with "Partnership for Clean Indoor Air" (PCIA), which was published as version 4.1.2 (2009). This new version includes instructions for emissions measurement and testing of non-woody solid, liquid or gaseous fuels [55]. In 2009, Colorado State University and "Shell Foundation", in collaboration with cookstoves manufacturers Philips and Envirofit, developed their own protocol, called Emission & Performance Test Protocol (EPTP) [56], on the basis of the updated WBT, but aimed at optimising repeatability. In 2010 two more protocols were released by research teams from different continents: the Adapted Water Boiling Test (AWBT) [57] was developed by the "Group for the Environment, Renewable Energy and Solidarity" (GERES) Cambodia as a modification of the WBT 3.0 including some user-centred considerations; the "Sustainable energy Technology and Research" (SeTAR) Centre, based in Johannesburg, also developed a new testing protocol in the framework of the Pro-BEC project on domestic stoves in 2010. It was called the Heterogeneous Testing Procedure (HTP) [58], and differs from WBT in terms of procedure concept, equipment and calculated parameters.

2.5. The ISO process for the definition of a new standard

Finally, in 2012, an ISO-IWA (ISO-International Workshop Agreement) was held in The Hague, gathering more than 90 stakeholders from 23 countries [35]. It was hosted by the GACC and the Partnership for Clean Indoor Air, and chaired by the American National Standards Institute. The workshop provided interim guidance for rating cookstoves on four performance indicators: (i) efficiency, (ii) total emissions, (ii) indoor emissions, and (iv) safety. For each indicator, multiple Tiers of Performance (0–4) were defined, to set a hierarchy in the ICSs technological advancement. The most recent version of WBT 4.2.3 (2014) includes results from the ISO-IWA meeting and Tiers of Performance. Still, criticism of the WBT has increased as further comparative studies against real-cooking performance have been released [59,60], and as researchers raise questions about the rationale of some performance metrics [61]. Different authors [23,61] declare that it is impossible to predict ICS field performance without a user-centred approach, properly accounting for local burn sequences and practices. Based on this trend, the lack of an accepted standard and the spread of alternative protocols represent a major challenge for the success and effective evaluation of stove dissemination programs and carbon-financed projects. For this reason, the ISO Technical Committee 285 is currently working in order to develop new and effective protocols and to solve the issue of the lab-field gap [62].

3. Protocols review

The comparative analysis of testing protocols is performed by setting the Water Boiling Test (WBT) as the benchmark. Other protocols can be in fact easily presented as variations or upgrades of the WBT. The only exceptions are the Indian and Chinese standards, which have been developed independently. For each protocol, four indicators are critically discussed: (i) *real-life relevance* of results, (ii) *repeatability* of the procedure, (iii) *metrics* meaningfulness and rigorousness, and (iv) *statistical significance* of results.

3.1. Water Boiling Test (WBT)

The WBT 4.2.3 is composed of three phases performed in sequence (Fig. 2):

1. *Cold Start*: the fuel in a stove initially at room temperature, is ignited to heat up and bring to the boil a measured quantity of water (2.5 or 5 L, depending on the size of the stove);
2. *Hot Start*: the stove, still hot from the previous phase, is re-ignited with a fresh fuel load to heat up and bring to the boil a measured quantity of fresh water;
3. *Simmering*: the stove is operated to maintain the quantity of water left from the previous phase just below the boiling temperature for 45 min.

"Cold Start" and "Hot Start" are defined *high-power* phases, while simmering *low-power* phase. Following the protocol, a complete assessment of a stove is performed through the evaluation of a number of performance metrics for each of the three phases, even though their general formulation keeps uniform throughout the test, with only slight differences in a few indicators. The key indicators evaluated for all the 3 phases, pivotal for foreseeing the impact of an ICS on the reduction of fuel consumption, health benefits and social aspects, are respectively:

– *Thermal Efficiency* (η), which is calculated as the ratio of the heat absorbed by water and the heat produced by combustion (Table C3);
– *Emission factors (EF)*, that are calculated for three kinds of emissions – PM, CO, CO_2 – as the average mass (grams) of pollutant emitted, normalised per mass (kilogram) of burned fuel, per test phase, per volume (L) of water boiled, per energy (MJ) delivered to the pot, per energy (MJ) released from the fuel, or per unit time (Table C4);
– Time to boil (t_b), which is the time the stove employs to bring to the boil a certain quantity of water in the pot.

The WBT is the most detailed protocol officially published: testing concept, procedure details, emissions equipment, rationale and formulation of metrics are thoroughly exposed, in a simple and clear manner (for details on the equipment and operations, and the performance metrics, refer to Appendices A, B and C). Protocol variations are provided to account for different fuels or different stove types, thus allowing for adaptation to different testing needs. Furthermore, an Excel spreadsheet for the calculation of the test results is downloadable from the GACC website and ready to use. Changes between different updates of the protocol are accounted for in a separate document titled "WBT 4.2.3 Spreadsheet Changes" [63], though there are some discrepancies in the version number between the latest Excel spreadsheet and the mentioned document [64].

Some of WBT critical issues remain unsolved. In particular, the main weakness of the WBT concerns its *real-life relevance*. As a

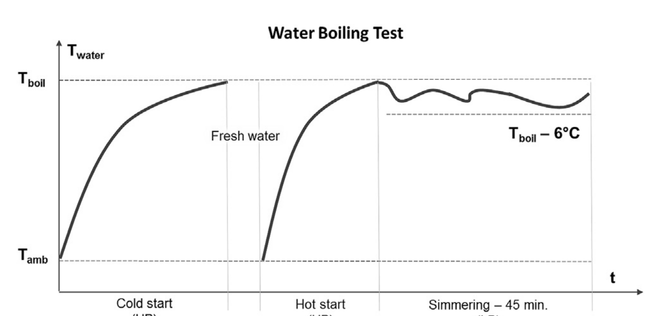

Fig. 2. Scheme of the WBT procedure, adapted from "Water Boiling Test 4.2.3" [55].

matter of fact, the protocol prescribes to test the stove for a fixed combination of burn sequence (high-power and low-power) and pot dimensions. This limits the test relevance to just that particular setting: when any variation of these parameters – which deeply affect the system performance – is introduced, results may significantly differ [21,23,25,26,59,65].

Criticism about WBT concerns also the *repeatability* of the protocol, with a number of researchers claiming that it would need to be reviewed in terms of accuracy [27,66,67]. The WBT is a controlled laboratory test, thus supposed to be characterized by good repeatability and to be effective in comparing different stove designs. Nevertheless, the choice to bring the water to the boil preventing the use of the lid – that is made in order to better approximate a typical cooking task – is not functional to this purpose. As a matter of fact, uncertainties related to temperature reading and vaporisation in the boiling region lead to high variability between test replicates [68]. In order to improve the reliability and replicability of the tests, the latest version of the protocol includes the section "*Changes to Testing Conditions to Improve Repeatability*", which refers to different fuels and pot characteristics. However, eventual changes involving other parameters (pot insulation and maximum water temperature), which deeply affect test variability, are not mentioned. Such problems give evidence of an unsolved conflict between the declared purpose of the WBT (*viz.* to be a design-phase test, not intended to be representative of real-use performance), and a general tendency to use the results form WBT as a significant means to select the most appropriate cooking stove for a given context [24,28,41].

A lot of debate has been made around formulation of *metrics*, primarily on thermal efficiency (Table C3), which is often interpreted as the most immediate and distinctive stove performance parameter. Studies from Bailis et al. [69] highlighted how relying on WBT thermal efficiency outputs, regardless of the relative importance of high and low power cooking tasks among the target population, can lead to misleading interpretations. Furthermore, Zhang et al. [61] and Jetter et al. [20] questioned the scientific meaningfulness of thermal efficiency at simmering. Indeed, this phase is characterised by highly variable steam production, which represents a heat loss in the energy balance but positively contributes to the efficiency value in the actual formulation of thermal efficiency.

Finally, some unsolved issues concerning *statistical significance* of data are worth mentioning. WBT 4.2.3 includes "*Statistic Lessons for Performance Testing*" as Appendix 5[55]. The appendix specifies that the minimum number of test replicates for each model of stove should be three, although it is reported that this number of replicates is not necessarily sufficient to determine a stove performance within a certain confidence interval. Nevertheless, Wang et al. [67] noticed how a great majority of published studies are performed using a number of replicates that is equal or less than three, perhaps due to a misinterpretation of the Appendix message as "only three tests are needed", regardless of variability and confidence interval. Wang et al. investigated this topic using a simplified version of the WBT 3.0 and demonstrated that more than 5 replicates are likely to be required to avoid impractically large 95% confidence intervals and that even more replicates may be required to demonstrate a statistically significant difference in performance between two or more stoves [67].

3.2. Emissions & Performance Test Protocol (EPTP)

The Emissions & Performance Test Protocol (EPTP) was developed as an improvement of the WBT 3.0 (2007). It is proposed as a standardized and replicable test to compare different cookstoves for different cooking applications, with the aim of helping stove designers in the study of both heat transfer and particulate emissions [56]. Similarly to the WBT, the EPTP is composed of three phases (*cold start*, *hot start* and *simmering* – Fig. 3), but it was specifically developed to address the issue of repeatability of tests. For this reason, it includes some peculiar modifications intended to reduce sources of uncertainty: pots are insulated by a floating layer of «closed-cell foam» [56] during high-power phases, and water is heated only up to 90 °C rather than to boiling point.

The intrinsic weakness of the WBT, in correlation to the *real-life relevance* of results, is embodied in the EPTP: the protocol does not propose a solution for the issue of testing the stove only for a fixed sequence (*viz.* high-power and low-power), fixed pot and fixed fuel type, that contributes to make unreliable predictions of the real field performance of a stove.

