Development Engineering 6 (2021) 100065

Contents lists available at ScienceDirect

Development Engineering

journal homepage: www.elsevier.com/locate/deveng

A glimpse into real-world kitchens: Improving our understanding of cookstove usage through in-field photo-observations and improved cooking event detection (CookED) analytics

Evan R. Coffey [a,*], Elise C. Mesenbring [a], Maxwell Dalaba [e], Desmond Agao [b], Rex Alirigia [a], Taylor Begay [a], Ali Moro [b], Abraham Oduro [b], Zachary Brown [c], Katherine L. Dickinson [d], Michael P. Hannigan [a]

[a] *University of Colorado Boulder; Boulder, CO, 80309, USA*
[b] *Navrongo Health Research Centre; Navrongo, Ghana*
[c] *North Carolina State University; Raleigh, NC, 27607, USA*
[d] *Colorado School of Public Health; Aurora, CO, 80045, USA*
[e] *University of Health and Allied Sciences, Ho, Ghana*

ARTICLE INFO

Keywords:
stove usage
cookstoves
field validation
cooking detection
algorithm

ABSTRACT

The combustion of solid fuels in residential cookstoves is a global health and climate issue, and expanded use of improved cookstoves could have significant benefits locally and globally. Evaluating impacts of improved cookstove programs requires more accurately measuring stove use patterns. This work builds on and improves existing stove use monitoring methods. First, we introduce and describe a novel, in-field photo-observation sampling method designed to capture near-continuous, real-world, ground-truth stove usage information. These measurements are used to validate predictions made by electronic stove use monitors (SUMs). Second, we present Cooking Event Detector (CookED), a SUM algorithm that translates stove-temperature measurements into classifications of cooking or not-cooking. The predictive performance of the new algorithm is evaluated using results from the photo-observations and compared to existing algorithms. CookED demonstrates considerable improvement over some methods for all five types of improved and traditional stoves monitored in the study. Overall minute-level predictive accuracy of CookED ranges from 95.6% to 98.4%, depending on the stove type, while Matthews correlation coefficients range from 72.8% to 88.3%. Comparisons between predicted and observed average cooking event durations show high correlation (Pearson's $r = 0.85$). These methods can be applied in a wide variety of applications, including research studies linking behavior, technology, exposure, and human and environmental health, as well as operational programs that aim to scale up improved cookstove adoption and quantify benefits.

1. Introduction and background

To this day nearly 3 billion people rely on solid fuels for residential cooking services (World Health Organization, 2018b; The World Bank, 1098; Smith et al., 2014). When combusted, these fuels contribute to household and ambient air pollution which is estimated to contribute to 2.3 and 4.1 million annual premature deaths, respectively, and disproportionately impact women and children from low and middle-income countries (LMIC) (Smith et al., 2014; Cohen et al., 2017; Balmes, 2019; Babatola, 2018; Murray et al., 2020). Roughly 91% of the world's population live in places exceeding World Health Organization (WHO) air quality guidelines (World Health Organization, 2018a). Moreover, the inefficient combustion of solid fuels has been linked to adverse climate impacts, deforestation from nonrenewable harvesting of fuelwood, and a myriad of health equity and development impacts (World Health Organization, 2018b). Globally, residential cooking accounts for a significant proportion of these unhealthy activities and immense efforts have been made and are underway to study and curtail the impacts. Efforts financed through carbon credits, to reduce/offset greenhouse gas emissions via more efficient cookstoves, are often required to provide

* Corresponding author.
E-mail address: evan.coffey@colorado.edu (E.R. Coffey).

https://doi.org/10.1016/j.deveng.2021.100065
Received 1 April 2021; Received in revised form 10 May 2021; Accepted 21 May 2021
Available online 26 May 2021

independent evaluations of measured or observed traditional and improved stove usage.

Interventions focusing on how and to what degree improved cookstoves can reduce fuel expenditures and pollutant emissions and exposures must understand how and when intervention and traditional stoves are used. Monitoring the use and long-term adoption of intervention technologies is vital to ensure that current and future funding are spent wisely. As succinctly summarized by Wilson et al., understanding the challenge of cookstove adoption is underscored by the complex and heterogeneous nature of cooking owing to widespread cultural and environmental variation in culinary practices (Wilson et al., 2015). For example, stove or fueling stacking, or use of multiple stove or fuel types, is a prevalent global phenomenon (Shankar et al., 2020; Ruiz-Mercado and Masera, 2015). Knowledge of the composition and characteristics of a household's "stack" can substantially bolster our understanding of the interrelationships between cooking behavior, exposure and health outcomes (Masera et al., 2000). Explicitly, a deeper and more nuanced comprehension of household stove usage is fundamental to the task of characterizing the linkages between the adoption of new (cleaner) technologies and disadoption of traditional stoves and their associated health and climate impacts.

1.1. Measurement methods of cookstove usage and adoption

To date, cookstove usage has been primarily measured in four ways: through participant surveys, observations-upon-visit, participant diaries or direct physical measurements. Each method has its strengths and limitations.

1.1.1. Surveys

Surveys traditionally measure study participants' self-reported use of stoves over discrete periods, and across multiple time points by asking respondents about usage patterns the day of or prior to the survey or in the past week (Dickinson et al., 2015, 2018; Piedrahita et al., 2016). In doing so, surveys are inherently prone to respondent recollection error and bias. Even where studies have shown reasonable agreement between surveys and electronic stove use monitoring for cooking event detection (Piedrahita et al., 2016), there are other important parameters that surveys are not well designed to measure, such as cooking event durations. Additional drawbacks of surveys include relatively large financial and logistical resources (e.g., personnel, time, transportation). At the same time, surveys can yield in-depth information of usage behaviors (e.g., what meal type was prepared on which stove using what fuel(s) and during what time of day and for how many people) that can be important for assessing *why* certain stove use patterns emerge (Dickinson et al., 2019; Rhodes et al., 2014; Williams et al., 2020; Hollada et al., 2017). Extensive glimpses into individual cook's preparatory preferences and cooking behaviors are made tangible through surveys.

1.1.2. Observations-upon-visits

Observations-upon-visit typically consist of research field staff planning and visiting a study home and noting whether or not each stove is 'on' with accompanying information similar to the survey (e.g., meal type, fuel type, time of day etc.). The main strength of this method is accuracy (e.g., minimizing recollection error and reducing bias) in addition to contextual data. Precision of this method depends on the subjectivity among multiple observers of what constitutes a stove usage event, a context-specific definition which has been shown to be interpreted a number of different ways depending on the culture and individual (Piedrahita et al., 2020; Graham et al., 2014). However, observational visits come with many of the same shortcomings of surveys and have been shown to produce a significant behavioral bias in cookstove users when they perceive their actions as being observed (Simons et al., 2017, 2018), known as the Hawthorne effect. Others point out that these observational measurements can be inconvenient to participants and are limited to time periods where household visits are appropriate (e.g., daytime/business hours, when a participant is home) and often do not encompass entire cooking events from start to finish (Simons et al., 2014).

1.1.3. Diaries

Diaries have also been used to capture participants' self-recorded activities including stove usage (Pilishvili et al., 2016). Diaries (electronic or paper/pencil) offer temporally resolved information of when and for how long certain stoves were operated in addition to what meals were prepared. However, drawbacks to this method include challenges making diaries useful in low-literacy samples, data prone to human bias, and burden on participants to complete at every meal or at predefined intervals.

1.1.4. Direct physical measurements

The fourth method - direct physical measurements - relies upon installing devices, commonly referred to as stove use monitors (SUMs), at study households to record direct measurements (e.g., temperature, light, pressure, electrical current etc.) that correlate with stove usage. This approach proffers unobtrusive, objective measurements of continuous stove usage information leading to more precise estimates of cooking duration and time of day usage trends (Piedrahita et al., 2016; Pillarisetti et al., 2014). Stove use monitoring systems (SUMS) refers to the overall stove usage monitoring instrumentation, data analysis and reporting platforms and frameworks (hardware and software) (Ruiz-Mercado et al., 2012).

The most common stove-usage correlate is temperature (typically of the stove body or a locale where temperature variation is explained by stove operation). To our knowledge, the monitoring of cookstove usage via temperature measurements was pioneered by Ruiz-Mercado and colleagues – where traditional and improved Guatemalan cookstove usage were monitored in study households using small temperature loggers, called Thermochron iButtons, now ubiquitous to stove use monitoring (Ruiz-Mercado et al., 2008). They established standardized stove usage metrics to quantify adoption and allow comparisons of stove usage between and within study groups and households (Ruiz-Mercado et al., 2013). Over the last decade, others have improved upon SUMS technology, and best practices incorporating a wider range of usage indicators such as infrared light or electrical current (e.g., stoves with battery-powered fans) (Wilson et al., 2015, 2020; Piedrahita et al., 2016, 2020; Pillarisetti et al., 2017). Many of these device measurements require signal normalization or comparison to some reference measurement, say of the ambient environment, from which to derive meaningful usage information and account for variation in stove operation or environmental characteristics (Ruiz-Mercado et al., 2012; Pillarisetti et al., 2018; Sundararaman et al., 2016).

Drawbacks of SUMS include the upfront resources of acquiring and installing them in study households as well as often-overlooked analysis resources, including the technical expertise and time to retrieve, process and analyze the data. Extensive undertakings to provide open-sourced SUMS algorithms and advanced data processing tools to end users have spurred a new wave of interest in SUMS, significantly reducing the data analysis burden (see Geocene Studies, SUMSarizer and SUMIT) (Wilson et al., 2020). These valuable services have enabled practitioners the capacity to label (see TRAINSET), process and interpret vast amounts of SUMS data over the web using event-detection algorithms (e.g., SuperLearner, FireFinder, Threshold detectors etc.) which feature robust machine learning and context-specific analytics (Wilson et al., 2020). However, these methods often lack the rigorous in-field validation information required to train or adequately parameterize the algorithms, relying on the intuition of expert users of what temperature data constitutes the start and end of a cooking event.

SUMS data attrition is also a widespread issue documented by many research groups (Wilson et al., 2015; Piedrahita et al., 2016; Simons et al., 2014; Pillarisetti et al., 2018). Data attrition is typically the result of poorly operating devices in extreme cooking environments, and resources to identify and replace such devices in the field can be quite substantial (Pillarisetti et al., 2014). Significant device improvements have been made including wireless communications, increased durability and reliability, and manufacturing advancements reducing unit costs.

On their own, SUMS data also lack key information on meal types, quantity of food prepared, and fuels used, such that pairing SUMS measurements with diaries or surveys can provide a more comprehensive picture of stove use patterns. In addition, time-resolved stove usage measurements can be paired with participants' contemporaneously measured location (within or beyond their homes) and exposure to air pollution, opening the door to a much deeper understanding of the effects interventions may have on behavioral outcomes such as how much time cooks spend in various proximity to their stoves, time-of-day analyses combining exposure outcomes and stove usage metrics, and analyses attributing exposure to specific cooking behaviors (Piedrahita et al., 2019).

Given the strengths and limitations of stove usage monitoring methods, a combination of approaches is often implemented (e.g., devices alongside surveys) for a well-rounded, more comprehensive understanding of usage and adoption (Piedrahita et al., 2020; Stanistreet et al., 2015). Leveraging multiple methods is also instrumental in assessing data quality - comparing outcomes across two or more methods.

1.2. The need for improved, in-field cookstove ground-truthing for SUMS event detection algorithm validation

A review of the SUMS literature to date yields few examples of rigorous in-field validation of published SUMS event detection algorithms. Methods to relate SUM signals to cooking status have predominantly been conducted in laboratory or simulated real-world settings, while field-based SUMS studies usually lack an observational ground-truth of stove usage due to cost and time constraints - leaving practitioners reliant upon expert-user intuition of cooking periods (Pillarisetti et al., 2017, 2018). Surveys and robust SUMS results have been compared but neither is sufficient to act as a ground-truth for its counterpart. Simon and colleagues performed in-field, observation-upon-visit ground-truthing of iButton SUMs in Uganda by visual inspection of cooking status during household visits (Simons et al., 2014). These observations consisted of recording the time, location, stove ID and cooking status at discrete time points during normal household visitations. A series of logistic regressions were used to transform 30-min SUM temperature readings on three-stone fires (TSF) and Envirofit stoves to probabilities of stove usage using observations from the field as the ground-truth. Hour of the day and lead/lagged temperature readings variables improved the performance of the model to an extent. The out-of-sample predictive accuracy of the best model was 87.8% and 91.1% for the TSF and Envirofit stoves, respectively (Simons et al., 2014). However, as the authors noted, the sensitivity of the model (i.e., probability of predicted "cooking" when observed as "cooking") was only 56.5% for TSF and 77.6% for Envirofit stoves emphasizing how unbalanced sample classes (i.e., an overwhelming majority of the time a stove is not on) can result in varying misclassification rates.

Pillarisetti et al. performed in-field validation of SUMs (The Pink Key, a meal counter) installed on liquid petroleum gas (LPG) stoves implemented to allow conditional cash transfers based on the number of recorded cooking events. The validation consisted of in-field observations of controlled cooking events on instrumented LPG stoves as well as duplicate measurements of usage using thermocouple-based SUMs (Pillarisetti et al., 2018). Data from the thermocouple-based SUMs (Wellzion) underwent a threshold algorithm to estimate meal counts, which was then used as the reference for comparison to the Pink Key. Results from the controlled cooking tests where only one burner was used showed the Pink Key accurately predicting one meal made in 53/59 cases (90%), inaccurately measuring no meals in 3/59 cases (5%), and inaccurately measuring two meals in 3/59 cases (5%). When using 2 burners, the performance was slightly worse (9/12 correct). Agreement between the two SUM methods was fair ($r^2 = 0.35$) with the Pink Key underestimating meal counts by 3.5 meals, on average, over 6-day periods (Pillarisetti et al., 2018).

Graham and colleagues performed validation of a wireless cookstove sensor coupled with a custom SUM ("J-bar") on forced-draft improved cookstoves in two households during kitchen performance tests (KPTs) in Uttar Pradesh, India. Decision tree architecture was used to classify SUM temperature data as "cooking" or "not cooking," and aggregated comparisons between calculated event duration and observed event duration showed high correlation ($r^2 = 0.82$) and moderate absolute error (15%)(Graham et al., 2014). However, without comparisons event by event, it is difficult to say how accurately the start and end of each event were classified. Also, capturing variation in stove construction and household cooking preferences was limited to the two households. Sundararaman et al., applied machine learning techniques to SUMS data collected from Ugandan-made Makaa improved cookstoves (Sundararaman et al., 2016). They compiled target data with which to train the algorithm by manually identifying start and end times of cooking events informed by temperature thresholds or peak detection - piloting a hybrid approach to SUMS validation. To our knowledge, no study has collected near-continuous, in-field stove usage observational data with which to validate SUMS methods.