Since the declared purpose of the EPTP is the improvement of results *repeatability*, all changes in the procedure, as compared to WBT, are motivated by this goal. L'Orange et al. [68] carried out studies comparing results from three EPTP replicates on the Envirofit International B1100 cookstove to standard WBT results on the same stove. They compared the *Dry fuel use* and CO emissions output parameters: no statistically significant difference between results from the two protocols emerged, although the EPTP was found to reduce the coefficient of variation (COV) for «nearly every stove performance metric tested». The results from this study proved the effectiveness of the EPTP in decreasing variability thanks to procedural changes introduced (*viz.* heating to 90 °C and foam insulation) and revealed a reduction in the total time needed to complete the test, as compared to WBT.

Concerning performance *metrics*, the EPTP only slightly differs from the WBT. In particular, the EPTP avoids ambiguities related to thermal efficiency evaluation at low power, as it does not define any efficiency parameter for the simmering phase, relying solely on *Burning Rate* and *Firepower* (details on the equipment requirements, the stove operation and the performance metrics are presented in Appendices A, B and C).

The approach to the issue of *statistical significance* and number of replicates slightly differs from the WBT. Indeed, an exhaustive appendix (*viz. Appendix G: Statistical Considerations* [55]) is included in the protocol, explaining the basic principles of confidence intervals and the influence of replicates number. In this case, the protocol also provides a formula to calculate «the number of test replicates required to determine significant difference in the performance of two test stoves» [56]. However, in the section *Overview of the EPTP*, presenting the test sequence, it is simply affirmed that each stove is tested «three times in a sequence» [56], possibly leading to the same ambiguity discussed for the WBT.

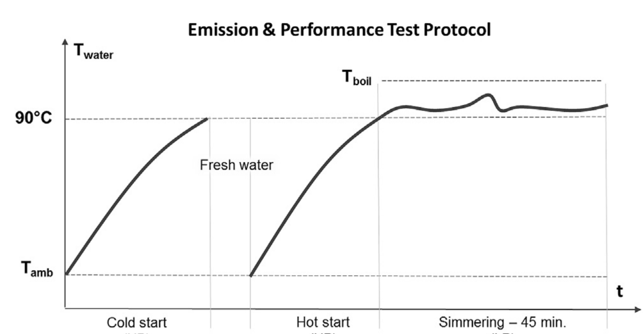

Fig. 3. Scheme of the EPTP procedure.

3.3. Adapted Water Boiling Test (AWBT)

The Adapted Water Boiling Test protocol [57] was proposed as an evolution of the previous "*Protocole de Test d'Ébullition de l'Eau Comparatif*" [70] – also known as Comparative Water Boiling Test –, and as an alternative to WBT 3.0. According to GERES, this protocol was designed «to ease its implementation in developing countries, reduce errors, and take into account local methods of cooking», as well as to be more «accessible to local development agencies and organizations working on the evaluation and dissemination of cookstoves» [57].

Differently from the WBT, the AWBT only consists of two phases which should be performed in a single continuous sequence, namely *Cold-Start High-Power* and *Evaporating High*-Power (Fig. 4). The two phases consist respectively in bringing a not-fixed volume of water to the boil, and consequently keeping the water temperature within a maximum range of 3 °C below the boiling point. Since the fuel amount is definite, and assessed *a priori* in response to local habits, the test ends once the fuel is exhausted and the temperature of the water drops 3 °C below the boiling point. Following the protocol, a complete assessment of a stove is performed through the evaluation of the useful energy provided to the water and the time needed to complete the different test phases. Details on the equipment requirements (*pot*, *fuel*, *insulation*), the stove operation and the performance metrics are presented in Appendices A, B and C, but it is worth noting that no precise prescriptions are given regarding fuel type, since any local fuel is admitted for testing. Similarly, pot material and dimensions, as well as the volume of water, depend on local practices and are not fixed, although a volume of water between 2 and 3 L is suggested.

As regards the real-life relevance *of results*, the AWBT is based on the assumption that taking into account typical cooking practices of the target area and including them in the testing procedure – in terms of fuel amount, pot type, etc. – may help in predicting average field performance. Johnson et al. [23] and Pemberton-Pigott [71] consider this approach a necessary step to improve current testing methodologies. Nevertheless, the AWBT procedure is still based on a fixed burn sequence, instead of following a typical sequence in use by the target community, and thereby not achieving an actual and complete user-centred approach.

The protocol embodies also the WBT's *repeatability* issues, *i.e.* the uncertainties related to temperature readings and vaporisation in the boiling region [68]. Furthermore, doubts may arise from the criterion chosen to determine the boiling point. In fact, the procedure indicates that the boiling point has been reached once the temperature readings have been constant for 10 s, which seems dependent on tester discretion, leading to further variability and errors.

A peculiarity of the AWBT, in terms of performance *metrics*, is that performance evaluation is focused on the concept of "improvement" in relationship to the baseline stove used for comparison, which is an important parameter for fostering the success of stove dissemination programs. Nevertheless, only time and fuel savings are investigated, which are a few parameters compared to other protocols, whilst emissions are not measured at all. Furthermore, this difference in performance is assessed by testing both stoves simultaneously, filled with an identical amount of fuel. Beritault et al. [72] report such practice may lead to errors in case of testing batch-feed charcoal stoves, when the optimum amount of fuel for one stove can be different from the other one.

As regards *statistical significance* of results, the AWBT suggests to perform a minimum of three comparative tests to assess a stove, despite Wang et al. proving how the minimum number of test replicates needed to obtain a statistically significant result cannot be assumed *a priori* equal to three, since it may be much higher depending on the cooking systems being compared [67]. Furthermore, the AWBT only reports that «results are considered statistically valid if the Coefficient of Variation (COV) for the useful energy of each cookstove is below 10%». However, though a low COV indicates a low variability of results, it does not represent a substitute for an inference about the significance level of results, since it does not provide information about their confidence interval.

3.4. Heterogeneous Testing Procedure (HTP)

The Heterogeneous Testing Procedure was developed as a response to the increasing need for the certification of stoves under both Clean Development Mechanism (CDM) and voluntary market projects, calling for the creation of testing protocols capable of simulating real-world use of stoves [58,66]. The main peculiarity of the HTP is to test the stove over a range of three different power levels – that are *high*, *medium* and *low* – and with more than one pot size. This approach is based on the assumption that pollutant emissions and thermal heat transfer mechanism vary with power levels and the size of the pot [66]. Differently from the WBT, the HTP consists of a single continuous phase (Fig. 5), during which the three different power levels are tested by means of *pot swapping*, *i.e.* the subsequent use of three water pots. During the whole test duration, the stove and the pot are placed on an appropriate scale, which tracks in real-time the changes in the weight of the apparatus. It is important to underline that in the case of HTP, differently from the WBT, the pot is placed on the stove with the lid. Once the fire is lit and the stove is fuelled at the highest power, the water temperature and total mass of the apparatus are recorded; the mass of the burned fuel is also measured by temporarily lifting the pot from the stove. As the water temperature reaches 80 °C, the power

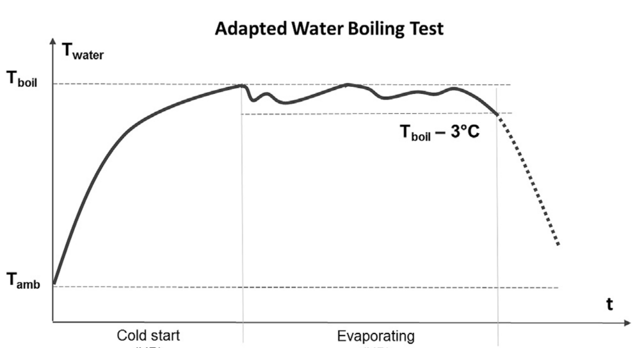

Fig. 4. Scheme of the AWBT procedure.

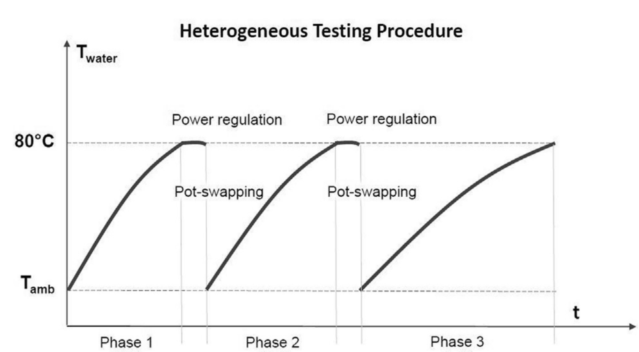

Fig. 5. Scheme of the HTP procedure. The procedure is repeated for any pot/fuel/stove combination tested.

is turned down to the midpoint between the lowest and the highest possible. The pot is replaced with an identical one, filled with fresh water, and the stove is fuelled at medium power until the water temperature reaches 80 °C again. At this stage, the same sequence of operations is repeated for the low-power setting, using a third pot. According to the protocol, a complete assessment of a stove is performed by repeating this procedure for each combination of fuel, pot and stove, at least three times. Performance metrics are missing in the original HTP document [58], and have been partially recovered from papers by Makonese and Pemberton-Pigott in Refs. [66,73]. For further details on the equipment requirements (*pot*, *fuel*, *insulation*), the stove operation and the performance metrics refer to Appendices A, B and C.

As concerns the *real-life relevance* of results, the HTP introduced a number of innovations in the world of ICSs performance evaluation. In particular: (i) testing the stove on three power levels and for different pot sizes, and (ii) providing as a result a set of performance curves covering a range of cooking conditions. These peculiarities are based on the key idea that testing an ICS merely on high-power and simmering tasks – as prescribed by the WBT – cannot provide a complete assessment of the actual stove performance, as stated by previous studies [23,37]. Conversely, the protocol considers performance curves to provide more representative «predictions of emissions and performance when conducting cooking tasks or combinations of tasks» [66]. Still, the HTP does not reproduce the real burn sequence of the target population, which shall be possibly composed of different combinations of high, medium and low-power tasks with different durations. Unfortunately, the literature is missing case studies comparing HTP lab results with field performance, not allowing for a precise assessment of the effectiveness of those innovations.

The HTP also tries to improve *repeatability* avoiding the thermodynamic sources of uncertainty related to temperature reading and vaporization at temperatures close to boiling point. To this end, the protocol prescribes lid insulation and limits the maximum temperature to 80 °C. However, a detailed theoretical study of the specific impact of these expedients on testing variability is missing, as well as an evaluation of possible changes in performance output between this procedure and the traditional "boiling" procedure. Criticalities can be identified in the procedure and in the experimental set-up needed to perform the protocol. Indeed, the HTP adds complexity to the experimental set-up since it requires real-time weighing; furthermore, the pot needs to be lifted every 60 s in order to read the changes in the fuel mass. The impact of these practices on uncertainty should be better evaluated. Also, while the feasibility of operating a stove at three different power levels may be reasonable for devices equipped with a power-level regulation (e.g. some models of ethanol gel stoves), the same practice is arguable when testing most common wood stoves, and criteria to identify power settings should be more precisely discussed [66].

Furthermore, the HTP also avoids all the ambiguities related to the low power *metrics* formulation of the WBT: it does not try to approximate any phase where the water is "simmering" at constant temperature around the boiling point, but rather reproduces a heat transfer procedure where the stove is fuelled at a low-power level to bring the water from ambient temperature to 80 °C. According to Zhang et al. [61], this should be the proper way to evaluate low-power parameters, as *low-power* and *simmering* are not synonyms. As a matter of fact, a simmering task only involves water evaporation, maintaining water temperature constant for a definite period of time. In a low power phase, instead, the stove is operated at the lowest possible power in order to achieve a certain increase of water temperature.

Finally, the time needed to perform a complete testing sequence, *viz.* three times for each fuel/pot/stove combination, is high as compared to other protocols, regardless of *statistical significance* considerations.

3.5. Indian Standard on Solid Biomass Chulha-specification (BIS)

This protocol was designed specifically for the testing of Chulhas Improved Stoves in India. It does not account for stoves using fuels other than wood or multiple pots, while it accounts for emission testing [50]. Sutar et al. [51] and Raman et al. [74] mention a "Revised version" of BIS protocol, dated 2013, also giving a few details about changes in the protocol. Unfortunately, only a draft version of the protocol is available from the Bureau of Indian Standards website [52]. The present analysis will thus refer to the original BIS version dated 1991.

The BIS protocol is based on a different concept than WBT. It does not try, in fact, to approximate a real burn sequence (boiling and simmering), but rather aims at studying an ideal heat transfer process. Fuel amount is calculated *a priori* as a function of the stove firepower and multiple lidded pots are used in a sequence until fuel exhaustion (*pot swapping*, conceptually is the same procedure adopted by the HTP protocol – Fig. 6).

Moreover, the BIS protocol does not repeat the process for different power levels. If testing is contextual to the Indian region, this choice may have a limited impact on the *real-life relevance* of results, as Indian cooking cycles are dominated by high power phases [59]. Conversely, completely avoiding low power testing can lead to very misleading interpretations of stove performance when real burn sequences are likely to require multiple power levels [28]. Furthermore, attention should be paid on fuel prescriptions: although the choice to use completely dried fuel for testing (Table A1) is consistent with the test rationale of reproducing an ideal heat transfer, it may lead to results that are very unrepresentative of typical field usage, since fuel moisture content has been proved to highly influence stove performance [31,68].

Rigid procedure requirements are provided to avoid tester discretion and to improve *repeatability* (see Table A1). In particular, the fixed amount of fuel is highlighted by Arora et al. [75] as an important factor reducing variability as compared to WBT. However, doubts arise from the choice to set a temperature limit to 5 °C below boiling point (Table A1), which is still too close to boiling point to avoid the related thermodynamic sources of uncertainty, as documented by L'Orange et al. [68]. Other protocols performing *pot swapping* and trying to reproduce an ideal heat transfer process, in fact, set the maximum temperature at much lower values (*e.g.* 70 °C). Another possible source of variability comes from the choice to set the *End of Test* as there is «no visible flame» in the stove body: the assessment of this parameter is subjective and variable with stove operator and design.

Finally, no indications are provided regarding the minimum

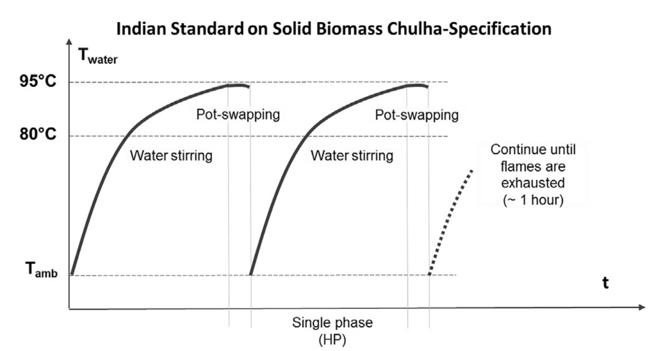

Fig. 6. Scheme of the BIS procedure.

number of replicates.

3.6. Chinese standard

This protocol is designed for the testing of «household stoves»: biomass stoves for cooking or heating, with a power up to 50 kW [53]. The following analysis will take into account «cooking stoves» only.

A remarkable aspect is that the protocol gives a set of requirements an ICS model should satisfy to be approved for being disseminated; such requirements regard the design safety, the appearance, the firepower, the efficiency and the emissions of the stove. The protocol also suggests the stove to be equipped with a chimney, to maintain a temperature of less than 60 °C in normal functioning conditions, without specifying which surface the temperature is referred to, and to ensure a lifetime of at least 3 years under normal operation – even if no method is provided for the estimation of this parameter. Such requirements were developed specifically to support the CNISP and to test Chinese cookstoves. As a consequence, the protocol is not readily adaptable to different stove designs or regions. Furthermore, testing parameters (*viz.* fuel and water amount, pot dimensions) are chosen as a function of the «nominal cooking power», which should be indicated by the manufacturer, meaning that artisanal stoves could not be tested following the same procedure. The amount of fuel derived on the basis of the Cooking Power determines the duration of the test, as the procedure – made up of a single continuous phase in which boiling and simmering are subsequent – is stopped when the fuel is insufficient to maintain water temperature at 95 °C (Fig. 7).

Although simple and quick to realise, the procedure is based on a fixed sequence, whose *real-life relevance* depends on its adherence to the target population's burn sequence [23].

As regards *repeatability,* the protocol is missing some procedure details (*e.g.* unclear ignition step), leading to possible ambiguities and giving space to tester's discretion, which is a source of variability. In addition, it is lacking in description of reasons behind some methodological choices (*e.g.* use of lids), which also have an impact on repeatability.

Questions arise as well on performance *metrics*, and thermal efficiency in particular: all other protocols evaluate separately thermal efficiencies for different power settings, although the same formulation of the parameter is generally maintained throughout all phases; here, a single efficiency parameter is defined from the start to the end of the test, without distinction between high-power and simmering performance (Table C3). Considering that problems related with the rigorousness of thermal efficiency at simmering have been well documented [61], and that the WBT 4.2.3 itself advises to view this parameter with caution [55], the choice to merge into a single value high power and simmering efficiencies seems arguable. Similar considerations can be made about the cooking power metric (Table C2), based on identical data in the numerator of the equation.

No indications are provided about the minimum number of needed replicates.

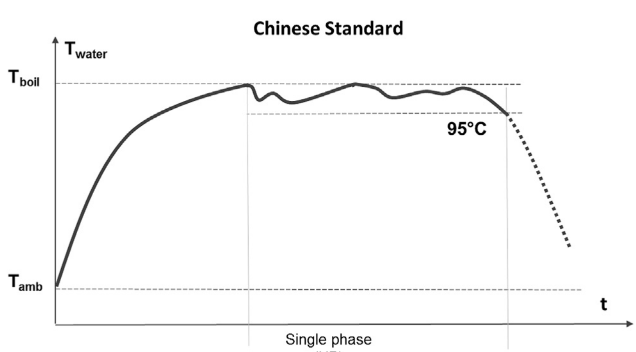

Fig. 7. Scheme of the CS procedure.

3.7. Protocols under development

In recent years new approaches to cookstove testing, namely the Burn Cycle Test (BCT) [23], the Water Heating Test (WHT) [71] and the Firepower Sweep Test (FST) [76], started being proposed as a response to the growing need for protocols capable of predicting average field performance [20]. Details on these protocols are not given within this study, since the BCT remained in the form of a proposal and the other two are still work in progress. However, some of the researchers and experts who contributed to those studies are now actively involved in the process led by the International Organization for Standardization (ISO) for the definition of new testing standards. In fact, the Technical Committee 285 - Working Group 2 on "Clean cookstoves and clean cooking solutions" is developing two new lab protocols (defined as *General laboratory test sequence* and *Contextual laboratory test sequence*), trying to improve the reliability of testing results and their translatability to field contexts [77]. The following overview on the mentioned studies aims therefore at providing a first reference of the concepts that should be formalised under the ISO process.

3.7.1. Burn Cycle Test (BCT) and Firepower Sweep Test (FST)

The BCT was proposed in 2010 in a work by Johnson et al. [23]. It was the first study to highlight that «specific tasks cannot encompass the variety of daily stove use activities, with up to 90% of stove tasks in some regions not involving boiling water». According to the authors, in order to reduce the gap between lab and field performance, greater attention should be paid on the burn sequence, which should be representative of that which occurs during daily cooking activities in homes. A meaningful protocol should therefore test the stove over the same average daily burn sequence that is commonly used by the target households. The authors proposed a procedure that can be summarized in the following steps:

1) the average daily burn sequence is derived from sample households in the field, using gas analysers and $CO_2/(CO+CO_2)$ ratio as a proxy for combustion efficiency;
2) using similar fuel type and moisture content as in the field, 1 kg of wood is split into 5 or 6 equal parts and used to feed the baseline stove (three-stone fire or other stove) in order to reproduce the same distribution of emission rates and combustion efficiencies of the field burn sequence;
3) ICSs are tested over this previously defined lab burn sequence, during the design phase.