1.2.1. Filling in the gaps towards improving best practices

In this work we extend the literature on stove usage and adoption by introducing and describing a novel, in-field validation methodology our team piloted in Northern Ghana, incorporating an observational, photo-documentation system (i.e., camera) of cooking activities. This method provides a more substantial temporally-resolved ground-truth of stove usage for three types of improved stoves and two traditional stoves. In addition, we describe and evaluate the performance of a new cooking event detection and duration estimating algorithm, termed CookED (Cooking Event Detector), and demonstrate potential generalizability beyond our study households, stove types, and culinary cultures.

2. Materials and methodology

2.1. Instrument selection/description

This work took place as part of the Prices, Peers and Perceptions (P3) biomass and LPG cookstove study in Northern Ghana which involved 600 households evenly split between the biomass ("P3 Bio")(Dickinson et al., 2018) and LPG ("P3 Gas") (Dalaba et al.,) arms. Briefly, through use of household surveys, SUMS, and low-cost, portable monitoring equipment, the P3 study measured how prices and peers' experience affect perceptions of stove quality, the decision to purchase a stove, use of improved and traditional stoves over time, and personal exposure to air pollutants from the stoves. Here we focus on the SUMS methodology developed during this study.

Thermocouple data loggers (Thermocouple Temperature Data Logger SSN-61, Wellzion, USD21.5) were used as SUMs with Type K thermocouples (1M K Screw Thermocouple and 2M Customized K Thermocouple, Wellzion 1.5USD). The SUMs are battery-powered (18-month battery life), log 32,000 recordings at a user-defined interval (1sec-12hrs) and have a measurement range of approximately −270 °C to 1200 °C, which allows them to be placed as close to the hottest portion of the cookstove as desired with less risk of overheating, resulting in clearer designations between cooking events and temperature effects from other activities (e.g., other nearby stoves, heating from direct sunlight or ambient temperature rises) a challenge documented nearly a decade ago (Ruiz-Mercado et al., 2012). Furthermore, a total of 5 stove types were expected to be encountered in our study region from baseline surveys (Dickinson et al., 2018; Dalaba et al.,), including the traditional three-stone fire (TSF) and a locally made charcoal stove called "coalpot" as well as the intervention stoves: the Greenway Jumbo rocket stove, ACE1 semi-gasifier stove and 2 variations of LPG stoves - a locally made 1-burner and a variety of 2-burner models (Fig. 1, additional images of installations can be found in Section 1 of the Supplementary Information).

Thermocouple probe type was chosen following testing of both short, and long threaded probes on all cookstove types in this study (Jumbo, ACE, LPG stoves, various coalpots, and three stone fires). It was found that for the Jumbo, ACE, LPG stoves, and coalpots, the 1M K Screw Thermocouple best captured a large enough temperature increase to recognize a cooking event while performing test-cooking at CU Boulder in the sun or shade. When testing thermocouple types on three stone fires in Navrongo, Ghana, the 6.5-inch 2M Customized K Thermocouple was able to be placed further into the fires and thus captured these cooking events best. Each data logger was enclosed in a standard-size (2-inch diameter, 9-inch length) PVC pipe unit to protect from water damage, direct sun exposure or tampering. The units were sealed permanently at one end and equipped with a watertight screw-top adapter on the other end to allow the field team access to the data logger for data downloads in the field. A slit in each unit allowed the stainless-steel thermocouple sheath containing the wires to exit the unit and extend towards the cookstove. This slit was waterproofed with silicone sealant.

Data Graph software (PC only) provided by Wellzion was used to set up the SUM logging interval as well as download and plot the SUMs temperature data. The SUMs were set to log a temperature reading every 5 min, balancing stove usage resolution and frequency of field visits (111 days of available memory).

2.2. Device installation and data organization

Lab testing at CU Boulder allowed us to determine the proper placement of the thermocouple probe on each stove type. Each stove was used for test-cooking with various probe placements on the stove until a 'best-case' placement was determined to see the highest rise in temperature during cooking. The Ghanaian field team was trained in these specific stove installation protocols for each stove type they expected to encounter at households. Complete coverage, or monitoring of all stoves (intervention or traditional, Fig. 1) at each household, was achieved by field staff performing household walkthroughs with each primary cook at each visit and identifying all stoves for SUM-monitoring.

For the ACE and Jumbo stoves, SUMs were secured to the side of the stove using metal tube brackets screwed directly into the stove body. The sheet metal on the stove bodies were prepared (i.e., tapped) for screwing of the thermocouple probe approximately one third of the way up the stove. For coalpots, the field team's discretion was needed. If the coalpot had a handle close enough to the portion of the stove where coals were placed, a small bracket was attached to the handle and the screw-probe was secured into this bracket. If there was no existing handle or bar to mount a bracket, the team obtained permission to drill one hole into the stove into which the screw probe could be secured. Eighteen-gauge stainless steel wire was then used to further secure each SUM to the coalpot and extend the PVC far enough from the stove body to avoid melting. Each burner on LPG stoves were monitored by installing a SUM with a thermocouple screw-probe terminating 1 inch away from each burner element. Each SUM unit was then secured underneath or on the side of each LPG stove with metal tube brackets. The SUM installation on the locally made 1-burner intervention LPG stove was similar to that of the coalpot with the screw-probe mounted on a bracket adjacent to the burner element. Thread locker was used on all screw-probes to prevent loosening over time and tampering.

SUM placement for TSFs was more challenging due to inherent variation in stone layouts, proximity to walls as well as ground type - often leading to more complications in TSF event detection (Simons et al., 2014). Robust SUM thermocouple probes and wire sheathing allowed for more placement options than more vulnerable tools with lower operating temperatures and weaker materials. To standardize the placement of TSF SUMs to the greatest extent, thermocouple probe ends were placed within six inches of the center of the fire onto an area of the

Fig. 1. Stove usage monitors (SUMs), comprised of thermocouple dataloggers (SSN-61) enclosed in PVC piping (2-inch diameter, 9-inch length), installed on 6 varieties of stove types encountered in the Prices, Peers and Perceptions (P3) cookstove study. Clockwise from the top left; Greenway Jumbo rocket stove (thermocouple logger visible), ACE1 semi-gasifier stove, 2-burner LPG stove, locally made 1-burner LPG stove, locally made charcoal-burning 'coalpot' and the traditional three stone fire (TSF) with standard placement of probe terminus 6 inches from center of combustion area. Additional instrumentation images can be found in SI Section 1).

ground that was soft enough to be drilled (Fig. 1). The thermocouple probes and wires were secured with multiple ground staples, which often had to be drilled for adequate depth. The SUM PVC enclosure was set back from TSF 3–4 feet to avoid heat damage.

Each SUM, PVC enclosure, and monitored stove was labeled with a waterproof, UV-resistant quick response (QR) barcode sticker and metal ‘dog’ tag to keep track of which SUM ID was associated with which stove ID. QR codes were labeled with unique, 3-character IDs (ranging from 001 to 600) such that QR code stickers with the same number would be placed on the SUM and PVC enclosure. A separate QR code was affixed to the monitored stove. Metal ‘dog’ tags were inscribed with the same code and affixed to each stove for redundancy.

At SUM installation as well as at each data download visit, an electronic survey through the program Open Data Kit (ODK) was completed with the primary cook to record the following information for each stove: reported stove use, any issues with the stove, stove location, fuel consumption, and any problems with the SUM. The ODK survey also allowed field workers to scan the QR code for each SUM and its corresponding stove (or a replacement SUM and its corresponding stove) to continuously keep track of which SUM was attached to which stove. Field workers downloaded each SUM data file using the USB interface and DataGraph software and exported them as a comma-separated-value (CSV) file with a name corresponding to the SUM ID from the QR code or tag.

2.3. In-field SUMS validation methodology

2.3.1. Materials and procedures for in-field photo data collection

During routine visits to SUM-instrumented study households enrolled in 48hr exposure monitoring, battery-powered outdoor cameras (Apeman, USD42, Fig. 2a) were programmed to collect images of stove activity for validation purposes (Fig. 2b). These images were compared to temperature data collected from the SUMs to train and evaluate performance of an existing event detection algorithm and the one described here. Two cameras were deployed at each consenting household; one placed in the primary cooking area and a second in the secondary cooking area, according to the primary cook. Households in this study area typically have two cooking areas - one covered or located inside a structure and a second outdoors. The cameras were attached to tripods roughly 0.5 m tall, placed approximately 4 m from the cooking area and pointed towards where the cookstoves would be used.

A 32 GB SD card was used in each camera, allowing storage for approximately 48 h of photographs taken, silently, every minute. The limiting factor in monitoring was how long the 8 rechargeable AA batteries lasted in each camera, resulting in a variable number of photographs at each deployment. A precautionary measure of setting video length and interval to the shortest time options was used in case the user accidently selected a video option as opposed to a photograph option in the camera. When activated, the camera took an inaudible, locally timestamped photograph every 60 s of the kitchen area. The camera utilizes infrared light, not visible to the human eye, to document cooking in dark settings. After the cameras were set up, the participants were asked to stand near the cookstove(s) for 2 min after which the camera was checked to ensure images were being properly captured and the cookstove(s) and participant were visible. A field record sheet was populated with the date, camera ID, kitchen area descriptors and SUM IDs for stoves at the household.

After approximately 48 h, field enumerators returned to the households to collect exposure tools and participants were given the choice of having any of the images deleted. Data from cameras were downloaded to password-protected servers.

2.3.2. Classifying cooking status from field images

Images from the cameras and information from the camera log sheets were used to classify usage of each stove visible within the frames at each available minute throughout the deployment (Fig. 3). Specifically, three codes were used to classify usage; ‘cooking’, ‘not cooking’ or ‘undefined’. Indicators of cooking included smoke plumes, visible flames, shifting pots and lids as well as evidence of refueling of stoves. Subjectivity of what constituted an active cooking event on a stove was minimized by reducing the number of independent validation classifiers (research staff). Research staff members, following a consistent set of classification criteria, inspected each image and recorded usage codes of each stove in the frame for the corresponding time indicated in the time-stamped image. If a stove was out of the camera frame or no classification could be made, the cooking classification at that time was set to ‘undefined’ and not used in subsequent validation. In cases where an observational cooking transition (i.e., start or end point) was unclear, a transition point delineation was estimated by marking the middle point of the unclear interval (median:6 min, mean: 11min SD: 10min) and labeling classifications on either side of that point accordingly.

2.4. New stove event/duration identification algorithm

Next, we describe and evaluate a new stove usage algorithm, henceforth referred to as CookED, designed for generalizability across many SUMS temperature loggers, stove types and study environments. Building off past SUMS event-detection algorithms, CookED incorporates an array of user-defined parameters tailored to systematically and automatically differentiate temperature signals associated with stove usage from those which are not. Parameter values can be refined or optimized on a case-by-case basis using results from in-field validation methods (described above). This algorithm was developed in the Matlab environment (Mathworks, R2019b).

2.4.1. Description of CookED

Simply put, CookED inspects temperature traces (temporally-resolved temperature measurements from SUMs) for distinct features in

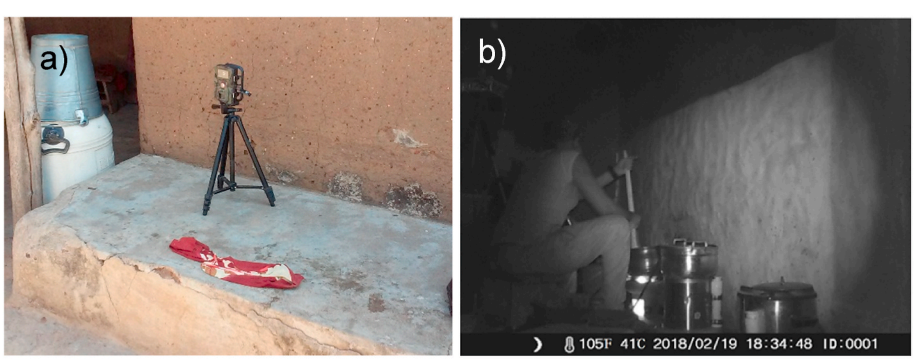

Fig. 2. a) the battery-powered, outdoor camera used to collect in-field, stove usage photo-observations and b) an example time-stamped image taken after dark in one of the rural study households showing meal preparation on a coalpot stove (the Infrared spotlight visible in this image was not visible to participants).

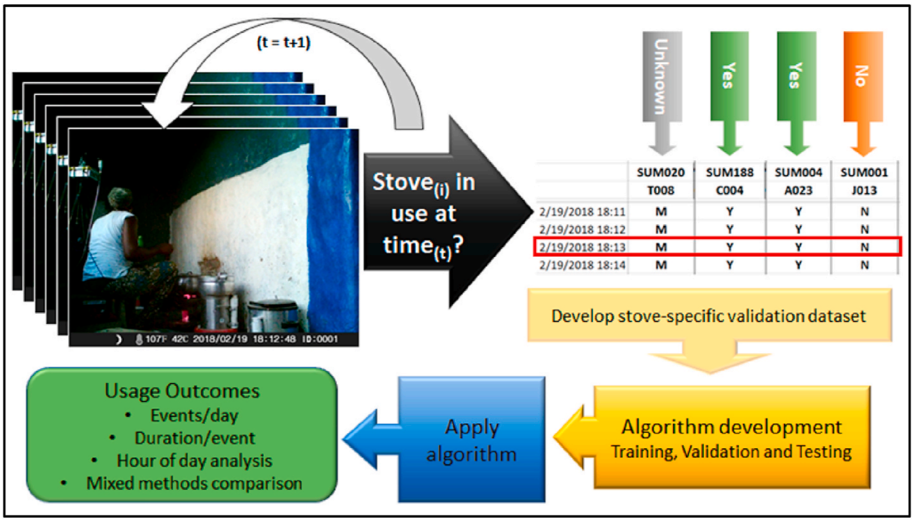

Fig. 3. Schematic of stove usage classification from the photo-observations. Each stove in each image is inspected for indicators of usage and classification of 'cooking', 'not cooking' and 'undefined' status are recorded for each time. These ground-truth data are available to train, validate and test event-detection algorithms towards estimating relevant stove usage outcomes.

the first and second derivatives with respect to time. Peaks in temperature are used to identify potential cooking events (or clustered events) whereas the first and second derivatives assist in defining the beginning and end of each event. Clustered events within a user-defined period of time are combined and all minutes are classified as either 'cooking', 'not cooking' or 'not logging'. An overview of the algorithm is described in Fig. 4. All 6 steps shown in Fig. 4 are automated but CookED input parameters are manually defined.

Step 1. Prefilter: data flagging and interpolation

CookED first performs prefiltering to discern valid temperature data from data to omit. The temperature loggers used in this study have an error reading of −270 °C which typically indicates a poor connection between thermocouple leads. A robust, yet conservative filter was used to flag intermittent connectivity resulting in sporadic temperature readings. A low temperature parameter (user-defined) is used to flag any data falling below a realistic temperature threshold for the study environment (0 °C in this study). Flagged temperature indices indicate times when the SUM was not logging properly and thus set to "NaN" (not a number). After flagging invalid data, interpolation between logged temperature recordings is optional and may provide more accurate estimates of cooking event duration yet caution is to be exercised in the extent of interpolation and note that some parameters could require adjustments depending on the logging interval selected (e.g. modifying a 5 °C/5-min slope to reflect a 1 °C/1-min slope). In this work, 5-min readings were replaced by 1-min readings using linear interpolation.

Step 2. Calculating time derivatives of temperature trace

Following flagging and interpolation, the first (dT/dt, T′) and second (d^2T/dt^2 , T'') time derivatives of temperature are calculated by twice differentiating the temperature trace (Fig. 5). Positive values of the first derivative indicate heating while negative values indicate cooling ($°C{*}min^{-1}$). The second derivative describes curvature of the temperature trace or the rate of temperature change of the probe ($°C{*}min^{-2}$). CookED utilizes the magnitude of the second derivative (|T"|, absolute values) to identify instances of large rates of temperature change. A third variable keeps track of when the first derivative changes sign (from heating to cooling or vice versa) indicative of turning points (e.g., local minima or maxima) in the trace.

Step 3. Identifying peaks in temperature traces indicative of cooking events

Peak values in SUM temperature traces are often indicative of cooking activities. CookED was designed to identify these peaks as a first-pass proxy for possible events. The built-in Matlab function, *findpeaks*, was used to locate peak values in the temperature trace as well as the first and second time derivatives (see Fig. 5). *Findpeaks* utilizes 4

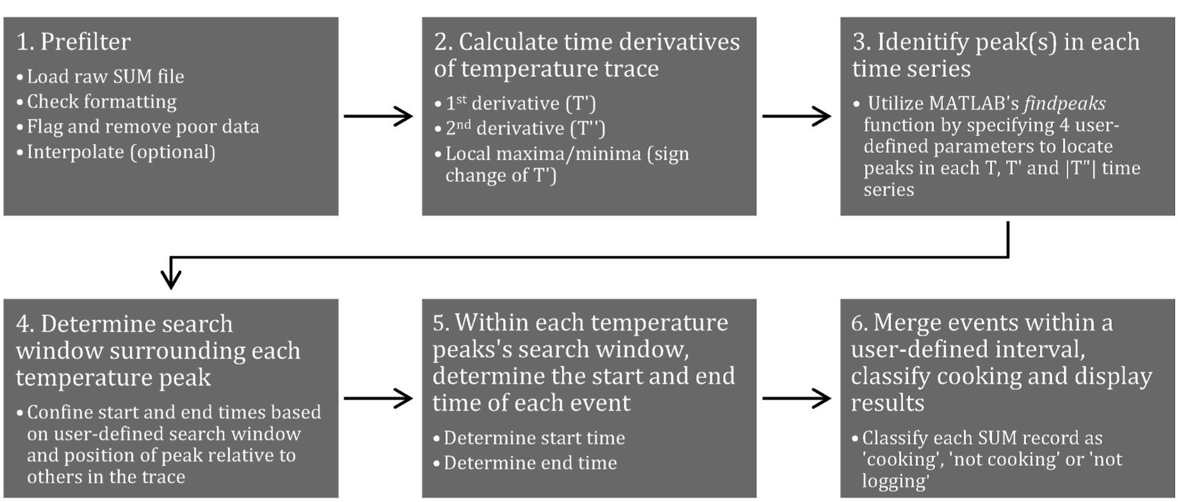

Fig. 4. Automated processing diagram for CookED delineated into 6 stages.

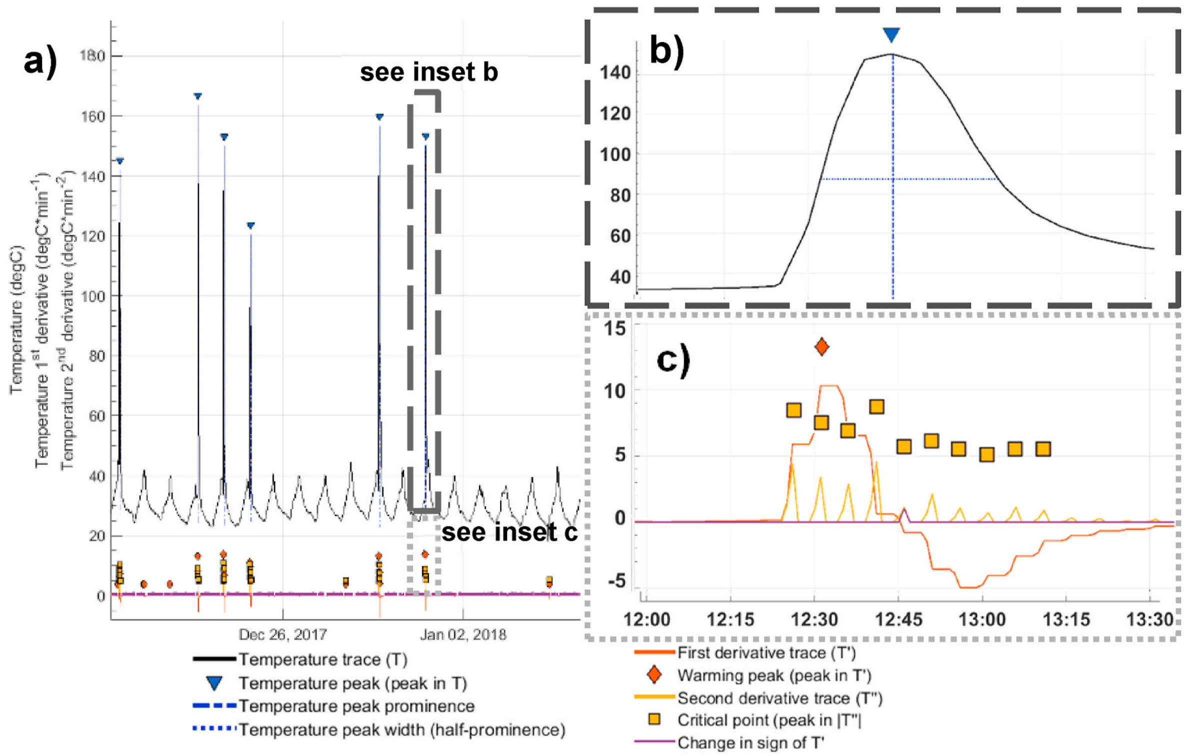

Fig. 5. Panel a) an example SUM time series indicating a prefiltered temperature trace (black soild line, T°C), first (orange solid line, T'°C*min^{-1}) and second (yellow solid line, |T''|°C*min^{-2}) derivative traces and peaks (colored and shaped markers) located in all three traces. Changes in the sign of T' (purple solid line) represent turning points in the temperature trace. This trace has a total of 6 candidate cooking events. Panel b) an enlarged upper section of the last temperature peak and panel c) an enlarged lower section of the first and second derivative traces and associated features of the last peak.

user-defined parameters (these should be tailored to each stove type monitored) to identify and distinguish peaks associated with potential cooking events from errant peaks caused by other factors (direct sun radiation, heat from nearby stoves, etc.). The 4 parameters describe minimum peak height (i.e., minimum temperature to be considered a peak), the prominence of a peak (i.e., height relative to nearby baseline temperature values), the minimum width of a peak (i.e., the time interval between points on the trace where half the prominence is bisected by a horizontal line) and minimum distance between peaks (i.e., minimum time interval between consecutive peaks). The locations (time) of the peaks in each trace and values of peaks (temperature °C, heating/cooling rate °C*min^{-1}, rate of temperature change °C*min^{-2}) are returned from the function. Discussed below, the subsequent analyses of the locations of these peak values, relative to one another, yield start and stop times of individual events.

Step 4. Determine the search window surrounding each temperature peak

Before individual event start and stop times can be estimated, a search window around each peak is automatically defined and this search window depends on how that individual peak is classified. The 4 classes of individual temperature peaks in a trace are described below.

There is the case where no peaks are identified in a trace - this is a simple case where no evidence is found implying an event. Then, there are 4 possible cases for any given peak in a temperature trace containing at least one peak; 1) it is the first peak of a trace containing two or more peaks 2) it is the last peak of a trace containing two or more peaks 3) it is not the first nor last peak of a trace containing two or more peaks 4) it is the only peak in the trace. Each one of these cases determines how the study-and-stove-specific, default search window parameter will be modified to home in on start and end times of that prospective event identified by the peak. Namely, this search window helps refine the times at which an event starting time or ending time could take place and isolates each temperature peak (identified in the previous step) for individual analysis. For this work, the default search window was set to approximate the longest event described by a single peak in temperature for each stove type monitored - which was estimated to be 240 min (4hrs) for all stove types.

In case 1, the search window of time spans 120 min before the peak (no previous peak to consider) and 120 min after the peak unless the next peak is within 120 min in which case the window ends midway between the two peaks. In case 2, there is no subsequent peak to consider, so a window ending 120 min past the peak is established and the start of the window is set to the end of the previous cooking event if that event ended within 120 min of the peak. In case 3, the search window is modified using a combination of conditionals from cases 1 and 2 where the window boundaries are 120 min in both directions from the peak unless a previous cooking event or a subsequent peak overlaps with this window and it is shortened. Case 4 is a simple case where the window is 120 min before and after the only peak.

Step 5. Within each temperature peaks's search window, determine the start and end time of each event

Now that a search window is defined around each temperature peak, start and end times of individual events can be determined by iterating through each peak and assessing relationships among the variables described above (i.e., T′, |T"|, sign switch of T′). These relationships are assessed within each peak's search window.

Estimating the start time of an event:

Within the search window around each peak, an approximate start time of an event is estimated by searching for specific criteria of T′, |T"| and the instances of when the sign of T′ switches (T' = 0). Specifically, the algorithm first searches for the critical point (peak in |T"|) directly preceding the nearest local temperature minima (T' = 0, indicated by sign of T′ switching) before the first warming peak (peak in T′). In other words, the algorithm is searching for the time at which the curvature of the temperature trace is highest, prior to the point at which the temperature is increasing rapidly (due to the heat of combustion) directly preceding the nearest local minima temperature to said point. If this point in time is identified with these criteria, this time is recorded as the approximate start of the event. If no approximate start time is identified, a secondary set of criteria are assessed. These criteria attempt to locate the critical point preceding the time at which only heating (T'>0) occurs until the first warming peak. If no critical point is located, the last criterion locates the first critical point preceding the warming peak or in rare cases, simply the first critical point in the search window. If no approximate event start time is identified, the user is notified, and the proxy-event is flagged.

The event starting time is estimated by adjusting the approximate start time (earlier) by applying a thermal lag parameter, or the average number of minutes it takes for heat within the combustion zone to register a temperature increase at the SUM measurement point. The thermal lag parameter is dependent on stove type as well as placement and response time of the device. Stoves with larger thermal mass will have larger thermal lags. Stoves in this work had thermal lag parameters ranging from 8 to 10 min estimated from the pilot testing of SUMs and comparing photo-observations to temperature trace data.

Estimating the end time of an event:

The end time of an event is identified using analogous sets of criteria to account for the high variability among characteristics of individual temperature traces. First, CookED checks if there is at least one warming peak after the temperature peak. If there are none, the temperature peak time is considered the event end time. If there is a warming peak after the temperature peak, the algorithm searches for the last warming peak and identifies the nearest, subsequent local maximum temperature (T' = 0, indicated by sign of T' switching) as the event end time. If no end time is identified at this point, the algorithm then searches for the nearest critical point (peak in |T"|) following the last warming peak. If these criteria do not result in an end time, the time of the temperature peak, itself, is used. Fig. 6 illustrates the process of identifying the start and end of an example event.

Step 6. Merge events within a user-defined interval, classify cooking and display results

Clustering events into a single cooking event:

Individual events, detected in previous steps, occurring within a user-defined interval of time (CookED parameter: *meal_spacer*) are automatically merged together to form one continuous cooking event. This was achieved by iteratively comparing start and end times of adjacent events and adjusting start and end times accordingly. This step is critical in accurately estimating the number of cooking events and therefore the average event duration. In this work, a duration of 30 min was determined to adequately distinguish consecutive events on a single stove.

Handling data indicative of rare scenarios:

The algorithm was also designed to handle and alert the user of rare events; events starting at the beginning of a trace or occurring up until the end of a trace. These instances occur when data is retrieved from a SUM during a cooking event interrupting the temperature time series. These occurrences can be limited when durations between SUM data retrievals are maximized or when data retrieval methodologies account for active cooking upon arrival perhaps by waiting for active cooking to end before downloading data.

Visualizing cooking events:

After merging and establishing final start and end points of cooking events, each measurement in the trace is categorized as either 'cooking', 'not cooking' or 'not logging'. A plot of the categorized temperature trace is generated for a quick assessment of cooking events (see Fig. 7).