The BCT approach was meant to allow for a clear analogy between the lab and the field, as stoves would have been tested over a burn sequence calibrated on the target population. Moreover, the authors suggested that preliminary estimates of GHGs emissions would have been enabled through prediction models, as CO_2-equivalent emissions are linearly linked to combustion efficiency. However, the BCT remained in the form of a proposal, and Michael Johnson eventually joined Bilsback et al. in the development of the Firepower Sweep Test (FST), which can be regarded as a further development of the studies carried out by the BCT authors. In fact, one of the key innovations proposed by the FST is the use of prediction models in order to estimate to what extent the emissions assessed in the lab may vary in real-use conditions [76]. However, differently from the BCT, the protocol is no more based on a representative burn sequence derived from field observations but rather on a generalised sequence trying to encompass a more representative range of power settings than the WBT's. The FST concept was presented at the ETHOS Conference 2016 but is still a work in progress, which should possibly converge into the ISO process.

3.7.2. *Water Heating Test (WHT)*

The Water Heating Test is a revised version of the HTP, which has been already used as a performance evaluation tool in some large scale programs (the most relevant is the Indonesian Clean Stove initiative [27,71], started in 2012 and funded by the World Bank), though it is still a work in progress only partially formalised in 2014 [71]. C. Pemberton-Pigott, from the SeTAR Centre, is the main author of this new approach. The WHT objective is «to evaluate biomass fuel burning cooking appliances in a realistic manner such that their future performance in the hands of a given community is reasonably predicted» [71]. The idea, similarly to the BCT, is to derive a contextual burn sequence that is representative of the average local cooking experience. In this case, the burn sequence is not derived from emission profiles in the field but rather from an evaluation by social scientist of two or more meals or cooking patterns that should be representative of all the different power levels required by a specific target population, *viz.* of the typical burn sequence. Through a sequence of intermediate tests, the typical daily average burn sequence of the specific target user is reproduced, finally becoming the overall sequence over which ICSs will be tested in the lab. The WHT testing procedure reflects the HTP's (ideal heat transfer and pot-swapping) and is kept always identical, while the burn sequence is directly derived from the local context.

As the ISO Process is still a work in progress, conclusions are postponed to the official publication of the new protocols. However, it can be already noted that the common idea of these two approaches is to widen the range of conditions under which the stove is evaluated and to include considerations about the final context of use.

4. Discussion

The comparative analysis of all existing lab protocols carried out in the previous section was meant to provide a precise overview of each test, highlighting its purpose, procedural concept, and strengths and weaknesses. For an overall comparison of the different features considered by the analysed protocols, refer to Tables A1, A2, and B1. The analysis allows in particular for the identification of some critical issues – summarised in Fig. 8 – that are common to all protocols and that represent key points to be addressed by further studies in order to develop new and effective standards.

The first common issue to be highlighted is that none of the current laboratory-based protocols can provide results that are representative of average stove performance in a real context of use. In fact, real-use performance are not merely concerned with the stove, but are rather contingent on different factors, namely: stove design, pot type, fuel and moisture content, burn sequence [6,20,25–31]. All those factors together should be treated as an integrated thermal system [25] or "cooking system" [26], with the last three being strongly dependent on the local context. If the dependence of results from pot shape/volume might be avoided by means of changes in the formulation of metrics, fuel type and burn sequence should still be studied in relationship to the real context of use, and integrated into testing, in order to have a chance of predicting average field performance in the lab [59]. Instead, as emerging from the present analysis, most of current testing protocols are performed fixing those factors – this is for the sake of repeatability and comparison of results between different laboratories –, resulting in performance ratings that are only valid for that particular setting, and that are untranslatable to performance in the field. Consequently, a stove that is highly rated by one of current lab test might be poorly performing under different circumstances. Therefore, if the purpose of the tester is to assess average field performance of a stove in a specific local context, the utilisation of any of the current lab testing protocols may lead to inappropriate estimations. Indeed, it would be necessary to look at approaches under development considering the whole "cooking system" and reproducing it into the lab. Nevertheless, it is important to underline that the results of such approaches may still differ from real-life performance. As a matter of fact, factors such as the user ability to operate the stove, or the size, shape and airflow of the room where the stove is used, have a strong influence on performance and cannot be taken into account in a laboratory test [26,34]. To this regard, a pilot assessment in the field may still be needed.

An important consequence of these considerations is that the Tier of Performance of a given stove, assigned based on the results from WBT or similar protocols, may result in a not reliable performance indicator for technology selection. This concept should be stressed as, although almost all of current lab protocols underline that their usefulness is limited to the *design evaluation* of a stove, and that results should not be intended as real performance indicators, there is still a great misunderstanding about their role. In fact, lab tests results are very often used as the only performance indicators for stoves selection. In order to push technological development in a more user-oriented direction, alternative Tiers of Performance may be defined based on testing approaches considering the whole "cooking system". In this case, the Tiers would allow for a comparison between different ICS models that is relevant in correlation to the target population on which the "cooking system" is calibrated.

The second common issue regards the use of current protocols to determine the effect of design alterations on performance or to identify the best stove designs, which is claimed as their main role

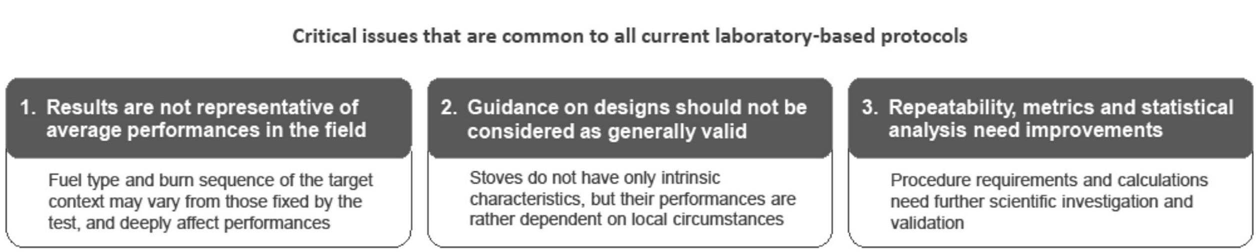

Fig. 8. Summary of key common issues of the laboratory-based protocols considered in the review.

and purpose. Actually, *ceteris paribus*, a given stove will perform better or worse than another one while varying its design. However, a design which is best for a given, fixed combination of all the factors, could not at all be optimised when a variation of one of the other parameters which compose the overall "cooking system" is introduced. For example, as proven by Bhattacharya et al. [31] and L'Orange et al. [68], CO emissions are strongly influenced by fuel moisture content. Therefore, testing one stove design for a fixed moisture content may result in a certain performance, which however will not be representative of the stove performance for a higher or lower moisture content. To avoid such problems, some protocols, as the WBT itself, suggest using typical local fuel for testing. Nevertheless, they still rely on classifying a stove design on a "Performance Tier" as if the results were inherent to that particular design, independently from the fuel, the pot and the burn sequence, leading to misinterpretations and failures in technology selection. A large part of tests published on the GACC Clean Cooking Catalog [78] are in fact missing wood type and moisture content details, yet Tiers ratings are evaluated in any case. The key concept to be highlighted is that no design can be considered as generally valid, since stoves do not have only intrinsic characteristics, but their performance are rather dependent on local circumstances, particularly as regards fuel consumption and emissions [23]. «There cannot be a single universally efficient cooking stove» [25].

In addition to the above considerations, further reflections should concern the repeatability of results, the formulation of performance metrics and the statistical analysis of data. Traditional boiling procedures are in fact characterised by large variability in test results [68], which other protocols try to address focusing on thermodynamic sources of uncertainty (*viz.* boiling regime and evaporation), and in particular limiting the temperature to a given threshold and insulating the pot. However, further theoretical and experimental studies would be needed to evaluate the effectiveness of those changes in improving the repeatability of test results and to define an optimal procedure. The mentioned changes are also often criticised as they might lead to different performance than those evaluated by means of traditional boiling procedures, yet differences might be easily avoided by adopting a more rigorous reference for the definition of typical performance parameters. Finally, statistical data analysis should be structurally integrated into the protocol procedure, as well as clear recommendations regarding the number of replicates needed and how to report tests results, since neglecting the epistemic statistic uncertainties originating from lab tests might lead to misinterpreted evaluations of ICSs' performance [79]. Recent studies suggest that a combination of different statistical approaches – possibility theory and t-student approach – might be adopted depending on the number of replicates: Riva et al. [79] suggest relying on possibility theory when imprecise knowledge and epistemic uncertainty related to small sample sizes arise – *viz.* fewer than five replicates at least, according to Wang et al. [67]. Nevertheless, they suggest further studies would be needed to clearly define a threshold number of replicates below or above which it may be preferable to use possibility theory or the traditional t-student approach, respectively.

5. Conclusions

The findings of this work should contribute to avoid the present confusion about the role of lab tests and the misinterpretation of their results as performance indicators for stoves selection. In addition, the analysis allows for the identification of areas for further research to be addressed within the framework of the ISO process for the development of new standards. To this end, we provide a possible research path for the definition of a standard:

(i) performing theoretical and experimental studies to clearly identify the sources of variability in a testing procedure, the external parameters on which those sources are dependent and the most effective method to control them;
(ii) based on the cited critical studies in the literature, defining meaningful and rigorous metrics, which should be independent from pot dimensions and water volume;
(iii) integrating statistical analysis of data when publishing testing results, avoiding the publication of non-significant ones. The analysis should take into account both possibility theory and *t*-test depending on the number of replicates;
(iv) ensuring the real-life relevance of results through the integration with different burn sequences and fuels, calibrated on the target user. Research should focus on a cost-effective procedure to realise this integration;
(v) defining a new rating system, to allow for a comparison of different stove designs that is real-life relevant in correlation to the specific target user;

It is also worth noting that, as a result of this path, the affordability of the standard would be achieved as well. Indeed, a high degree of repeatability ensures that a reasonable number of replicates is sufficient to draw statistically significant conclusions. In addition, a cost-effective testing procedure would reduce the time needed to perform a single test replicate. However, field testing would be still needed to take into account user variability and boundary conditions in a room or in an outdoor setting.

Appendix A. Requirements on procedure and equipment

Table A1
Procedure and equipment details for all laboratory protocols considered.