2.4.2. Deploying CookED on P3 SUM data

Initial model parameter values were informed by a variety of classified SUM measurements collected during the in-field, photo-observations validation process as well as from aggregated ambient air temperature measurements made by an air quality sampler in 62 study household kitchen areas over the same time period (Coffey et al., 2019). A detailed description of each parameter and suggested specifications can be found in SI Section 2. In brief, the 99.9th percentile of ambient temperatures (Fig. S1) and ambient temperature instantaneous slopes (Fig. S2) were used as initial values for findpeaks parameters

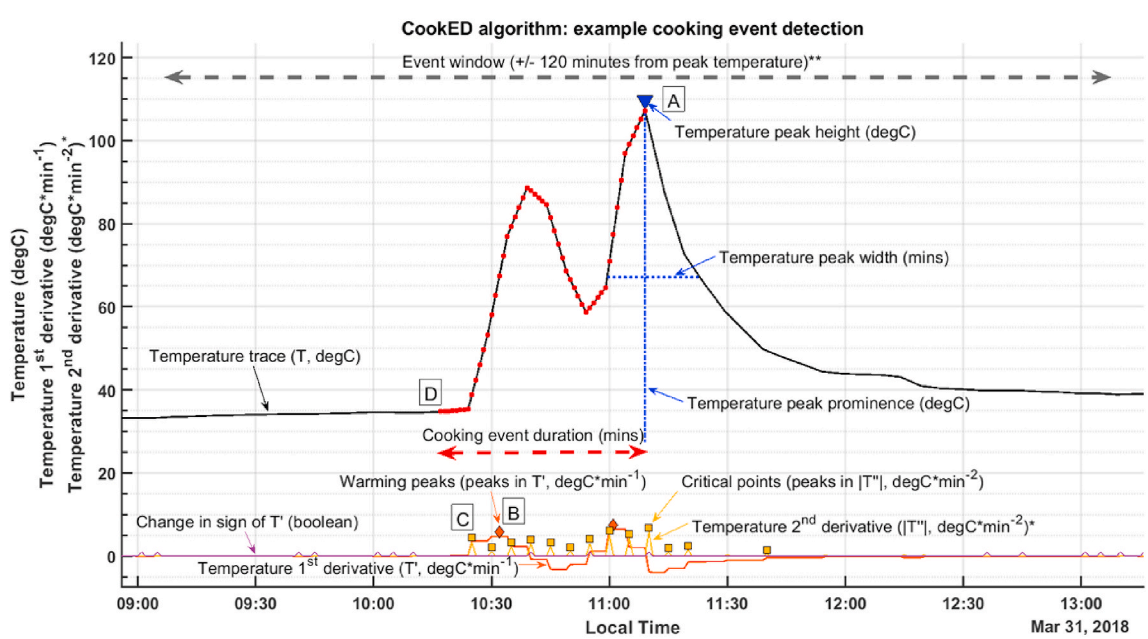

Fig. 6. Example SUM temperature data timeseries depicting the key parameters of the CookED. The temperature peak ("A", blue triangle) is identified using peak height (blue triangle), peak width (horizontal dotted blue line) and peak prominence (vertical dashed blue line) thresholds. Within the event search window (grey dashed line), peaks (orange diamonds) in the temperature slope (orange, T'), peaks (yellow squares) in the change in temperature slopes (yellow, |T"|) and sign changes in temperature slopes (purple) are used to identify event start ("D") and end ("A") times. In this case, the time corresponding to the initial critical point ("C") prior to the first warming peak ("B") is adjusted for stove thermal lag to estimate the event start time ("D"). The end of the event ("A") is identified as the time of the temperature peak itself as no subsequent warming peaks were detected. Minutes corresponding to stove usage are marked with red circles resulting in an estimated cooking duration (red dashed line).

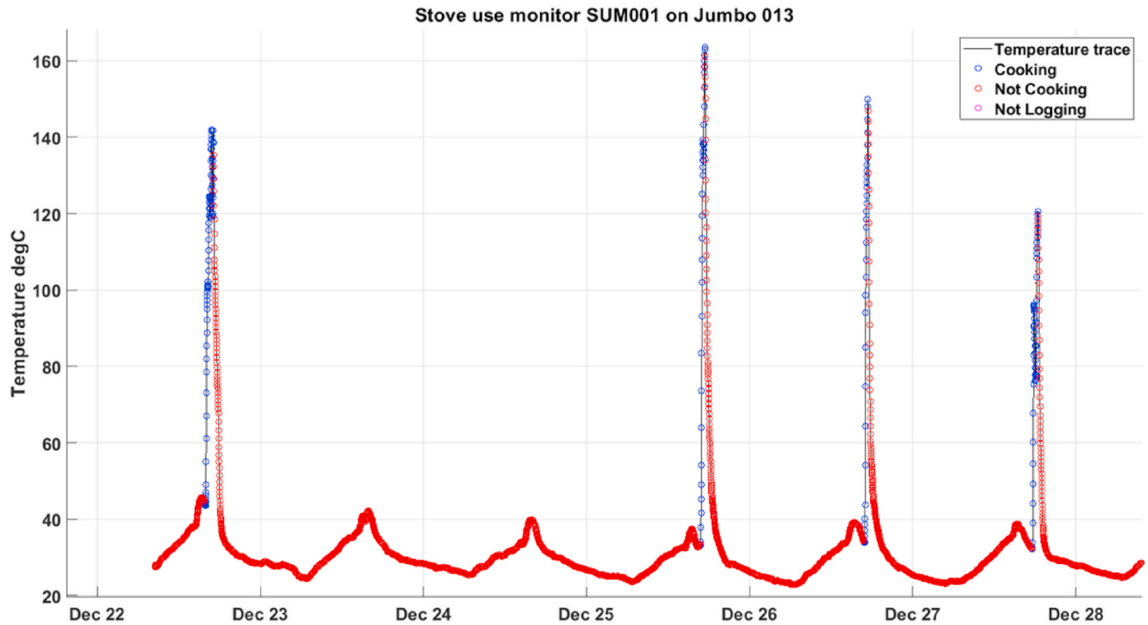

Fig. 7. An automatically generated CookED-classified SUM temperature trace from a Jumbo rocket stove indicating cooking events (blue) and non-cooking (red) spanning one week.

Minpeakheight (50C) and *Minpeakheight_prime* (0.5C/min), respectively (SI). To account for narrow peaks associated with short cooking events, conservative values for *Minpeakwidth* were chosen ranging from 3 to 8 min, again depending on the stove type. Stoves with SUMs experiencing rapid heating and cooling and/or stoves used for short cooking tasks (e.g., LPG stoves boiling water for tea) can have very short-term spikes in SUM temperature readings and therefore short peak widths. Accounting for the wide range in SUM temperature readings when stoves were not used and peak temperatures associated with cooking, the *minpeakprominence* parameter identifies the minimum temperature prominence of spikes above background levels which ranged between 22 and 35 degrees C. Additional temperature slope (T') and change in temperature slope (T") parameters (i.e., minpeakprominence_prime/prime2, minpeakdist_prime/prime2, minpeakwidth_prime/prime2, minpeakheight_prime2) were manually selected based on visual inspection of where and how many resulting peaks were identified. To be clear, these parameters - associated with the first and second derivative peaks – are primarily used to delineate the start and end of an event rather than detect the event itself (that is the primary objective of the temperature trace peak detection). An in-depth sensitivity analysis of model parameters or automated feature selection was not performed in this scope of work and could result in improvements.

CookED was applied to an estimated 167 million SUM minute-level values from data collected in Ghana from September 2017 through March 2019. These data comprised an estimated 116,550 stove-monitoring-days across 104 study households. Stove use classifications from 2-burner LPG stoves incorporated both SUM classifications indicating overall stove usage if either of the SUMs reported use at each time point.

2.5. Comparison to published event detection algorithms

Ruiz-Mercado et al. introduced the first event detection algorithm for cookstoves in the late 2000s (Ruiz-Mercado et al., 2008). Briefly, the Ruiz-Mercado (RM) algorithm identifies usage events by identifying locations along a SUM temperature trace where specific criteria are met signifying the minimum unit of stove use, a fueling event. Specifically, the algorithm identifies instances where a peak in temperature exceeds a minimum temperature threshold and slopes on either side of the peak temperature exceed specific slope thresholds. The temperature threshold is determined from the distribution of daily SUM temperature readings on reference days, or days in which cooking was not predicted. Slope thresholds are determined from the distribution of reference day temperature slopes. The RM method prescribes the 1st and 99th percentiles as the low and high slope thresholds, respectively. RM algorithm parameters are then refined through a recursive process. Fueling events clustered within a user-defined period of time (e.g., derived from surveys) are combined into one cooking event. A complete description of the RM method is provided elsewhere (Ruiz-Mercado et al., 2012).

Three different specifications of the RM method were deployed on the SUM data collected in this study for comparison purposes. Results from three separate outputs of the RM method are reported, each with slightly different initial parameters. The first RM specification uses the original, prescribed 1st and 99th percentiles of ambient temperature slopes as well as a default 2hr meal window (for clustered events). The second specification uses the minimum and maximum ambient temperature slopes and a 2hr meal window. The third uses the minimum and maximum ambient temperature slopes and average meal durations determined from the photo observations for each stove type.

In addition to the RM method, the Threshold and FireFinder event detectors available through SUMSarizer (R package and subsumed under Geocene Studies) were assessed. These options represent, perhaps, the most widely used event detection methods at the time of publication. The Threshold detector, representing the simplest option, classifies cooking based solely on a threshold temperature parameter and a direction qualifier (e.g., >45C). The authors of this method acknowledge this option as a poorly performing cooking event detector suggesting it instead be applied to detect and assess the rate of broken sensors, yet its performance has never been formally assessed against in-field observations. FireFinder, a more advanced deterministic method, involves tunable parameters to distinguish data associated with cooking events from non-cooking data. The tunable FireFinder parameters in the open-sourced SUMSarizer R package (available through GitHub) describe the temperature threshold above which cooking is likely to occur, the minimum duration of a cooking event as well as the minimum duration between distinct, consecutive events. FireFinder first assumes all points are not cooking. Next, it assumes all temperatures above the primary temperature threshold are cooking. Of these data, collections of points with negative slopes lasting an hour or more are classified as not cooking. Intuitively, the detector assumes points with positive slopes are related to cooking and negative slopes are related to not cooking. Events shorter than the user-defined minimum event length are removed (Wilson et al., 2020). The authors of Threshold and FireFinder methods suggest selecting parameter values and iteratively editing them based on expert intuition or known usage classifications with accompanying temperature trace data. The stove-type-specific parameters used for evaluation purposes here are displayed in Table 1. We arrived at these values following the selection suggestions (e.g., iterating) from SUMSarizer's authors but these parameters represent a single, thoughtful performance benchmark and not necessarily the optimum. SUMSarizer also offers users the ability to develop custom algorithms by training models using machine learning ensembles. This promising approach's authors suggest sufficient data be used to train (on the order of 5% or 25 files, whichever is larger) and subsequently validate the models. Five, independent models (one for each stove type) were trained and evaluated. However, due to a lack of training and validation data for each stove type and to avoid misrepresentation, the results from this approach were not included in the final analysis.

All three RM specifications as well as the Threshold and FireFinder algorithms benefited from the same raw data prefiltering described above as part of CookED. This allowed direct comparisons of the models from a refined state. Model performance was assessed at the minute level by means of predictive accuracies (positive and negative predictions). In addition, estimated mean event durations for each stove type, aggregated at the household level, were compared to observed durations to reflect overall accuracy in estimating cooking durations. Bias and error in estimated durations were investigated.

2.6. Ethical considerations

We are aware of the ethical considerations pertaining to the data collection protocols described. The study and all procedures performed in the study that involved human participants were approved by the Navrongo Health Research Centre Institutional Review Board and University of Colorado, Boulder Institutional Review Board. Informed consent was obtained from all individual participants included in the study and participants had the option to have data deleted upon request at the end of each deployment visit. All data collected were stored on password protected servers and reported in aggregate or de-identified.

3. Results and discussion

3.1. In-field, stove usage photo-observations

A summary of in-field, stove photo-validation observations is reported in Table 2. All 18 households invited to participate in the photo-measurements agreed to participate: 10 from the P3 Gas arm and 8 from the P3 Bio arm. From these households, a combined 43,029 photo-

Table 1
Parameter specification for SUMSarizer's Threshold and FireFinder event detection models (see (Wilson et al., 2020)).

Stove type	Detector model	(primary) temperature threshold (degC)	Min_event_sec	Min_break_sec
LPG	Threshold	55	–	–
	FireFinder	55	600 (10min)	1800 (30min)
Jumbo	Threshold	54	–	–
	FireFinder	54	600 (10min)	1800 (30min)
ACE	Threshold	56.7	–	–
	FireFinder	56.7	600 (10min)	1800 (30min)
TSF	Threshold	54	–	–
	FireFinder	54	600 (10min)	1800 (30min)
Coalpot	Threshold	50	–	–
	FireFinder	50	600 (10min)	1800 (30min)

Table 2
A summary of the stove usage photo-observation sampling by study arm and stove type.

	P3 Bio arm (rural)			P3 Gas arm (urban)	
Unique households visited	8			10	
Sample days	16			20	
Total # photo-observations	17,015			26,014	
	TSF	**Coalpot** charcoal stove	**ACE1** semi-gasifier	**Jumbo** rocket stove	**LPG**
Unique stoves observed	15	13	3	5	4
Cooking events observed	11	15	4	7	4
Minutes of overlap between photo-observations and valid SUMS measurements	15,188	15,390	6285	8110	4062

observations were taken across both camera systems. A total of 49,035 min of valid comparison data between SUM temperature measurements and classified photo images were collected; the data covered 40 unique cookstoves across 36 days. Placing cameras in both the primary and secondary cooking areas was vital to capturing the maximum number of validation observations as cooks often moved their portable stoves (e.g., ACE, Jumbo and coalpots) from one cooking area to another to operate. Another benefit of photo-observations is that a single image can document the usage of multiple stoves.

3.2. Performance of CookED

3.2.1. Event detection and duration estimation

Overall performance of CookED was first characterized by the fact that out of the 41 cooking events observed in the field, all but one event were successfully detected by CookED. The single event which was not identified, was on a coalpot stove and further investigation into the SUM temperature trace showed that this was the result of combining 2 observed events into one due to the events being clustered within 30 min. No single event was predicted which was not observed in the field. Household comparisons of mean observed and estimated event durations for each stove type monitored are illustrated in Fig. 8. Stove use duration data were aggregated and averaged at the household level, for each stove type, to showcase variation amongst households (disaggregated results can be found in SI, Fig. S3). Differences between stove-type-aggregated observed mean event durations and estimated durations were not significant at the 5% level and overall correlation (Pearson's r) between estimated and observed duration was 0.85. Predicted event duration bias and error were assessed for the 5 stove types. Estimated mean duration was positively biased (5–37%) deviating from mean observed durations by 4–28 min, varying by stove type (Table S2).