Parameters for comparison of testing procedures	WBT 4.2.3 (2014)	EPTP (2009)	AWBT (2010)	HTP (2010)	Indian BIS (1991)	Chinese Std. (2008)
n° phases	3 (2 HP + 1 LP)	3 (2 HP + 1 LP)	2 (1 HP + 1 LP)	3 (1 HP + 1 MP + 1 LP)	1 (HP)	2 (1 HP + 1 LP)
Pot Insulation	no	closed-cell foam (HP only)	no	lid	lid	lid (HP only)
Max value of water temperature	Boiling temperature	90 °C	Boiling temperature	80 °C	boiling temperature minus 5 °C	Boiling temperature
Water quantity	5 L	4-6 L	3-5 L	80% of the pot's volume	2-18 kg	(depends on power)

Table A1 (*continued*)

Parameters for comparison of testing procedures	WBT 4.2.3 (2014)	EPTP (2009)	AWBT (2010)	HTP (2010)	Indian BIS (1991)	Chinese Std. (2008)
Temperature of the water at beginning of test	Ambient temperature	4-30 °C	–	–	23 ± 2 °C	–
Ignition method	(depends on local habits) start test «after the fire has caught»	(depends on manufacturer's indications; if not present, kindling materials are suggested based on fuel type) start test before ignition	(depends on local habits) start test when kindling is exhausted	(depends on local habits)	use kerosene as kindling, start test after 30 s	(left to tester's discretion) start test «when the fuel starts to burn»
Phase duration (when time dependent)	45 min (LP)	45 min (LP)	–	–	>60 min.	–
Minimum n° pots	1	1	–	3	2	1
Pot volume	7 L (volume)	–	–	6.4 L or 3 L	(depends on power)	(depends on power)
Fuel cross-sectional dimensions	1.5 × 1.5 cm^2	1.5 × 1.5 cm^2	–	–	3 × 3 cm^2	–
Length of fuel pieces	–	–	–	–	half the diameter/ length of combustion chamber	–
Moisture content of fuel	6.5% or 10%	4-10%	15% (for fuelwood)5% (for charcoal)	–	0%	–
Pre-weighed bundles of fuel	5 kg	5 kg	(depends on local habits)	–	(depends on power)	(depends on power)
Gaseous emission equipment	(different options are presented with pros and cons)	NDIR (at least one replicate), electrochemical	–	NDIR, electrochemical	(different options are allowed)	–
Particulate matter equipment	(different options are presented with pros and cons)	gravimetric (at least one replicate), optical	–	gravimetric	gravimetric	–

WBT = Water Boiling Test; EPTP = Emissions & Performance Test Protocol; AWBT = Adapted Water Boiling Test; HTP = Heterogeneous Testing Procedure; Indian BIS = Indian Standard on Solid Biomass Chulha-Specification; HP = High Power; LP = Low Power; MP = Medium Power.

Appendix B. Energy and emission indicators

Table B1
Indicators of energy performance considered by different protocols.

Energy indicators	WBT 4.2.3 (2014)	EPTP (2009)	AWBT (2010)	HTP (2010)	Indian BIS (1991)	Chinese Std. (2008)
Fuel Consumption	✓ (Specific fuel consumption, Burning rate)	✓ (Burning rate)	✓ (Potential fuel differences)	✓ (Burn rate)*	–	–
Power	✓ (Fire power)	✓ (Overall fire power, Useful fire power)	–	✓ (Fire power)*	✓ (Burning capacity rate, Power output rating)	✓ (Cooking power)
Thermal efficiency	✓	✓	–	✓	✓	✓
Useful energy	–	–	✓	–	–	–
Turndown Ratio	✓	–	–	–	–	–
Time to boil	✓	✓**	✓	–	–	–
Total time of test	–	–	✓	–	–	–

*Not specified in the protocol but mentioned from the same authors in Makonese, Tafadzwa, et al. “Performance evaluation and emission characterisation of three kerosene stoves using a Heterogeneous Stove Testing Protocol (HTP).” Energy for Sustainable Development 16.3 (2012): 344–351.
**In this case “time to boil” refers to time needed to heat the water from its starting temperature to 90 °C.
✓ = indicator considered in the protocol.
() = the brackets include the name of the related energy indicator as defined in the original protocol.

Table B2
Indicators of emissions performance considered by different protocols.

Emission indicators	WBT 4.2.3 (2014)	EPTP (2009)	AWBT (2010)	HTP (2010)	Indian BIS (1991)	Chinese Std. (2008)
Emission Factor	✓ (CO, CO_2, PM)	–	–	✓ (CO, CO2)	–	–
Pollutant Mass produced	✓ (CO, CO_2, PM)	✓ (CO, PM)	–	–	✓ (PM)	–
Emission per Water Boiled	✓ (CO, CO_2, PM)	–	–	–	–	–
CO/CO2 ratio	–	–	–	–	✓	–

*Not specified in the protocol but mentioned from the same authors in Makonese, Tafadzwa, et al. "Performance evaluation and emission characterisation of three kerosene stoves using a Heterogeneous Stove Testing Protocol (HTP)." Energy for Sustainable Development 16.3 (2012): 344–351.
✓ = indicator considered in the protocol.
() = the brackets include the name of the related energy indicator as defined in the protocol.

Appendix C. Metrics for the measurement of energy and emission indicators

Performance metrics are usually calculated separately for each phase, or power setting, composing the procedure. However, their general formulation keeps uniform throughout a test, with only slight variations occurring in a few cases. Accordingly, metrics formulations are here reported only once, eventually highlighting remarkable exceptions. Furthermore, metrics formulation is reported in the same form as in the protocol, and terminology discrepancies may thus emerge between identical metrics.

Table C1
Metrics referred to fuel consumption for different protocols.

Energy indicator: fuel consumption		
WBT (v. 4.2.3)	$SC = \frac{f_d}{m_{w,f}} \quad [kg\ kg^{-1}]$	$Specific\ fuel\ Consumption = \frac{equivalent\ dry\ fuel\ consumed}{effective\ mass\ of\ water\ boiled}$
EPTP	$R_b = \frac{f_d}{t} \quad [kg\ s^{-1}]$	$Burning\ Rate = \frac{dry\ fuel\ consumed}{test\ duration}$
AWBT	$PFD\% = 100 \cdot \frac{UE_{ICS} - UE_{traditional}}{UE_{ISC}} \quad [-]$	$Potential\ fuel\ differences = \frac{Usefuel\ energy\ of\ the\ improved\ stove - Usefuel\ energy\ of\ the\ traditional\ stove}{Usefuel\ energy\ of\ the\ improved\ stove}$

Table C2
Metrics referred to power for different protocols.

ENERGY INDICATOR: Power		
WBT (v. 4.2.3)	$FP = \frac{f_d \cdot LHV_{wood,dry}}{t} \quad [kW]$	$FirePower = \frac{fuel\ energy\ consumed\ to\ boil\ or\ simmer\ water}{time\ to\ boil\ or\ simmer}$
Indian BIS	$BCR = \frac{2(m_{st.+w.,i} - m_{st.+w.,f}) \cdot LHV_{wood}}{3600} \quad [kW] P_0 = BCR \cdot \eta \quad [kW]$	$Burn\ Capacity\ Rate = \frac{2 \cdot \left(\Delta mass_{stove+wood,\ \frac{1}{2}\ hour}\right) \cdot wood\ calorific\ value}{1\ hour}$ $Power\ output\ rating = Burn\ Capacity\ Rate \cdot thermal\ efficiency$
Chinese Standard	$P_C = \frac{m_{w,i} c_{p,w} (T_b - T_i) + (m_{w,i} - m_{w,f}) h_{lv}}{t}** \quad [kW]$	$Cooking\ Power = \frac{fuel\ energy\ consumed\ to\ boil\ or\ simmer\ water}{time\ to\ boil\ or\ simmer}$
EPTP	$FP_o = \frac{f_d \cdot LHV_{wood,dry}}{t} \quad [kW] FP_u = FP_o \cdot \eta \quad [kW]$	$FirePower\ (overall) = average\ power\ from\ fuel\ combustion$ $FirePower\ (useful) = average\ power\ transferred\ to\ the\ pot$

** The metric is calculated jointly for the high power and low power phases. The mass of water vaporised includes, therefore, also the mass lost during the simmering phase.

Table C3
Metrics referred to thermal efficiency for different protocols.

	Energy indicator: thermal efficiency	
WBT (v. 4.2.3)	$\eta = \frac{m_{w,i} c_{p,w} (T_b - T_i) + m_{eva} h_{lv}}{f_d \cdot LHV_{wood,dry}} *$	$\frac{energy\ to\ bring\ initial\ water\ to\ boil + energy\ due\ to\ evaporation}{energy\ of\ equivalent\ dry\ fuel}$
Indian BIS	$\eta = \frac{(n-1)(m_{pot+lid} c_{p,Al} + m_{w,i} c_{p,w})(T_{limit} - T_i) + (m_{pot+lid} c_{p,Al} + m_{w,i} c_{p,w})(T_f - T_i)}{m_{wood} \cdot LHV_{wood} + m_{kindling} \cdot LHV_{kindling}}$	$\frac{en.\ to\ heat\ [(n-1)(pots + lid + w.) up\ to\ T_{limit} + last\ (p. + l. + w.)\ up\ to\ T_f]}{energy\ of\ fuel + energy\ of\ kindling}$
Chinese Standard	$\eta = \frac{m_{w,i} c_{p,w} (T_b - T_i) + (m_{w,i} - m_{w,f}) h_{lv}}{m_{wood} \cdot LHV_{wood} + m_{kindling} \cdot LHV_{kindling}} ***$	$\frac{energy\ to\ bring\ initial\ water\ to\ boil + energy\ due\ to\ evaporation}{energy\ of\ fuel + energy\ of\ kindling}$
EPTP	$\eta = \frac{m_{w,i} c_{p,w} (T_{limit} - T_i) + m_{eva} h_{lv}}{f_d \cdot LHV_{wood,dry}}$	$\frac{energy\ to\ bring\ initial\ water\ to\ 90^\circ C + energy\ due\ to\ evaporation}{energy\ of\ equivalent\ dry\ fuel}$
HTP	$\eta = \frac{(m_{w,i} c_{p,w})(T_{limit} - T_i) + m_{eva} h_{lv}}{f_d \cdot LHV_{wood,dry}}$	$\frac{energy\ to\ bring\ initial\ water\ to\ 80^\circ C + energy\ due\ to\ evaporation}{energy\ of\ fuel}$

*In the simmering phase, the mass of water simmered is not the initial mass of water, but the average of the initial and final masses of water in the pot.
** In the protocol, it is not indicated whether the mass of water in the formula refers to the initial or final mass. However, since the procedure prescribes to weigh the pot only at the beginning of the test, the authors concluded that this should be the initial mass of water.
*** The metric is calculated jointly for the high power and low power phases. The mass of water vaporised includes, therefore, also the mass lost during the simmering phase.

Table C4
Metrics referred to other energy performance indicators for different protocols.

	Energy indicator: useful energy	Energy indicator: turndown ratio
WBT (v. 4.2.3)		$TDR = \frac{FP_{cold-start}}{FP_{simmering}}$ $\frac{high\ firepower}{low\ firepower}$
AWBT	$UE = m_{w,i} c_{p,w} (T_b - T_i) + m_{eva} h_{lv}$ $[kJ]$ $en.\ to\ bring\ initial\ water\ to\ boil + en.\ due\ to\ evap.$	

Table C5
Metrics referred to emission factors for different protocols for different protocols.

Emission indicator: emission factor	
WBT 4.2.3	$EF_{CO} = \frac{[CO_{test}] - [CO_{bk}]}{[C]} \cdot \frac{28}{12} \cdot fuelFracC \cdot 1000$ $[g\ kg^{-1}]$
	$EF_{CO_2} = \frac{[CO_{2,test}] - [CO_{2,bk}]}{[C]} \cdot \frac{44}{12} \cdot fuelFracC \cdot 1000$ $[g\ kg^{-1}]$
	$EF_{PM} = \frac{PM_{test} - PM_{bk}}{C} \cdot fuelFracC \cdot 10^{-3}$ $[g\ kg^{-1}]$
HTP	$EF = \frac{[Pollutant]}{[C]} \cdot fuelFracC$ $[g\ kg^{-1}]$

Table C6
Metrics referred to the total mass of a specific pollutant produced for different protocols.

Emission indicator: pollutant mass produced	
WBT 4.2.3	$m_{pollutant} = dryfuel \cdot EF_{pollutant}$ $[g]$
EPTP	$m_{CO} = \Delta t \cdot \sum_{i=0}^{n-1} \dot{m}_{CO,i} = \Delta t \cdot \sum_{i=0}^{n-1} \left(\dot{V} \cdot \frac{[CO]_i}{10^6} \cdot \frac{p_{e,i}}{R_{CO,i} T_{e,i}} \right)$ $[g]$
	$m_{PM} = m_{filter,f} - m_{filter,i} - \dot{m}_{PM,bk} \cdot t$ $[g]$
Indian BIS	$TSP = Total\ Suspended\ PM = \frac{(m_{filter,f} - m_{filter,i}) \cdot 10^6}{\dot{V}_{air} \bullet 60}$ $[mg\ m^{-3}]$

Table C7
Metrics referred to emissions per water boiled for different protocols.

Emission indicator: emission per water boiled	
WBT 4.2.3	$E_{pollutant} = \frac{m_{pollutant}}{m_{w,f}} \cdot 1000$ $[g\ dm^{-3}]$

Nomenclature

BCR Burn Capacity Rate $[g\ kg^{-1}]$
C Total carbon in exhaust $[g\ m^{-3}]$
[C] Average exhaust carbon concentration $[mg\ kg^{-1}]$
$[CO_{test}]$ Average exhaust CO concentration $[mg\ kg^{-1}]$
$[CO_{bk}]$ Average background CO concentration $[mg\ kg^{-1}]$
$[CO_{2,test}]$ Average exhaust CO2 concentration $[mg\ kg^{-1}]$
$[CO_{2,bk}]$ Average background CO2 concentration $[mg\ kg^{-1}]$
$c_{p,Al}$ Specific heat capacity of Aluminium $[kJ\ kg^{-1} K^{-1}]$
$c_{p,w}$ Specific heat capacity of water $[kJ\ kg^{-1} K^{-1}]$
$E_{pollutant}$ Emission of a specific pollutant per water boiled $[g\ dm^{-1}]$
EF_x Emission Factor referred to pollutant x $[g\ kg^{-1}]$
f_d Equivalent dry fuel consumed $[kg]$
FP Firepower $[kW]$
FP_o Overall Firepower $[kW]$
FP_u Useful Firepower $[kW]$
$fuelFracC$ Fraction of Carbon in the fuel $[g\ g^{-1}]$
h_{lv} Specific enthalpy of evaporation of water $[kJ\ kg^{-1}]$
$LHV_{kindling}$ Lower Heating Value of the kindling used to light the fire (as received) $[MJ\ kg^{-1}]$
LHV_{wood} Lower Heating Value of the fuelwood (as received) $[MJ\ kg^{-1}]$
$LHV_{wood,dry}$ Lower Heating Value of the fuelwood (dry wood) $[MJ\ kg^{-1}]$
$\dot{m}_{CO}$ Instantaneous mass flow rate of carbon monoxide $[g\ s^{-1}]$
m_{eva} Total mass of water evaporated during a test phase $[kg]$
$m_{filter,f}$ Mass of particulate filter at the end of a test phase $[g]$
$m_{filter,i}$ Mass of particulate filter at the beginning of a test phase $[g]$
$m_{kindling}$ Mass of kindling used to light the fire $[kg]$
$m_{w,i}$ Mass of water in the pot at the beginning of a test phase $[kg]$
$m_{w,f}$ Mass of water in the pot at the end of a test phase $[kg]$
m_{wood} Mass of wood $[kg]$
$\dot{m}_{PM,bk}$ Average background particulate collection rate $[g\ s^{-1}]$
$m_{pollutant}$ Total mass of a pollutant emitted during a test phase $[g]$
$m_{pot+lid}$ Mass of the pot, including lid $[kg]$
$m_{st.+w.,i}$ Mass of the stove, including wood at the beginning of the test $[kg]$
$m_{st.+w.,f}$ Mass of the stove, including wood at the end of the test $[kg]$
n Number of pots utilised $[-]$
P_C Cooking Power $[kW]$
p_e Instantaneous pressure at exhaust sampling location $[Pa]$
$PFD\%$ Potential fuel differences $[\%]$
PM_{test} Average exhaust mass concentration of particulate matter $[\mu g\ m^{-3}]$
PM_{bk} Average background mass concentration of particulate matter $[\mu g\ m^{-3}]$
P_O Power Output Rating $[kW]$
$[Pollutant]$ Average exhaust pollutant concentration $[mg\ kg^{-1}]$
R_b Burning rate $[kg\ s^{-1}]$
SC Specific fuel consumption $[kg\ kg^{-1}]$
t Duration of a test phase $[s]$
t_b Time to boil $[s]$
T_b Local boiling temperature $[^\circ C]$
TDR Turn Down Ratio $[-]$
T_e Instantaneous temperature at exhaust sampling location $[K]$
T_i Temperature of water in the pot at the beginning of a test phase $[^\circ C]$
T_f Final temperature of water at the end of a test phase $[^\circ C]$

T_{limit} Threshold temperature for a test phase [°C]
TSP Total Suspended Particulate Matter [$mg\ m^{-3}$]
$\dot{V}$ Volumetric flow rate of the emissions collection hood [$m^3 s^{-1}$]
$\dot{V}_{air}$ Volumetric flow rate of ambient air [$dm^3 s^{-1}$]
UE Useful Energy [kJ]
Δt Time between sample points [s]
η Thermal Efficiency [–]

Acronyms – Subscripts
AWBT Adapted Water Boiling Test
BCT Burn Cycle Test
BIS Bureau of Indian Standards
CCT Controlled Cooking Test
CDM Clean Development Mechanism
CNISP Chinese National Improved Stoves Programme
CO Carbon Monoxide
CO_2 Carbon Dioxide
COV Coefficient of Variation
CS Chinese Standard
EPTP Emissions & Performance Test Protocol
ETHOS Engineers in Technical and Humanitarian Opportunities of Service
FST Firepower Sweep Test
GACC Global Alliance for Clean Cookstoves
GERES Group for the Environment, Renewable Energy and Solidarity
GHG Green House Gas
HP High Power
HTP Heterogeneous Testing Procedure
ICS Improved Cooking Stove
ISO International Organization for Standardization
ITDG Intermediate Technology Development Group
LP Low Power
NGO Non-Governmental Organisation
NPIC National Programme on Improved Chulhas
PCIA Partnership for Clean Indoor Air
PM Particulate Matter
SeTAR Sustainable energy Technology and Research
TNMHC Total non-methane hydrocarbons
VITA Volunteers in Technical Assistance
WBT Water Boiling Test
WHT Water Heating Test