3.2.2. Minute-level stove usage classification

A performance summary of the model predicting minute-level stove use is presented in Table 3. The sensitivity, or the proportion of minute-to-minute positives that are correctly identified, ranged from 78.7% for ACE stoves to 92.4% for LPG stoves. Selectivity, or the proportion of negatives that are correctly identified, was high, ranging from 96.5% to 99.2% driven, in large part, by the substantial fraction of observed minutes without cooking. The predictive accuracy, shown in Table 3, combines selectivity and sensitivity to describe the proportion of total samples accurately classified. Values of predictive accuracy for CookED ranged from 95.8% to 98.7%, again dominated by the high occurrence of true negatives. Matthew's correlation coefficient (MCC), is a useful

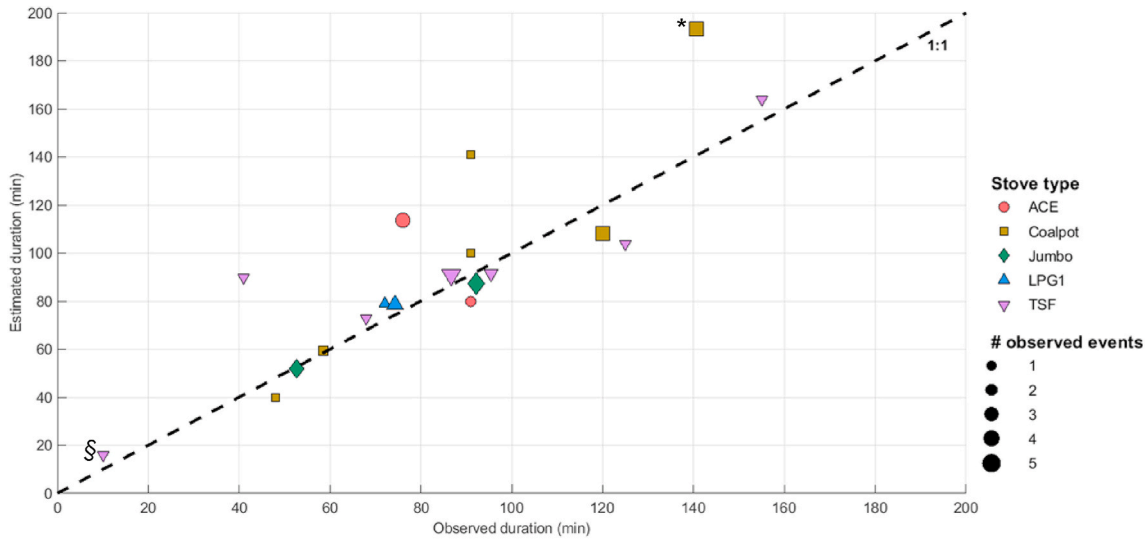

Fig. 8. Comparison of CookED's mean estimated cooking duration to mean observed cooking duration (Pearson's r = 0.85) aggregated at the household level for each stove type (LPG1 signifies a 1-burner gas stove). Each marker represents a study household's mean cooking duration, in minutes, for a given stove type (marker color and shape). The size of the marker is proportional to the number of observed events. § a partial event was observed lasting only 10 min * Over-predicted average duration at this household was a result of incorrectly combining two events into one.

Table 3
Minute-level stove usage predictive performance of CookED as percentages.

Minute Classification Performance	TSF	Coalpot charcoal stove	ACE1 semi-gasifier	Jumbo rocket stove	LPG
Sensitivity	84.7	83.7	78.7	81.8	92.4
Selectivity	98.9	97.2	96.5	99.2	99.2
Positive predictive value	84.9	90.0	73.5	91.7	90.0
Negative predictive value	98.2	97.2	98.4	98.1	99.4
Overall predictive accuracy	97.9	96.9	95.8	98.4	98.7
Matthew's Correlation Coefficient (MCC)	86.1	88.3	75.8	90.3	90.4

diagnostic for evaluating binary classification models with the main advantage of providing a more comprehensive evaluation than overall predictive accuracy, especially when class samples are unbalanced as they are here (Chicco and Jurman, 2020). Values of MCC can range from −1 to 1; 1 being a perfect classifier and 0 being the expected value for a coin toss classifier. The stove-type-specific MCCs ranged from 0.758 for ACE stoves to 0.904 for LPG stoves.

To further explore where (or more aptly when) misclassifications along a SUM temperature trace occurred, plots of SUM temperature colored by the outcome classification/misclassification are presented. In Fig. 9, a ~45hr temperature trace for a Jumbo stove is presented with minute-level classification/misclassification categories. A total of four events were observed on this stove over this period, each with unique temperature patterns and overall duration. The first event was marked by a slightly early estimated start time and subsequent early stop time resulting in false positives (predicting cooking when no cooking is

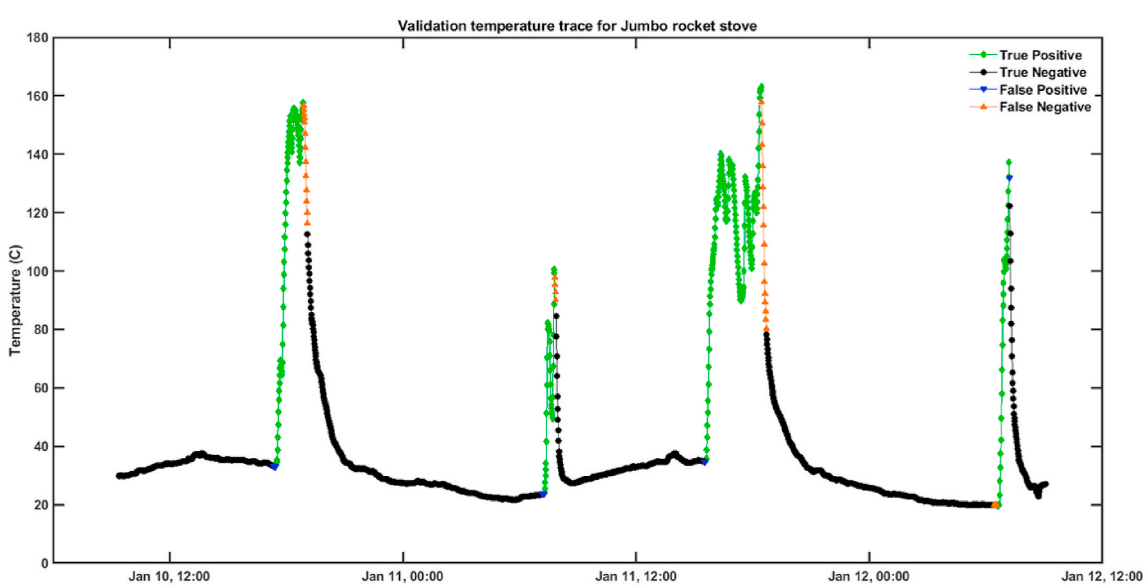

Fig. 9. CookED minute SUM temperature trace values from a Jumbo rocket stove spanning 2 days are colored by mis/classifications using in-field photo-observations as the ground truth.

observed, blue triangles) and false negatives (predicting no cooking when cooking is observed, orange triangles) at the head and tail of the peak, respectively. Active stove use was observed in the tail following the peak as the stove cooled down but was not well captured by CookED. Now, compare this event to the second where the estimated cooking ended within a few minutes of observed cooking just past the peak temperature. The third event is similar to the first with multiple, less prominent temperature peaks, mid-event, and a tail of false negatives following the estimated end. The fourth event has a nearly 20-min period of observed cooking before the temperatures begin to rise; half of which were estimated as cooking. This may be the result of a cook attempting to light her Jumbo stove which can be more challenging and/or time consuming than lighting a TSF. Over-fueling a rocket style stove like the Jumbo can present ignition challenges as a sufficient natural draft is required for the stove to continue operating. White smoke seen in the photo-observations suggested the fire was poorly lit requiring the cook to tend more to the stove while the SUM temperature measurements indicated no heating during this period and therefore no predicted use. Ignition of solid fuels is associated with markedly high pollutant emissions (Fedak et al., 2018) (i.e., particulate matter) and accurate classification of stove usage during this dirty activity can play a crucial role in understanding exposure outcomes linked to specific cooking behaviors.

Fig. 10 depicts a 48hr coalpot SUM temperature trace with the same prediction classification indicators. Again, there were varying degrees of accuracy in predicting event end points – resulting in mostly false negatives. However, in the third event, there were a series of false positives demonstrating the challenges of accurately determining the conclusion of an event. On the other hand, the beginning of each event was accurately estimated within 2–13 min with most predicted start times occurring slightly after observed start times.

3.3. Comparison to published SUM algorithms

Results of applying the three RM method specifications to the SUM data are shown alongside those from CookED in Table 4. Overall, CookED more accurately identified individual events as well as start and end times than the RM specifications. Gradual improvements in predictive accuracy for each stove type are made between the RM methods when accounting for 1) minimum and maximum ambient temperature slopes (RM2) over 1st and 99th percentiles (RM1) as well as 2) specifying stove-type-specific meal duration values (RM3) for clustering purposes over default values of 2 h (RM1 and RM2). More revealing are the differences in MCC and event prediction accuracy across the different RM model specifications. Changing the temperature slope parameters - for both heating and cooling - from RM1's 99th/1st percentile of ambient temperature slopes to RM2's maximum and minimum resulted in substantial predictive improvement ($MCCs_{RM2}$:0.32–0.76 compared to $MCCs_{RM1}$:0.26–0.54). Note that this improvement was less drastic for LPG stoves as they are typically located inside the home, out of direct sunlight and often not positioned directly next to other stoves reducing the likelihood of capturing fugitive heat from other sources. Interestingly, MCC fell for the ACE from RM1 to RM2 despite improved accuracy. This is the result of the higher magnitude slope thresholds in RM2 detecting fewer of the observed events reducing the occurrence of false positives ultimately improving overall accuracy. In other words, although RM1 and RM2 both incorrectly predicted 2 ACE events, RM2 accurately classified more individual minutes than RM1. Granted the RM slope threshold values originated from a distribution of ambient air temperature measurements rather than a distribution of non-cooking SUM measurements more prone to the erroneous effects of sunlight, the difference in performance of these two model specifications demonstrates how sensitive results can be to slight changes in parameter values highlighting the value in-field, ground-truth observations can offer towards optimizing these features.

All three RM specifications consistently overpredicted the duration of events exemplified most by RM1 and RM2 utilizing a default 2hr event-width parameter (Figs. S4–6). The number of events predicted compared to the number of events observed (Table 3) can be quite drastic resulting in large discrepancies between predicted and observed average event duration (Tables S4–6). Moreover, the ratio of the number of estimated cooking minutes to observed cooking minutes exceeded 1 in all scenarios except for RM1 and RM2 applied to ACE stoves. These biases should be considered carefully when extrapolating event-wise durations across multiple stoves and households to study wide summaries of usage.

SUMsarizer's Threshold method performed similarly to RM method specifications with varied performance between stove types (Table 4). As for most algorithms, the Threshold method's minute-level classification accuracy was higher for improved stoves (LPG, Jumbo and ACE) than their traditional counterparts (TSF and Coalpots). Variability among traditional stove characteristics, household culinary practices

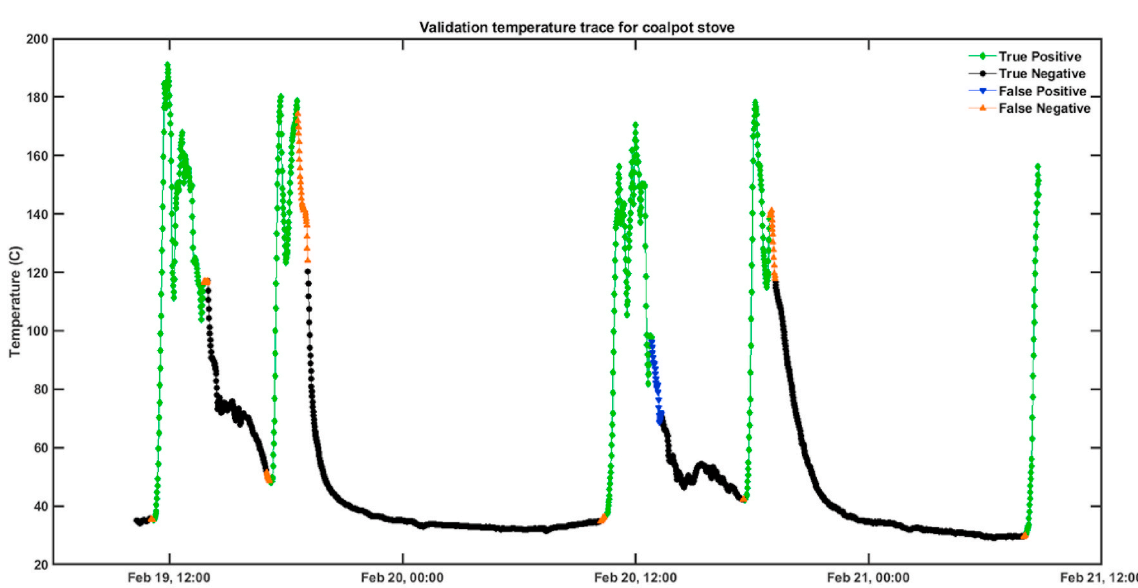

Fig. 10. CookED minute SUM temperature trace values from a coalpot (charcoal) stove colored by mis/classifications using in-field photo-observations as the ground truth.

Table 4
Summary of performance diagnostics for CookED, SUMSarizer's FireFinder and Threshold detectors and three specifications of the Ruiz-Mercado (RM) algorithm.

			TSF		Coalpot		ACE		Jumbo		LPG	
CookED	% min true	Events est./obs.	97.8%	11/11	96.3%	14/15	95.6%	4/4	97.8%	7/7	98.4%	4/4
	% min false	Cooking min est./obs.	2.2%	0.93	3.7%	0.99	4.4%	1.32	2.2%	0.91	1.6%	1.03
	MCC		0.85		0.86		0.73		0.88		0.88	
SUMSarizer FireFinder	% min true	Events est./obs.	82.9%	[a]13/11	92.8%	[a]19/15	95.5%	4/4	98.1%	7/7	98.3%	4/4
	% min false		17.1%		7.2%		4.5%		1.9%		1.7%	
	MCC	Cooking min est./obs.	0.65	3.03	0.76	1.32	0.65	1.13	0.88	0.77	0.88	0.95
SUMSarizer Threshold	% min true	Events est./obs.	68.8%	[a]11/11	88.5%	[a]24/15	92.3%	4/4	96.0%	8/7	96.8%	4/4
	% min false		31.2%		11.5%		7.7%		4.0%		3.2%	
	MCC	Cooking min est./obs.	0.42	5.15	0.67	1.79	0.52	1.93	0.79	1.14	0.79	1.20
RM 3	% min true	Events est./obs.	93.3%	[a]10/11	92.6%	15/15	96.0%	[a]2/4	95.8%	7/7	94.2%	4/4
	% min false		6.7%		7.4%		4.0%		4.2%		5.8%	
	MCC	Cooking min est./obs.	0.61	1.31	0.77	1.63	0.35	0.54	0.82	1.56	0.72	1.80
RM 2	% min true	Events est./obs.	91.6%	[a]10/11	91.6%	15/15	95.0%	[a]2/4	91.6%	7/7	89.3%	4/4
	% min false		8.4%		8.4%		5.0%		8.4%		10.7%	
	MCC	Cooking min est./obs.	0.58	1.65	0.76	1.73	0.32	0.82	0.74	2.29	0.60	2.48
RM 1	% min true	Events est./obs.	70.6%	[a]20/11	78.9%	[a]13/15	85.5%	6/4	79.3%	[a]13/7	86.5%	[a]6/4
	% min false		29.4%		21.1%		14.5%		20.7%		13.5%	
	MCC	Cooking min est./obs.	0.26	4.99	0.50	2.97	0.49	3.86	0.29	4.19	0.54	2.86

RM: Ruiz-Mercado SUM event-detection algorithm, RM1 is the original (Ruiz-Mercado et al., 2012) and RM2 and RM3 are adaptations.
MCC: Mathew's correlation coefficient (see Chicco and Jurman, 2020).
"min" refers to minutes, "est.", estimated by algorithm; "obs.", from in-field photo-observations.
[a] Occurrences of event false positives/negatives found are not reflected in accuracy yet are reflected in MCC.

preferentially carried out on traditional stoves and less-controlled SUM placement likely account for some of these discrepancies. Threshold-estimated event durations tended to be considerably longer than those observed even after considering the higher number of predicted events, especially for TSF and coalpots (Table S5 and Fig. S7). This resulted in the ratio of the number of estimated cooking minutes to observed cooking minutes to regularly exceed unity (1.14–5.15 depending on stove type, Table 4). As suggested by the algorithm's authors, the Threshold algorithm appears to be more useful in quantifying and ascertaining SUM data quality than detecting and accurately estimating cooking events across a diverse array of stove types.

FireFinder, in contrast, had much higher performance than the Threshold method, accurately identifying all events observed on improved stoves whilst simultaneously avoiding event false positives on improved stoves. FireFinder and CookED performed similarly for improved stoves, predicting minute-level usage on ACE, Jumbo and LPG stoves with high accuracy (>95%) within 0.3 percentage points of one another. Comparing MCCs, CookED had markedly higher performance than FireFinder for ACE stoves (0.73 vs. 0.65) and equivalent values for LPG and Jumbo stoves (0.88). Differences in performance between CookED and FireFinder for these improved stoves manifest in slight differences in predicting event start and end times accentuated by the ratios of predicted to observed cooking minutes.

However, classifying the use of traditional stoves proved more challenging for FireFinder. It predicted several events on traditional stoves which were not observed resulting in MCCs of 0.65 and 0.76 for TSF and coalpots, respectively (Table 4). FireFinder accuracy and MCCs for traditional stoves were similar to those from the RM3 specification. That said, FireFinder had some of the smallest biases in estimating mean event durations (Table S6 and Fig. S8) while CookED had some of the overall lowest errors (Table S1).

The versatility and complexity of these algorithms vary greatly. For example, the Threshold method involves two parameters (one being a qualifier) compared to FireFinder's three, the RM method's four and CookED's sixteen. Tradeoffs between methods manifest among complexity, accuracy, and bias. Moreover, these tradeoffs occur across different levels of analysis. For example, event quantity (# events) and duration (length of event) are required to estimate total cooking time for any given stove. Take an algorithm that estimates the duration of detected events perfectly yet fails to detect half of the observed events. Such an algorithm will yield inaccurate total cooking time for that stove which may influence intermediate study outcomes. Computational time to execute these methods are similar with data prefiltering comprising the majority of executable time. Needless to say, models with more input parameters often require more tuning time. The largest differences amongst the methods in tuning capabilities (and thus time) seem to originate in the extent to which start and end times of events are estimated (rather in identifying an event). CookED offers greater adjustability in this domain but only at the discretion of the user. Furthermore, observation-based, ground-truth stove usage information paired with SUM measurements may be utilized to calculate or tune parameters towards optimizing specific usage indicators with the prospects of mitigating some of the burden associated with implementing complex and demanding algorithms. For example, CookED's thermal lag parameter can be tuned to optimize event-wise duration estimation, perhaps reducing duration bias, when research or evaluative priorities focus on cooking duration. Yet, where explicit research questions require upmost predictive accuracy and over short timescales – for example, linking complex cooking behaviors (user's proximity to a stove, fuel ignition patterns and active fire tending) to acute exposure outcomes – algorithm specificity and sensitivity become essential.

3.4. Generalizability and accessibility

CookED was developed to classify minute temperature data from five stove types used in the P3 study region of Northern Ghana into cooking or not-cooking, encompassing a wide variety of stove construction types, environments of stove usage (e.g., indoors/outdoors, day/night) and unique household cooking behaviors. These differences are likely to translate into substantial variation in temperature traces recorded by SUMs ultimately offering a more robust validation set for SUMS analytics: algorithm training, testing and evaluation. That being said, CookED's tunable parameters are also likely to vary across these contextual variables. Rather than be designed for a specific type of stove, CookED was created to capture expected trends found across numerous unique temperature traces indicative of residential cooking activities.

Efforts to translate CookED to an open-sourced software environment is underway.

3.5. Limitations and challenges

A challenge inherent to all stove usage monitoring, including photo-observational methods described here, is discerning usage from non-usage in real-world environments. These environments include cooks who frequently leave an active stove unattended to address other household needs, or utilize the residual heat from a diminishing fire after having prepared a meal. In fact, many cooks in this study were observed to have extinguished the coals of the fire after having prepared a meal to conserve fuel - creating a plume of steam - while others appeared to let the coals burn out. These two behaviors can result in large differences in observed event duration and require a consistent set of classification protocols or clear explanations of what constitutes 'usage' in a study context. Another limitation of the photo-observation method is representative sampling. Kitchen characteristics can vary widely between households in a single village, and some urban kitchens were reported as too confined for valid photo-observations. Using cameras with wider fields of view (e.g., 'fish-eye lens') could reduce these occurrences and provide more representative sampling.

In addition to uncertainty classifying photo-observations, the number of in-field cooking observations was limited for some stove types. For example, both the ACE and LPG stove types were only observed being operated 4 times in the field, limiting the extent of validation. This was in part due to equipment malfunctions (camera battery failures) resulting in incomplete observation periods as well as general, infrequent use of ACE stoves. Stove types operated with more than one fuel type (e.g., firewood or charcoal in ACE) require a larger number of field observations as variation in SUM temperature traces are likely to result from the differences in operation between fuels. Although proving extremely valuable, the manual classification of photo-observations was time-intensive. This process initially took 40 min of classifying time per deployment day per stove, and by the end was reduced to nearly 15 min. Automated image processing/classification - or algorithms designed to identify cooking from photo-observations - are promising techniques which could expedite this process warranting future research.

On the issue of observation bias, our untested assumption is that the less-intrusive, motionless, and silent cameras used in this work are likely to reduce behavioral biases (Hawthorne effect) compared to other in-home stove usage monitoring methods. Additional research into this untested assumption is warranted.

3.6. Conclusion

In this work we laid the groundwork for a new, in-field cooking event observation method enabling the collection of near-continuous, ground-truth stove usage information during real-world situations. These data, in turn, can inform the development and evaluation of SUM event detection algorithms used in virtually all studies implementing SUMs. Photo-observations of cookstoves operated by end-users in the field offer perhaps the least intrusive and most realistic glimpse into real-world cooking patterns and behaviors. The benefits of this validation approach also extend beyond SUMs to assess other usage monitoring methods like cooking diaries and surveys granting practitioners yet another mechanism to substantiate findings. A global repository of in-field photo-observations paired with temperature traces from associated SUMs could advance SUMS analytics across studies, extend data transparency, and allow for more standardized approaches to installing, operating, processing and interpreting vast amounts of SUM data.

In addition, we introduced, described and evaluated a Cookstove Event Detection algorithm, CookED, designed to translate minute-level SUM temperature measurements into classifications of 'cooking' or 'not cooking' for 5 unique stove types with potential generalizability beyond the P3 study. Rather than replace existing tools, CookED can be added to current stove usage analytic toolkits. CookED demonstrated substantial improvement detecting and accurately estimating duration of cooking events over the implementation of three specifications of the 2012 Ruiz-Mercado algorithm. Performance of SUMSarizer's Threshold and FireFinder event detectors were also examined, highlighting variability in accuracy among stove types; traditional TSF and charcoal coalpots being the most challenging to characterize. CookED came out as a leading option. Further refinement of CookED's parameters through optimization and sensitivity analyses using classifications from photo-observations is likely to yield increased performance; a future endeavor to advance any SUM algorithm.

Current projections indicate considerable population growth in sub-Saharan Africa through the end of the century (United Nations, 2019). As a region relying heavily on solid fuels for household energy needs, a more comprehensive understanding of residential stove usage patterns is pivotal to combating the global public health challenge of human exposure to household air pollution and foremost, from the combustion of solid fuels. Insights into the linkages between stove usage, human behavior, exposure and ultimately health outcomes can be enhanced with more accurate sensors and accompanying analytics. Sensor devices are becoming ubiquitous within global development monitoring activities with expressed need in remote and power-constrained environments(Andres et al., 2018). Efforts exploring the effects of providing incentives (e.g. conditional cash transfers) to households using improved stoves or discontinuing use of traditional stoves currently rely on electronic SUMs to gauge stove usage and initiate incentives. The tools and methodology presented here can play an important role in substantiating the performance of these devices and expedite the technological and behavioral changes needed to curtail the effects of poor combustion of solid fuels in residential cookstoves.

Data availability statement

Some or all data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies. (https://osf.io/e24s3/).

Funding

This research was supported by grants from the US National Institutes of Health Clean Cooking Implementation Science Network and the National Science Foundation (SES 1528811).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

A warm thanks goes out to all the P3 study participants for their hospitality and patience. Thank you to all the field staff at the Navrongo Health Research Centre. Special thanks to friends, family and community for support and guidance. Publication of this article was funded by the University of Colorado Boulder Libraries Open Access Fund. This work was reviewed and approved by the Institutional Review Boards at the University of Colorado Boulder and the Navrongo Health Research Center, a part of the Ghana Health Service.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.deveng.2021.100065.

References

Andres, L., Boateng, K., Borja-Vega, C., Thomas, E., 2018. A review of in-situ and remote sensing technologies to monitor water and sanitation interventions. Water 10 (6), 756. https://doi.org/10.3390/w10060756.

Babatola, S.S., 2018. Global burden of diseases attributable to air pollution. J. Publ. Health Afr. 9 (3) https://doi.org/10.4081/jphia.2018.813.

Balmes, J.R., 2019. Household air pollution from domestic combustion of solid fuels and health. J. Allergy Clin. Immunol. 143 (6), 1979–1987. https://doi.org/10.1016/j.jaci.2019.04.016.

Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genom. 21 (1), 6. https://doi.org/10.1186/s12864-019-6413-7.

Coffey, E.R., Pfotenhauer, D., Mukherjee, A., Agao, D., Moro, A., Dalaba, M., Begay, T., Banacos, N., Oduro, A., Dickinson, K.L., Hannigan, M.P., 2019. Kitchen area air quality measurements in northern Ghana: evaluating the performance of a low-cost particulate sensor within a household energy study. Atmosphere 10 (7), 400. https://doi.org/10.3390/atmos10070400.

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., Dingenen, R. van, Donkelaar, A. van, Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. Lancet 389 (10082), 1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6.

Dalaba, M.; Coffey, E. R.; Brown, Z.; Alirigia, R.; Awaregya, J.; Aburiya, J. B. A.; Ali, M.; Agao, D.; Banacos, N.; Jones, M.; Hannigan, M. P.; Oduro, A.; Dickinson, K. L. Cleaner Cooking with Liquefied Petroleum Gas (LPG) in Northern Ghana: Interdisciplinary Experiment Measuring Willingness to Pay, Behavior Change, and Impacts on Personal Exposure. In preparation..

Dickinson, K.L., Kanyomse, E., Piedrahita, R., Coffey, E., Rivera, I.J., Adoctor, J., Alirigia, R., Muvandimwe, D., Dove, M., Dukic, V., Hayden, M.H., Diaz-Sanchez, D., Abisiba, A.V., Anaseba, D., Hagar, Y., Masson, N., Monaghan, A., Titiati, A., Steinhoff, D.F., Hsu, Y.-Y., Kaspar, R., Brooks, B., Hodgson, A., Hannigan, M., Oduro, A.R., Wiedinmyer, C., 2015. Research on emissions, air quality, climate, and cooking technologies in northern Ghana (reaccting): study rationale and protocol. BMC Publ. Health 15 (1), 126. https://doi.org/10.1186/s12889-015-1414-1.

Dickinson, K.L., Dalaba, M., Brown, Z.S., Alirigia, R., Coffey, E.R., Mesenbring, E., Achazanaga, M., Agao, D., Ali, M., Kanyomse, E., Awaregya, J., Adagenera, C.A., Aburiya, J.B.A., Gubilla, B., Oduro, A.R., Hannigan, M.P., 2018. Prices, peers, and perceptions (P3): study protocol for improved biomass cookstove project in northern Ghana. BMC Publ. Health 18 (1), 1209. https://doi.org/10.1186/s12889-018-6116-z.

Dickinson, K.L., Piedrahita, R., Coffey, E.R., Kanyomse, E., Alirigia, R., Molnar, T., Hagar, Y., Hannigan, M.P., Oduro, A.R., Wiedinmyer, C., 2019. Adoption of improved biomass stoves and stove/fuel stacking in the REACCTING intervention study in northern Ghana. Energy Pol. 130, 361–374. https://doi.org/10.1016/j.enpol.2018.12.007.

Fedak, K.M., Good, N., Dahlke, J., Hecobian, A., Sullivan, A., Zhou, Y., Peel, J.L., Volckens, J., 2018. Chemical composition and emissions factors for cookstove startup (ignition) materials. Environ. Sci. Technol. 52 (16), 9505–9513. https://doi.org/10.1021/acs.est.8b02218.

Graham, E.A., Patange, O., Lukac, M., Singh, L., Kar, A., Rehman, I.H., Ramanathan, N., 2014. Laboratory demonstration and field verification of a wireless cookstove sensing system (WiCS) for determining cooking duration and fuel consumption. Energy for Sustainable Development 23, 59–67. https://doi.org/10.1016/j.esd.2014.08.001.

Hollada, J., Williams, K.N., Miele, C.H., Danz, D., Harvey, S.A., Checkley, W., 2017. Perceptions of improved biomass and liquefied petroleum gas stoves in puno, Peru: implications for promoting sustained and exclusive adoption of clean cooking technologies. Int. J. Environ. Res. Publ. Health 14 (2). https://doi.org/10.3390/ijerph14020182.

Masera, O.R., Saatkamp, B.D., Kammen, D.M., 2000. From linear fuel switching to multiple cooking strategies: a critique and alternative to the energy ladder model. World Dev. 28 (12), 2083–2103. https://doi.org/10.1016/S0305-750X(00)00076-0.