References

[1] IEA, World Energy Outlook 2015-Biomass Database, 2015. http://www.worldenergyoutlook.org/media/weowebsite/2015/CopyofWEO2015Biomassdatabase.xlsx.
[2] WHO, Household Air Pollution and Health, 2016. http://www.who.int/mediacentre/factsheets/fs292/en/ (accessed July 11, 2016).
[3] R. Bailis, R. Drigo, A. Ghilardi, O. Masera, The carbon footprint of traditional woodfuels, Nat. Clim. Chang. 5 (2015) 266–272, http://dx.doi.org/10.1038/nclimate2491.
[4] O.R. Masera, R. Bailis, R. Drigo, A. Ghilardi, I. Ruiz-Mercado, Environmental burden of traditional bioenergy use, Annu. Rev. Environ. Resour. 40 (2015) 121–150, http://dx.doi.org/10.1146/annurev-environ-102014-021318.
[5] Z.A. Chafe, M. Brauer, Z. Klimont, R. Van Dingenen, S. Mehta, S. Rao, K. Riahi, F. Dentener, K.R. Smith, Household cooking with solid fuels contributes to ambient PM2.5 air pollution and the burden of disease, Environ. Health Perspect. 122 (2014) 1314–1320, http://dx.doi.org/10.1289/ehp.1206340.
[6] M. Johnson, R. Edwards, C. Alatorre Frenk, O. Masera, In-field greenhouse gas emissions from cookstoves in rural Mexican households, Atmos. Environ. 42 (2008) 1206–1222, http://dx.doi.org/10.1016/j.atmosenv.2007.10.034.
[7] J.E.M. Arnold, G. Köhlin, R. Persson, Woodfuels, livelihoods, and policy interventions: changing perspectives, World Dev. 34 (2006) 596–611, http://dx.doi.org/10.1016/j.worlddev.2005.08.008.
[8] T. Bensel, Fuelwood, deforestation, and land degradation: 10 years of evidence from Cebu province, the Philippines, L. Degrad. Dev. 19 (2008) 587–605, http://dx.doi.org/10.1002/ldr.862.
[9] G. Hiemstra-van der Horst, A.J. Hovorka, Fuelwood: the “other” renewable energy source for Africa? Biomass Bioenergy 33 (2009) 1605–1616, http://dx.doi.org/10.1016/j.biombioe.2009.08.007.
[10] N. Hosonuma, M. Herold, V. De Sy, R.S. De Fries, M. Brockhaus, L. Verchot, A. Angelsen, E. Romijn, An assessment of deforestation and forest degradation drivers in developing countries, Environ. Res. Lett. 7 (2012) 44009, http://dx.doi.org/10.1088/1748-9326/7/4/044009.
[11] C.A. Cuvilas, R. Jirjis, C. Lucas, Energy situation in Mozambique: a review, Renew. Sustain. Energy Rev. 14 (2010) 2139–2146, http://dx.doi.org/10.1016/j.rser.2010.02.002.
[12] S.M. Oduori, F. Rembold, O.H. Abdulle, R. Vargas, Assessment of charcoal driven deforestation rates in a fragile rangeland environment in North Eastern Somalia using very high resolution imagery, J. Arid. Environ. 75 (2011) 1173–1181, http://dx.doi.org/10.1016/j.jaridenv.2011.05.003.
[13] S.M. Oduori, F. Rembold, O.H. Abdulle, R. Vargas, Assessment of charcoal driven deforestation rates in a fragile rangeland environment in North Eastern Somalia using very high resolution imagery, J. Arid. Environ. 75 (2011) 1173–1181, http://dx.doi.org/10.1016/j.jaridenv.2011.05.003.
[14] P.L. Sankhayan, O. Hofstad, A village-level economic model of land clearing, grazing, and wood harvesting for sub-Saharan Africa: with a case study in southern Senegal, Ecol. Econ. 38 (2001) 423–440, http://dx.doi.org/10.1016/S0921-8009(01)00189-6.
[15] S. Mandelli, J. Barbieri, L. Mattarolo, E. Colombo, Sustainable energy in Africa: a comprehensive data and policies review, Renew. Sustain. Energy Rev. 37 (2014) 656–686, http://dx.doi.org/10.1016/j.rser.2014.05.069.
[16] M.P. Kshirsagar, V.R. Kalamkar, A comprehensive review on biomass cookstoves and a systematic approach for modern cookstove design, Renew. Sustain. Energy Rev. 30 (2014) 580–603, http://dx.doi.org/10.1016/j.rser.2013.10.039.
[17] World Bank, Household Cookstoves, Environment, Health, and Climate Change: a New Look at an Old Problem, Washington DC, USA, 2010, http://documents.worldbank.org/curated/en/732691468177236006/Household-cookstoves-environment-health-and-climate-change-a-new-look-at-an-old-problem.
[18] I. Ruiz-Mercado, O. Masera, Patterns of stove use in the context of fuel–device stacking: rationale and implications, Ecohealth 12 (2015) 42–56, http://dx.doi.org/10.1007/s10393-015-1009-4.
[19] International Energy Agency, Africa energy outlook. A focus on the energy prospects in sub-Saharan Africa, world energy outlook spec. Report, Int. Energy Agency Publ. (2014) 1–237. https://www.iea.org/publications/freepublications/publication/africa-energy-outlook.html.
[20] J. Jetter, Y. Zhao, K.R. Smith, B. Khan, T. Yelverton, P. Decarlo, M.D. Hays, Pollutant emissions and energy efficiency under controlled conditions for household biomass cookstoves and implications for metrics useful in setting international test standards, Environ. Sci. Technol. 46 (2012) 10827–10834, http://dx.doi.org/10.1021/es301693f.
[21] P. Lloyd, H. Annegarn, C. Pemberton-Pigott, Towards a standard for clean solid-fuelled cookstoves, in: Domest. Use Energy (DUE), 2015 Int. Conf., Cape Town, 2015: pp. 13–17.http://dx.doi.org/10.1109/DUE.2015.7102957.
[22] G.L. Simon, R. Bailis, J. Baumgartner, J. Hyman, A. Laurent, Current debates and future research needs in the clean cookstove sector, Energy sustain. Dev. 20 (2014) 49–57, http://dx.doi.org/10.1016/j.esd.2014.02.006.
[23] M. Johnson, R. Edwards, V. Berrueta, O. Masera, New approaches to performance testing of improved cookstoves, Environ. Sci. Technol. 44 (2010) 368–374, http://dx.doi.org/10.1021/es9013294.
[24] C.M. Lee, C. Chandler, M. Lazarus, F.X. Johnson, Assessing the climate impacts of cookstove projects: issues in emissions accounting, Challenges Sustain 1 (2013) 53–71, http://dx.doi.org/10.12924/cis2013.01020053.
[25] C. Gupta, K.R. Usha, Improved chimneyless fuelwood cookstoves (Pondicherry region), Proc. Indian Acad. Sci. (Engg. Sci.) 5 (1982) 395–408.
[26] C. Roth, D. Andreatta, ETHOS Conference 2016: “Stoves 101,” in: vol. 2016. http://ethoscon.com/pdf/ETHOS/ETHOS2016/Roth.pdf.
[27] Yayasan Dian Desa, From Water Boiling Test to Water Heating Test Case Study of Indonesia, 2015. https://collaboration.worldbank.org/docs/DOC-15092.
[28] V.M. Berrueta, R.D. Edwards, O.R. Masera, Energy performance of wood-burning cookstoves in Michoacan, Mexico, Renew. Energy 33 (2008) 859–870, http://dx.doi.org/10.1016/j.renene.2007.04.016.
[29] C.A. Roden, T.C. Bond, S. Conway, A. Benjamin, O. Pinel, N. Maccarty, D. Still, Laboratory and field investigations of particulate and carbon monoxide emissions from traditional and improved cookstoves, Atmos. Environ. 43 (2009) 1170–1181, http://dx.doi.org/10.1016/j.atmosenv.2008.05.041.
[30] E.M. Carter, M. Shan, X. Yang, J. Li, J. Baumgartner, Pollutant emissions and energy efficiency of chinese gasifier cooking stoves and implications for future intervention studies, Environ. Sci. Technol. 48 (11) (2014) 6461–6467. http://dx.doi.org/10.1021/es405723w.
[31] S. Bhattacharya, D. Albina, A. Myint Khaing, Effects of selected parameters on performance and emission of biomass-fired cookstoves, Biomass Bioenergy 23 (2002) 387–395, http://dx.doi.org/10.1016/S0961-9534(02)00062-4.
[32] I. Ruiz-Mercado, E. Canuz, J.L. Walker, K.R. Smith, Quantitative metrics of stove adoption using Stove Use Monitors (SUMs), Biomass Bioenergy 57 (2013) 136–148, http://dx.doi.org/10.1016/j.biombioe.2013.07.002.
[33] R. Bailis, M. Ezzati, D. Kammen, Greenhouse gas implications of household energy technologies in Kenya, Environ. Sci. Technol. Sci. Technol. 37 (2003) 2051–2059, http://dx.doi.org/10.1021/es026058q.
[34] C. L'Orange, D. Leith, J. Volckens, M. DeFoort, A quantitative model of