Murray, C.J.L., Aravkin, A.Y., Zheng, P., Abbafati, C., Abbas, K.M., Abbasi-Kangevari, M., Abd-Allah, F., Abdelalim, A., Abdollahi, M., Abdollahpour, I., Abegaz, K.H., Abolhassani, H., Aboyans, V., Abreu, L.G., Abrigo, M.R.M., Abualhasan, A., Abu-Raddad, L.J., Abushouk, A.I., Adabi, M., Adekanmbi, V., Adeoye, A.M., Adetokunboh, O.O., Adham, D., Advani, S.M., Agarwal, G., Aghamir, S.M.K., Agrawal, A., Ahmad, T., Ahmadi, K., Ahmadi, M., Ahmadieh, H., Ahmed, M.B., Akalu, T.Y., Akinyemi, R.O., Akinyemiju, T., Akombi, B., Akunna, C.J., Alahdab, F., Al-Aly, Z., Alam, K., Alam, S., Alam, T., Alanezi, F.M., Alanzi, T.M., Alemu, B. wassihun, Alhabib, K.F., Ali, M., Ali, S., Alicandro, G., Alinia, C., Alipour, V., Alizade, H., Aljunid, S.M., Alla, F., Allebeck, P., Almasi-Hashiani, A., Al-Mekhlafi, H. M., Alonso, J., Altirkawi, K.A., Amini-Rarani, M., Amiri, F., Amugsi, D.A., Ancuceanu, R., Anderlini, D., Anderson, J.A., Andrei, C.L., Andrei, T., Angus, C., Anjomshoa, M., Ansari, F., Ansari-Moghaddam, A., Antonazzo, I.C., Antonio, C.A.T., Antony, C.M., Antriyandarti, E., Anvari, D., Anwer, R., Appiah, S.C.Y., Arabloo, J., Arab-Zozani, M., Ariani, F., Armoon, B., Ärnlöv, J., Arzani, A., Asadi-Aliabadi, M., Asadi-Pooya, A.A., Ashbaugh, C., Assmus, M., Atafar, Z., Atnafu, D.D., Atout, M.M. W., Ausloos, F., Ausloos, M., Quintanilla, B.P.A., Ayano, G., Ayanore, M.A., Azari, S., Azarian, G., Azene, Z.N., Badawi, A., Badiye, A.D., Bahrami, M.A., Bakhshaei, M.H., Bakhtiari, A., Bakkannavar, S.M., Baldasseroni, A., Ball, K., Ballew, S.H., Balzi, D., Banach, M., Banerjee, S.K., Bante, A.B., Baraki, A.G., Barker-Collo, S.L., Bärnighausen, T.W., Barrero, L.H., Barthelemy, C.M., Barua, L., Basu, S., Baune, B.T., Bayati, M., Becker, J.S., Bedi, N., Beghi, E., Béjot, Y., Bell, M.L., Bennitt, F.B., Bensenor, I.M., Berhe, K., Berman, A.E., Bhagavathula, A.S., Bhageerathy, R., Bhala, N., Bhandari, D., Bhattacharyya, K., Bhutta, Z.A., Bijani, A., Bikbov, B., Sayeed, M.S.B., Biondi, A., Birihane, B.M., Bisignano, C., Biswas, R.K., Bitew, H., Bohlouli, S., Bohluli, M., Boon-Dooley, A.S., Borges, G., Borzì, A.M., Borzouei, S., Bosetti, C., Boufous, S., Braithwaite, D., Breitborde, N.J.K., Breitner, S., Brenner, H., Briant, P.S., Briko, A.N., Briko, N.I., Britton, G.B., Bryazka, D., Bumgarner, B.R., Burkart, K., Burnett, R.T., Nagaraja, S.B., Butt, Z.A., Santos, F. L. C. dos, Cahill, L.E., Cámera, L.L.A., Campos-Nonato, I.R., Cárdenas, R., Carreras, G., Carrero, J.J., Carvalho, F., Castaldelli-Maia, J.M., Castañeda-Orjuela, C.A., Castelpietra, G., Castro, F., Causey, K., Cederroth, C.R., Cercy, K.M., Cerin, E., Chandan, J.S., Chang, K.-L., Charlson, F.J., Chattu, V.K., Chaturvedi, S., Cherbuin, N., Chimed-Ochir, O., Cho, D.Y., Choi, J.-Y.J., Christensen, H., Chu, D.-T., Chung, M.T., Chung, S.-C., Cicuttini, F.M., Ciobanu, L.G., Cirillo, M., Classen, T.K.D., Cohen, A.J., Compton, K., Cooper, O.R., Costa, V.M., Cousin, E., Cowden, R.G., Cross, D.H., Cruz, J.A., Dahlawi, S.M.A., Damasceno, A.A.M., Damiani, G., Dandona, L., Dandona, R., Dangel, W.J., Danielsson, A.-K., Dargan, P.I., Darwesh, A.M., Daryani, A., Das, J.K., Gupta, R.D., Neves, J. das, Dávila-Cervantes, C.A., Davitoiu, D.V., Leo, D.D., Degenhardt, L., DeLang, M., Dellavalle, R.P., Demeke, F. M., Demoz, G.T., Demsie, D.G., Denova-Gutiérrez, E., Dervenis, N., Dhungana, G.P., Dianatinasab, M., Silva, D. D. da, Diaz, D., Forooshani, Z.S.D., Djalalinia, S., Do, H. T., Dokova, K., Dorostkar, F., Doshmangir, L., Driscoll, T.R., Duncan, B.B., Duraes, A. R., Eagan, A.W., Edvardsson, D., Nahas, N.E., Sayed, I.E., Tantawi, M.E., Elbarazi, I., Elgendy, I.Y., El-Jaafary, S.I., Elyazar, I.R., Emmons-Bell, S., Erskine, H.E., Eskandarieh, S., Esmaeilnejad, S., Esteghamati, A., Estep, K., Etemadi, A., Etisso, A. E., Fanzo, J., Farahmand, M., Fareed, M., Faridnia, R., Farioli, A., Faro, A., Faruque, M., Farzadfar, F., Fattahi, N., Fazlzadeh, M., Feigin, V.L., Feldman, R., Fereshtehnejad, S.-M., Fernandes, E., Ferrara, G., Ferrari, A.J., Ferreira, M.L., Filip, I., Fischer, F., Fisher, J.L., Flor, L.S., Foigt, N.A., Folayan, M.O., Fomenkov, A. A., Force, L.M., Foroutan, M., Franklin, R.C., Freitas, M., Fu, W., Fukumoto, T., Furtado, J.M., Gad, M.M., Gakidou, E., Gallus, S., Garcia-Basteiro, A.L., Gardner, W. M., Geberemariyam, B.S., Gebreslassie, A.A.A.A., Geremew, A., Hayoon, A.G., Gething, P.W., Ghadimi, M., Ghadiri, K., Ghaffarifar, F., Ghafourifard, M., Ghamari, F., Ghashghaee, A., Ghiasvand, H., Ghith, N., Gholamian, A., Ghosh, R., Gill, P.S., Ginindza, T.G.G., Giussani, G., Gnedovskaya, E.V., Goharinezhad, S., Gopalani, S.V., Gorini, G., Goudarzi, H., Goulart, A.C., Greaves, F., Grivna, M., Grosso, G., Gubari, M.I.M., Gugnani, H.C., Guimarães, R.A., Guled, R.A., Guo, G., Guo, Y., Gupta, R., Gupta, T., Haddock, B., Hafezi-Nejad, N., Hafiz, A., Haj-Mirzaian, A., Haj-Mirzaian, A., Hall, B.J., Halvaei, I., Hamadeh, R.R., Hamidi, S., Hammer, M.S., Hankey, G.J., Haririan, H., Haro, J.M., Hasaballah, A.I., Hasan, M.M., Hasanpoor, E., Hashi, A., Hassanipour, S., Hassankhani, H., Havmoeller, R.J., Hay, S. I., Hayat, K., Heidari, G., Heidari-Soureshjani, R., Henrikson, H.J., Herbert, M.E., Herteliu, C., Heydarpour, F., Hird, T.R., Hoek, H.W., Holla, R., Hoogar, P., Hosgood, H.D., Hossain, N., Hosseini, M., Hosseinzadeh, M., Hostiuc, M., Hostiuc, S., Househ, M., Hsairi, M., Hsieh, V.C., Hu, G., Hu, K., Huda, T.M., Humayun, A., Huynh, C.K., Hwang, B.-F., Iannucci, V.C., Ibitoye, S.E., Ikeda, N., Ikuta, K.S., Ilesanmi, O.S., Ilic, I.M., Ilic, M.D., Inbaraj, L.R., Ippolito, H., Iqbal, U., Irvani, S.S.N., Irvine, C.M.S., Islam, M.M., Islam, S.M.S., Iso, H., Ivers, R.Q., Iwu, C.C.D., Iwu, C.J., Iyamu, I.O., Jaafari, J., Jacobsen, K.H., Jafari, H., Jafarinia, M., Jahani, M.A., Jakovljevic, M., Jalilian, F., James, S.L., Janjani, H., Javaheri, T., Javidnia, J., Jeemon, P., Jenabi, E., Jha, R.P., Jha, V., Ji, J.S., Johansson, L., John, O., John-Akinola, Y.O., Johnson, C.O., Jonas, J.B., Joukar, F., Jozwiak, J.J., Jürisson, M., Kabir, A., Kabir, Z., Kalani, H., Kalani, R., Kalankesh, L.R., Kalhor, R., Kanchan, T., Kapoor, N., Matin, B.K., Karch, A., Karim, M.A., Kassa, G.M., Katikireddi, S.V., Kayode, G.A., Karyani, A.K., Keiyoro, P.N., Keller, C., Kemmer, L., Kendrick, P.J., Khalid, N., Khammarnia, M., Khan, E.A., Khan, M., Khatab, K., Khater, M.M., Khatib, M.N., Khayamzadeh, M., Khazaei, S., Kieling, C., Kim, Y.J., Kimokoti, R.W., Kisa, A., Kisa, S., Kivimäki, M., Knibbs, L.D., Knudsen, A.K.S., Kocarnik, J.M., Kochhar, S., Kopec, J.A., Korshunov, V.A., Koul, P.A., Koyanagi, A., Kraemer, M.U. G., Krishan, K., Krohn, K.J., Kromhout, H., Defo, B.K., Kumar, G.A., Kumar, V., Kurmi, O.P., Kusuma, D., Vecchia, C.L., Lacey, B., Lal, D.K., Lalloo, R., Lallukka, T., Lami, F.H., Landires, I., Lang, J.J., Langan, S.M., Larsson, A.O., Lasrado, S., Lauriola, P., Lazarus, J.V., Lee, P.H., Lee, S.W.H., LeGrand, K.E., Leigh, J., Leonardi, M., Lescinsky, H., Leung, J., Levi, M., Li, S., Lim, L.-L., Linn, S., Liu, S., Liu, S., Liu, Y., Lo, J., Lopez, A.D., Lopez, J.C.F., Lopukhov, P.D., Lorkowski, S., Lotufo, P.A., Lu, A., Lugo, A., Maddison, E.R., Mahasha, P.W., Mahdavi, M.M., Mahmoudi, M., Majeed, A., Maleki, A., Maleki, S., Malekzadeh, R., Malta, D.C., Mamun, A.A., Manda, A.L., Manguerra, H., Mansour-Ghanaei, F., Mansouri, B., Mansournia, M.A., Herrera, A.M.M., Maravilla, J.C., Marks, A., Martin, R.V., Martini, S., Martins-Melo, F.R., Masaka, A., Masoumi, S.Z., Mathur, M.R., Matsushita, K., Maulik, P.K., McAlinden, C., McGrath, J.J., McKee, M., Mehndiratta, M.M., Mehri, F., Mehta, K.M., Memish, Z.A., Mendoza, W., Menezes, R. G., Mengesha, E.W., Mereke, A., Mereta, S.T., Meretoja, A., Meretoja, T.J., Mestrovic, T., Miazgowski, B., Miazgowski, T., Michalek, I.M., Miller, T.R., Mills, E. J., Mini, G.K., Miri, M., Mirica, A., Mirrakhimov, E.M., Mirzaei, H., Mirzaei, M., Mirzaei, R., Mirzaei-Alavijeh, M., Misganaw, A.T., Mithra, P., Moazen, B., Mohammad, D.K., Mohammad, Y., Mezerji, N.M.G., Mohammadian-Hafshejani, A., Mohammadifard, N., Mohammadpourhodki, R., Mohammed, A.S., Mohammed, H., Mohammed, J.A., Mohammed, S., Mokdad, A.H., Molokhia, M., Monasta, L., Mooney, M.D., Moradi, G., Moradi, M., Moradi-Lakeh, M., Moradzadeh, R., Moraga, P., Morawska, L., Morgado-da-Costa, J., Morrison, S.D., Mosapour, A., Mosser, J.F., Mouodi, S., Mousavi, S.M., Khaneghah, A.M., Mueller, U.O., Mukhopadhyay, S., Mullany, E.C., Musa, K.I., Muthupandian, S., Nabhan, A.F., Naderi, M., Nagarajan, A.J., Nagel, G., Naghavi, M., Naghshtabrizi, B., Naimzada, M. D., Najafi, F., Nangia, V., Nansseu, J.R., Naserbakht, M., Nayak, V.C., Negoi, I., Ngunjiri, J.W., Nguyen, C.T., Nguyen, H.L.T., Nguyen, M., Nigatu, Y.T., Nikbakhsh, R., Nixon, M.R., Nnaji, C.A., Nomura, S., Norrving, B., Noubiap, J.J., Nowak, C., Nunez-Samudio, V., Oţoiu, A., Oancea, B., Odell, C.M., Ogbo, F.A., Oh, I.-H., Okunga, E.W., Oladnabi, M., Olagunju, A.T., Olusanya, B.O., Olusanya, J.O., Omer, M.O., Ong, K.L., Onwujekwe, O.E., Orpana, H.M., Ortiz, A., Osarenotor, O., Osei, F.B., Ostroff, S.M., Otstavnov, N., Otstavnov, S.S., Øverland, S., Owolabi, M.O., M P, A., Padubidri, J.R., Palladino, R., Panda-Jonas, S., Pandey, A., Parry, C.D.H., Pasovic, M., Pasupula, D.K., Patel, S.K., Pathak, M., Patten, S.B., Patton, G.C., Toroudi, H.P., Peden, A.E., Pennini, A., Pepito, V.C.F., Peprah, E.K., Pereira, D.M., Pesudovs, K., Pham, H.Q., Phillips, M.R., Piccinelli, C., Pilz, T.M., Piradov, M.A., Pirsaheb, M., Plass, D., Polinder, S., Polkinghorne, K.R., Pond, C.D., Postma, M.J., Pourjafar, H., Pourmalek, F., Poznańska, A., Prada, S.I., Prakash, V., Pribadi, D.R.A., Pupillo, E., Syed, Z.Q., Rabiee, M., Rabiee, N., Radfar, A., Rafiee, A., Raggi, A., Rahman, M.A., Rajabpour-Sanati, A., Rajati, F., Rakovac, I., Ram, P., Ramezanzadeh, K., Ranabhat, C.L., Rao, P.C., Rao, S.J., Rashedi, V., Rathi, P., Rawaf, D.L., Rawaf, S., Rawal, L., Rawassizadeh, R., Rawat, R., Razo, C., Redford, S. B., Reiner, R.C., Reitsma, M.B., Remuzzi, G., Renjith, V., Renzaho, A.M.N., Resnikoff, S., Rezaei, N., Rezaei, N., Rezapour, A., Rhinehart, P.-A., Riahi, S.M., Ribeiro, D.C., Ribeiro, D., Rickard, J., Rivera, J.A., Roberts, N.L.S., Rodríguez-Ramírez, S., Roever, L., Ronfani, L., Room, R., Roshandel, G., Roth, G.A., Rothenbacher, D., Rubagotti, E., Rwegerera, G.M., Sabour, S., Sachdev, P.S., Saddik, B., Sadeghi, E., Sadeghi, M., Saeedi, R., Moghaddam, S.S., Safari, Y., Safi, S., Safiri, S., Sagar, R., Sahebkar, A., Sajadi, S.M., Salam, N., Salamati, P., Salem, H., Salem, M.R.R., Salimzadeh, H., Salman, O.M., Salomon, J.A., Samad, Z., Kafil, H.S., Sambala, E.Z., Samy, A.M., Sanabria, J., Sánchez-Pimienta, T.G., Santomauro, D.F., Santos, I.S., Santos, J.V., Santric-Milicevic, M.M., Saraswathy, S.Y.I., Sarmiento-Suárez, R., Sarrafzadegan, N., Sartorius, B., Sarveazad, A., Sathian, B., Sathish, T., Sattin, D., Saxena, S., Schaeffer, L.E., Schiavolin, S., Schlaich, M.P., Schmidt, M.I., Schutte, A.E., Schwebel, D.C., Schwendicke, F., Senbeta, A.M., Senthilkumaran, S., Sepanlou, S.G., Serdar, B., Serre, M.L., Shadid, J., Shafaat, O., Shahabi, S., Shaheen, A.A., Shaikh, M.A., Shalash, A.S., Shams-Beyranvand, M., Shamsizadeh, M., Sharafi, K., Sheikh, A., Sheikhtaheri, A., Shibuya, K., Shield, K.D., Shigematsu, M., Shin, J.I., Shin, M.-J., Shiri, R., Shirkoohi, R., Shuval, K., Siabani, S., Sierpinski, R., Sigfusdottir, I.D., Sigurvinsdottir, R., Silva, J.P., Simpson, K.E., Singh, J.A., Singh, P., Skiadaresi, E., Skou, S.T.S., Skryabin, V.Y., Smith, E.U.R., Soheili, A., Soltani, S., Soofi, M., Sorensen, R.J.D., Soriano, J.B., Sorrie, M.B., Soshnikov, S., Soyiri, I.N., Spencer, C.N., Spotin, A., Sreeramareddy, C.T., Srinivasan, V., Stanaway, J.D., Stein, C., Stein, D.J., Steiner, C., Stockfelt, L., Stokes, M.A., Straif, K., Stubbs, J.L., Sufiyan, M.B., Suleria, H.A.R., Abdulkader, R.S., Sulo, G., Sultan, I., Szumowski, Ł., Tabarés-Seisdedos, R., Tabb, K.M., Tabuchi, T., Taherkhani, A., Tajdini, M., Takahashi, K., Takala, J.S., Tamiru, A.T., Taveira, N., Tehrani-Banihashemi, A., Temsah, M.-H., Tesema, G.A., Tessema, Z.T., Thurston, G. D., Titova, M.V., Tohidinik, H.R., Tonelli, M., Topor-Madry, R., Topouzis, F., Torre, A.E., Touvier, M., Tovani-Palone, M.R.R., Tran, B.X., Travillian, R., Tsatsakis, A., Car, L.T., Tyrovolas, S., Uddin, R., Umeokonkwo, C.D., Unnikrishnan, B., Upadhyay, E., Vacante, M., Valdez, P.R., Donkelaar, A. van, Vasankari, T.J., Vasseghian, Y., Veisani, Y., Venketasubramanian, N., Violante, F.S., Vlassov, V., Vollset, S.E., Vos, T., Vukovic, R., Waheed, Y., Wallin, M.T., Wang, Y., Wang, Y.-P., Watson, A., Wei, J., Wei, M.Y.W., Weintraub, R.G., Weiss, J., Werdecker, A., West, J.J., Westerman, R., Whisnant, J.L., Whiteford, H.A., Wiens, K. E., Wolfe, C.D.A., Wozniak, S.S., Wu, A.-M., Wu, J., Hanson, S.W., Xu, G., Xu, R., Yadgir, S., Jabbari, S.H.Y., Yamagishi, K., Yaminfirooz, M., Yano, Y., Yaya, S., Yazdi-Feyzabadi, V., Yeheyis, T.Y., Yilgwan, C.S., Yilma, M.T., Yip, P., Yonemoto, N., Younis, M.Z., Younker, T.P., Yousefi, B., Yousefi, Z., Yousefinezhadi, T., Yousuf, A.Y., Yu, C., Yusefzadeh, H., Moghadam, T.Z., Zamani, M., Zamanian, M., Zandian, H., Zastrozhin, M.S., Zhang, Y., Zhang, Z.-J., Zhao, J.T., Zhao, X.-J.G., Zhao, Y., Zhou, M., Ziapour, A., Zimsen, S.R.M., Brauer, M., Afshin, A., Lim, S.S., 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. Lancet 396 (10258), 1223–1249. https://doi.org/10.1016/S0140-6736(20)30752-2.

Piedrahita, R., Dickinson, K.L., Kanyomse, E., Coffey, E., Alirigia, R., Hagar, Y., Rivera, I., Oduro, A., Dukic, V., Wiedinmyer, C., Hannigan, M., 2016. Assessment of cookstove stacking in northern Ghana using surveys and stove use monitors. Energy for Sustainable Development 34, 67–76. https://doi.org/10.1016/j.esd.2016.07.007.

Piedrahita, R., Coffey, E.R., Hagar, Y., Kanyomse, E., Verploeg, K., Wiedinmyer, C., Dickinson, K.L., Oduro, A., Hannigan, M.P., 2019. Attributing air pollutant exposure to emission sources with proximity sensing. Atmosphere 10 (7), 395. https://doi.org/10.3390/atmos10070395.

Piedrahita, R., Johnson, M., Bilsback, K.R., L'Orange, C., Kodros, J.K., Eilenberg, S.R., Naluwagga, A., Shan, M., Sambandam, S., Clark, M., Pierce, J.R., Balakrishnan, K., Robinson, A.L., Volckens, J., 2020. Comparing regional stove-usage patterns and using those patterns to model indoor air quality impacts. Indoor Air 30 (3), 521–533. https://doi.org/10.1111/ina.12645.

Pilishvili, T., Loo, J.D., Schrag, S., Stanistreet, D., Christensen, B., Yip, F., Nyagol, R., Quick, R., Sage, M., Bruce, N., 2016. Effectiveness of six improved cookstoves in reducing household air pollution and their acceptability in rural western Kenya. PloS One 11 (11). https://doi.org/10.1371/journal.pone.0165529.

Pillarisetti, A., Vaswani, M., Jack, D., Balakrishnan, K., Bates, M.N., Arora, N.K., Smith, K.R., 2014. Patterns of stove usage after introduction of an advanced cookstove: the long-term application of household sensors. Environ. Sci. Technol. 48 (24), 14525–14533. https://doi.org/10.1021/es504624c.

Pillarisetti, A., Allen, T., Ruiz-Mercado, I., Edwards, R., Chowdhury, Z., Garland, C., Hill, L.D., Johnson, M., Litton, C.D., Lam, N.L., Pennise, D., Smith, K.R., 2017. Small, smart, fast, and cheap: microchip-based sensors to estimate air pollution exposures in rural households. Sensors 17 (8), 1879. https://doi.org/10.3390/s17081879.

Pillarisetti, A., Gill, M., Allen, T., Madhavan, S., Dhongade, A., Ghorpade, M., Roy, S., Balakrishnan, K., Juvekar, S., Smith, K.R., 2018. A low-cost stove use monitor to enable conditional cash transfers. EcoHealth 15 (4), 768–776. https://doi.org/10.1007/s10393-018-1379-5.

Rhodes, E.L., Dreibelbis, R., Klasen, E., Naithani, N., Baliddawa, J., Menya, D., Khatry, S., Levy, S., Tielsch, J.M., Miranda, J.J., Kennedy, C., Checkley, W., 2014. Behavioral attitudes and preferences in cooking practices with traditional open-fire stoves in Peru, Nepal, and Kenya: implications for improved cookstove interventions. Int. J. Environ. Res. Publ. Health 11 (10), 10310–10326. https://doi.org/10.3390/ijerph111010310.

Ruiz-Mercado, I., Masera, O., 2015. Patterns of stove use in the context of fuel–device stacking: rationale and implications. EcoHealth 12 (1), 42–56. https://doi.org/10.1007/s10393-015-1009-4.

Ruiz-Mercado, I., Lam, N.L., Canuz, E., Davila, G., Smith, K.R., 2008. Low-cost temperature loggers as stove use monitors (SUMs). Boiling Point 55, 16–18.

Ruiz-Mercado, I., Canuz, E., Smith, K.R., 2012. Temperature dataloggers as stove use monitors (SUMs): field methods and signal analysis. Biomass Bioenergy 47, 459–468. https://doi.org/10.1016/j.biombioe.2012.09.003.

Ruiz-Mercado, I., Canuz, E., Walker, J.L., Smith, K.R., 2013. Quantitative metrics of stove adoption using stove use monitors (SUMs). Biomass Bioenergy 57, 136–148. https://doi.org/10.1016/j.biombioe.2013.07.002.

Shankar, A.V., Quinn, A.K., Dickinson, K.L., Williams, K.N., Masera, O., Charron, D., Jack, D., Hyman, J., Pillarisetti, A., Bailis, R., Kumar, P., Ruiz-Mercado, I., Rosenthal, J.P., 2020. Everybody stacks: lessons from household energy case studies to inform design principles for clean energy transitions. Energy Pol. 141, 111468. https://doi.org/10.1016/j.enpol.2020.111468.

Simons, A.M., Beltramo, T., Blalock, G., Levine, D.I., 2014. Comparing methods for signal analysis of temperature readings from stove use monitors. Biomass Bioenergy 70, 476–488. https://doi.org/10.1016/j.biombioe.2014.08.008.

Simons, A.M., Beltramo, T., Blalock, G., Levine, D.I., 2017. Using unobtrusive sensors to measure and minimize Hawthorne effects: evidence from cookstoves. J. Environ. Econ. Manag. 86, 68–80. https://doi.org/10.1016/j.jeem.2017.05.007.

Simons, A.M., Beltramo, T., Blalock, G., Levine, D.I., 2018. Sensor data to measure Hawthorne effects in cookstove evaluation. Data Brief 18, 1334–1339. https://doi.org/10.1016/j.dib.2018.04.021.

Smith, K.R., Bruce, N., Balakrishnan, K., Adair-Rohani, H., Balmes, J., Chafe, Z., Dherani, M., Hosgood, H.D., Mehta, S., Pope, D., Rehfuess, E., 2014. Millions dead: how do we know and what does it mean? Methods used in the comparative risk assessment of household air pollution. Annu. Rev. Publ. Health 35 (1), 185–206. https://doi.org/10.1146/annurev-publhealth-032013-182356.

Stanistreet, D., Hyseni, L., Bashin, M., Sadumah, I., Pope, D., Sage, M., Bruce, N., 2015. The role of mixed methods in improved cookstove research. J. Health Commun. 20 (Suppl. 1), 84–93. https://doi.org/10.1080/10810730.2014.999896.

Sundararaman, P., Gandhi, A., Frey, D., Ekolu, A.H., Ikalany, B., Hegde, M., Leith, K., Sweeney, D., 2016. Measuring usage and adoption of improved cookstoves in Ugandan households using quantitative and qualitative methods. In: 2016 IEEE Global Humanitarian Technology Conference (GHTC), pp. 136–142. https://doi.org/10.1109/GHTC.2016.7857271.

The World Bank. Household Cookstoves, Environment, Health and Climate Change: A New Look at an Old Problem https://openknowledge.worldbank.org/bitstream/handle/10986/27589/632170WP0House00Box0361508B0PUBLIC0.pdf?sequence=1&isAllowed=y (accessed Sep 30, 2020).

United Nations, 2019. World Population Prospects 2019: Highlights. https://population.un.org/wpp/Publications/Files/WPP2019_10KeyFindings.pdf. (Accessed 29 January 2021).

Williams, K.N., Kephart, J.L., Fandiño-Del-Rio, M., Condori, L., Koehler, K., Moulton, L. H., Checkley, W., Harvey, S.A., 2020. CHAP trial investigators. Beyond cost: exploring fuel choices and the socio-cultural dynamics of liquefied petroleum gas stove adoption in Peru. Energy Res Soc Sci 66. https://doi.org/10.1016/j.erss.2020.101591.

Wilson, D.L., Adam, M.I., Abbas, O., Coyle, J., Kirk, A., Rosa, J., Gadgil, A.J., 2015. Comparing cookstove usage measured with sensors versus cell phone-based surveys in darfur, Sudan. In: Hostettler, S., Hazboun, E., Bolay, J.-C. (Eds.), Technologies for Development. Springer International Publishing, Cham, pp. 211–221.

Wilson, D.L., Williams, K.N., Pillarisetti, A., 2020. An integrated sensor data logging, survey, and analytics platform for field research and its application in HAPIN, a multi-center household energy intervention trial. Sustainability 12 (5), 1805. https://doi.org/10.3390/su12051805.

World Health Organization, May 2, 2018. Ambient (outdoor) air pollution. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health. (Accessed 15 October 2020).

World Health Organization, May 8, 2018. Household air pollution and health. https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health. (Accessed 15 October 2020).