cookstove variability and field performance: implications for sample size, Biomass Bioenergy 72 (2015) 233–241, http://dx.doi.org/10.1016/j.biombioe.2014.10.031.
[35] ISO, International Workshop Agreement IWA 11 Guidelines for evaluating cookstove performance, 2012. https://www.iso.org/obp/ui/#iso:std:iso:iwa:11:ed-1:v1:en.
[36] Aprovecho Research Center, Testing Stoves with the Water Boiling Test and the Controlled Cooking Test, (n.d.). http://aprovecho.org/portfolio-item/stove-testing/(accessed September 19, 2016).
[37] Kirk R. Smith, Karabi Dutta, Chaya Chengappa, P.P.S. Gusain, Omar Masera, Victor Berrueta, et al., Monitoring and evaluation of improved biomass cookstove programs for indoor air quality and stove performance: conclusions from the household energy and health project, Energy sustain. Dev. 11 (2007) 5–18, http://dx.doi.org/10.1016/S0973-0826(08)60396-8.
[38] USAID, Berkeley Air Monitoring Group, In-home Emissions of Greenhouse Pollutants from Rocket and Traditional Biomass Cooking Stoves in Uganda, 2011. http://pdf.usaid.gov/pdf_docs/PA00JBR7.pdf.
[39] C. L'Orange, D. Leith, J. Volckens, M. DeFoort, A quantitative model of cookstove variability and field performance: implications for sample size, Biomass Bioenergy 72 (2015) 233–241, http://dx.doi.org/10.1016/j.biombioe.2014.10.031.
[40] J. Robinson, M. Ibraimo, The Uncontrolled Cooking Test: Measuring Three-stone Fire Performance in Northern Mozambique, 2011. http://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/82-1.pdf.
[41] R. Hanna, E. Duflo, M. Greenstone, Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves, Am. Econ. J. Econ. Policy 8 (2016) 80–144. https://doi.org/10.1257/pol.20140008.
[42] T.W. Aung, G. Jain, K. Sethuraman, J. Baumgartner, C. Reynolds, A.P. Grieshop, J.D. Marshall, M. Brauer, Health and climate-relevant pollutant concentrations from a carbon-finance approved cookstove intervention in rural India, Environ. Sci. Technol. 50 (2016) 7228–7238, http://dx.doi.org/10.1021/acs.est.5b06208.
[43] N. MacCarty, D. Ogle, D. Still, T. Bond, C. Roden, A laboratory comparison of the global warming impact of five major types of biomass cooking stoves, Energy sustain. Dev. 12 (2008) 56–65, http://dx.doi.org/10.1016/S0973-0826(08)60429-9.
[44] M. Johnson, R. Edwards, A. Ghilardi, V. Berrueta, O. Masera, Why current assessment methods may lead to significant underestimation of GHG reductions of improved stoves, Boil. Point 54 (2007) 11–14. http://db.sparknet.info/docs/BP54-JohnsonEtAl-4.pdf.
[45] S. Joseph, Y. Shanahan, Designing a Test Procedure for Domestic Woodburning Stoves, ITDG, 1980, in: http://www.pssurvival.com/PS/Stoves/Cooking/Designing_A_Test_Procedure_For_Domestic_Woodburning_Stoves_1980.pdf.
[46] K.K. Prasad, Some Studies on Open Fires, Shielded Fires and Heavy Stoves, 1981. http://alexandria.tue.nl/repository/books/140651.pdf.
[47] S.F. Baldwin, Biomass Stoves: Engineering Desing, Development, and Dissemmination, 1987. http://www.newdawnengineering.com/website/library/Papers+Articles/Biomass Stoves, Engineering Design, Development and Dissemination, Samuel Baldwin 1987.pdf.
[48] V.V.N. Kishore, P.V. Ramana, Improved cookstoves in rural India: how improved are they? A critique of the perceived benefits from the National Programme on Improved Chulhas (NPIC), Energy 27 (2002) 47–63, http://dx.doi.org/10.1016/S0360-5442(01)00056-1.
[49] O. Adria, J.P. Bethge, Chinese National Improved Stove Programme (CNISP), (n.d.)http://www.bigee.net/en/policy/guide/appliances/policy_examples/12/ (accessed July 11, 2016).
[50] Bureau of Indian Standards, Indian Standard on Solid Biomass Chulha-specification, New Delhi, 1991, http://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/78-1.pdf.
[51] K.B. Sutar, S. Kohli, M.R. Ravi, A. Ray, Biomass cookstoves: a review of technical aspects, Renew. Sustain. Energy Rev. 41 (2015) 1128–1166, http://dx.doi.org/10.1016/j.rser.2014.09.003.
[52] Bureau of Indian Standards, Draft Indian Standard, 2013. http://www.bis.org.in/sf/med/MED04_1157C.pdf.
[53] Bureau of Technical Supervision Beijing, General Specifications for Biomass Household Stoves-China, 2008. http://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/77-1.pdf.
[54] R. Bailis, D. Ogle, D. Still, R. Edwards, The Water Boiling Test Version 3.0: Cookstove Emissions and Efficiency in a Controlled Laboratory, 2007. http://www.pciaonline.org/files/WBT_Version_3.0_0.pdf.
[55] The Water Boiling Test 4.2.3, 2014. https://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/399-1.pdf (accessed July 11, 2016).
[56] M. DeFoort, C. L'Orange, C. Kreutzer, Stove Manufacturers Emissions & Performance Test Protocol (EPTP), 2009. http://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/75-1.pdf.
[57] GERES Cambodia, Adapted Water Boiling Test (AWBT), 2010. http://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/75-1.pdf.
[58] SeTAR Centre, The Heterogeneous Testing Procedure for Thermal Performance and Trace Gas Emissions, Johannesburg, 2012, http://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/74-1.pdf.
[59] P. Arora, S. Jain, K. Sachdeva, Laboratory based assessment of cookstove performance using energy and emission parameters for North Indian cooking cycle, Biomass Bioenergy 69 (2014) 211–221, http://dx.doi.org/10.1016/j.biombioe.2014.07.012.
[60] K. Lask, K. Booker, T. Han, J. Granderson, N. Yang, C. Ceballos, A. Gadgil, Performance comparison of charcoal cookstoves for Haiti: laboratory testing with water boiling and controlled cooking tests, Energy sustain. Dev. 26 (2015) 79–86, http://dx.doi.org/10.1016/j.esd.2015.02.002.
[61] Y. Zhang, C. Pemberton-Pigott, C. Pernberton-Piqott, Z. Zhang, H. Ding, Y. Zhou, R. Dong, Key differences of performance test protocols for household biomass cookstoves, in: Domest. Use Energy (DUE), 2014 Proc. Twenty-second, IEEE, Cape Town, 2014, pp. 1–11, http://dx.doi.org/10.1109/DUE.2014.6827753.
[62] ISO/TC 285, Clean cookstoves and clean cooking solutions, (n.d.) http://www.iso.org/iso/home/store/catalogue_tc/catalogue_tc_browse.htm?commid=4857971&development=on (accessed July 11, 2016).
[63] GACC, Spreadsheet Changes for WBT 4.2.3, 2014. http://cleancookstoves.org/binary-data/DOCUMENT/file/000/000/403-1.pdf.
[64] GACC, GACC - Protocols, (n.d.). http://cleancookstoves.org/technology-and-fuels/testing/protocols.html.
[65] P. Arora, S. Jain, A review of chronological development in cookstove assessment methods: challenges and way forward, Renew. Sustain. Energy Rev. 55 (2016) 203–220, http://dx.doi.org/10.1016/j.rser.2015.10.142.
[66] T. Makonese, J. Robinson, C. Pemberton-Pigott, H. Annegarn, A Preliminary Comparison between the Heterogeneous Protocols and the Water Boiling Test, 2010. http://www.newdawnengineering.com/website/library/Papers+Articles/SeTAR/20110413 DUE Preliminary comparison of WBT and HTP, Makonese et al.pdf.
[67] Y. Wang, M.D. Sohn, How many replicate tests are needed to test cookstove performance and emissions? — Three is not always adequate, Energy sustain. Dev. 20 (2014) 21–29, http://dx.doi.org/10.1016/j.esd.2014.02.002.
[68] C. L'Orange, M. DeFoort, B. Willson, Influence of testing parameters on biomass stove performance and development of an improved testing protocol, Energy sustain. Dev. 16 (2012) 3–12, http://dx.doi.org/10.1016/j.esd.2011.10.008.
[69] R. Bailis, V. Berrueta, C. Chengappa, K. Dutta, R. Edwards, O. Masera, D. Still, K.R. Smith, Performance testing for monitoring improved biomass stove interventions: experiences of the household energy and health project, Energy sustain. Dev. 11 (2007) 57–70, http://dx.doi.org/10.1016/S0973-0826(08)60400-7.
[70] J. Rozis, Protocole de Test d'Ebullition de l'Eau Comparatif, 2008. http://docplayer.fr/15372220-Protocole-de-test-d-ebullition-de-l-eau-comparatif.html#download_tab_content.
[71] C. Pemberton-Pigott, Indonesian Clean Stove Initiative Pilot Programme Water Boiling Test Methods, Product Evaluation Criteria. https://cleancookstoves.org/binary-data/RESOURCE/file/000/000/370-1.pdf, 2014.
[72] D. Beritault, V. Lim, Stove Performance Report - Tao Payat PTT3, Tao Dum, 2013. http://www.newdawnengineering.com/website/library/Stove Testing/Testing Protocols/AWBT Adapted Water Boiling Test/LIV_13Oct09_TestingReport_Laos_Eng_v2_DBT.pdf.
[73] T. Makonese, C. Pemberton-Pigott, J. Robinson, D. Kimemia, H. Annegarn, Performance evaluation and emission characterisation of three kerosene stoves using a Heterogeneous Stove Testing Protocol (HTP), Energy sustain. Dev. 16 (2012) 344–351, http://dx.doi.org/10.1016/j.esd.2012.06.002.
[74] P. Raman, N.K. Ram, J. Murali, Improved test method for evaluation of biomass cook-stoves, Energy 71 (2014) 479–495, http://dx.doi.org/10.1016/j.energy.2014.04.101.
[75] P. Arora, P. Das, S. Jain, V.V.N. Kishore, A laboratory based comparative study of Indian biomass cookstove testing protocol and water boiling test, Energy sustain. Dev. 21 (2014) 81–88, http://dx.doi.org/10.1016/j.esd.2014.06.001.
[76] K. Bilsback, K. Hixson, M. Johnson, J. Kodros, J. Liacos, E. Lipscky, C. L'Orange, J. Pierce, B. Reynolds, A. Robinson, ETHOS Conference 2016: "A Novel Approach to Cookstove Testing: The Firepower Sweep Test," in: 2016. http://ethoscon.com/pdf/ETHOS/ETHOS2016/Bilsback.pdf.
[77] ISO, Strategic, Business Plan - ISO/TC 285-Clean Cookstoves and Clean Cooking Solutions, 2015. http://isotc.iso.org/livelink/livelink?func=ll&objId=17164114&objAction=Open&nexturl=%2Flivelink%2Flivelink%2Ffetch%2F2000%2F2122%2F687806%2Fcustomview.html%3Ffunc%3Dll%26objId%3D687806%26objAction%3Dbrowse%26sort%3Dname.
[78] GACC, Clean Cooking Catalog, (n.d.). http://catalog.cleancookstoves.org/ (accessed March 1, 2016).
[79] F. Riva, F. Lombardi, C. Pavarini, E. Colombo, Fuzzy interval propagation of uncertainties in experimental analysis for improved and traditional three – stone fire cookstoves, Sustain. Energy Technol. Assessments 18 (2016) 59–68, http://dx.doi.org/10.1016/j.seta.2016.09.007.