View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Strathclyde Institutional Repository

Blodgett, Courtney and Dauenhauer, Peter and Louie, Henry and Kickham, Lauren (2017) Accuracy of energy-use surveys in predicting rural mini-grid user consumption. Energy for Sustainable Development, 41. pp. 88-105. ISSN 0973-0826 , http://dx.doi.org/10.1016/j.esd.2017.08.002

This version is available at https://strathprints.strath.ac.uk/61818/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

The Strathprints institutional repository (https://strathprints.strath.ac.uk) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde's intellectual output.

Energy for Sustainable Development 41 (2017) 88–105

Contents lists available at ScienceDirect

Energy for Sustainable Development

Accuracy of energy-use surveys in predicting rural mini-grid user consumption

Courtney Blodgett[a], Peter Dauenhauer[b], Henry Louie[c,*], Lauren Kickham[a]

[a] Vulcan Inc., Seattle, WA, USA
[b] Institute for Energy & Environment, University of Strathclyde, Glasgow, United Kingdom
[c] Seattle University, Seattle, WA, USA

ARTICLE INFO

Article history:
Received 9 May 2017
Received in revised form 24 July 2017
Accepted 5 August 2017
Available online xxxx

Keywords:
Micro-grids
Mini-grids
Rural electrification
Energy estimation
Energy-use survey
Solar power

ABSTRACT

Mini-grids for rural electrification in developing countries are growing in popularity but are not yet widely deployed. A key barrier of mini-grid proliferation is the uncertainty in predicting customer electricity consumption, which adds financial risk. Energy-use surveys deployed in the pre-feasibility stage that capture present and aspirational consumption are intended to reduce this uncertainty. However, the general reliability and accuracy of these surveys has not been demonstrated. This research compares survey-predicted electrical energy use to actual measured consumption of customers of eight mini-grids in rural Kenya. A follow-up audit compares the aspirational inventory of appliances to the realized inventory. The analysis shows that the ability to accurately estimate past consumption based on survey or audit data, even in a relatively short time-horizon is prone to appreciable error — a mean absolute error of 426 Wh/day per customer on a mean consumption of 113 Wh/day per customer. An alternative data-driven proxy village approach, which uses average customer consumption from each mini-grid to predict consumption at other mini-grids, was more accurate and reduced the mean absolute error to 75 Wh/day per customer. Hourly load profiles were constructed to provide insight into potential causes of error and to suggest how the data provided in this work can be used in computer-aided mini-grid design programs.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Current access to modern energy services is neither universal nor guaranteed. The World Energy Outlook (Birol, 2016) estimates that 1.2 billion people do not have access to electricity. Projections by the IEA suggest that efforts to electrify Africa will slightly outpace population growth but still leave leave half a billion still without electricity by 2030 (Birol, 2015). Overall investment levels are expected to reach US$30 billion annually in electricity access alone. Within the range of solutions, renewables-based mini-grids are a promising option owing to their relatively low initial investment levels, scalability, and suitability for rural areas.

While of considerable interest as a marquee solution, the practicalities of implementation and operation of mini-grids is challenging. Often located in remote areas with difficult terrain and impoverished customers, the sustainability of mini-grids is far from guaranteed. Issues such as limited local technical and managerial skills, low energy demand, poor availability of supply components, and unproven financing models have been noted as some of the problems facing rural mini-grids (Azimoh et al., 2017).

Mini-grids are often designed to be financially viable once installed and many developers are hoping this will pave the way for commercial funding. Mini-grids are established on the principle of providing affordable energy for the rural population in their service area while balancing the need for an acceptable level of reliability and financial viability. Estimates place the costs of mini-grid supplied energy above that from the centralized grid, ranging from US$1.35/kWh to US$2.04/kWh compared to US$0.41/kWh to US$0.80/kWh for grid extension (Action, 2016), although in some scenarios mini-grids are cost-competitive (Nerini et al., 2016). For proponents of mini-grids in developing countries, it is essential that financial viability is a top priority.

The design of a mini-grid determines critical project parameters such as mix of renewables, component sizing, and network design. Clearly, these choices have a large impact on the overall financial model and determine the cost of energy required to make the mini-grid profitable. A common design approach is to use software such as HOMER (Bekele and Tadesse, 2012; Díaz et al., 2011; Kolhe et al., 2015; Olatomiwa et al., 2015; Rajbongshi et al., 2017). However,

* Corresponding author.
E-mail addresses: CourtneyB@vulcan.com (C. Blodgett), peter.dauenhauer@strath.ac.uk (P. Dauenhauer), hlouie@ieee.org (H. Louie), LaurenK@vulcan.com (L. Kickham).

http://dx.doi.org/10.1016/j.esd.2017.08.002
0973-0826/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

this approach requires accurate predictions of consumption profiles to yield meaningful designs. Recent efforts to produce consumption profiles for design purposes have used simulation techniques to generate more realistic consumption profiles. However, the efforts are still reliant on surveyed data as a primary input data (Boait et al., 2015; Mandelli et al., 2016b). While the advantage of the simulation techniques is an increased accuracy (mainly by the inclusion of stochastic elements) of certain aspects of load variation, the fundamental inaccuracy of the primary input data is not addressed.

Operational information for existing mini-grids is sparse, but is becoming more common due to the use of low-cost data acquisition systems (Blanchard and Little, 2016; Louie et al., 2016). Technical design data that have been published are rarely accompanied with an evaluation of the design suitability after installation.

Neither over- nor under-designing systems is desirable and it may lead to a lack of sustainability. Given that many of the users will gain access to electricity for the first time in their lives, predicting consumption based on past use is impossible. Alternatively, a widespread practice is to employ energy-use surveys of potential customers to predict consumption (GIZ, 2016).

The energy-use survey approach is common but receives little coverage in the literature. As a result it is unclear how effective it is at predicting consumption. This paper draws attention, in particular, to the accuracy of the energy-use survey. It is of utmost importance how a few survey questions are converted into a consumption prediction for mini-grid sizing and the implications of errors in this process.

Importance of energy prediction in mini-grid design

Although consumption prediction is a key element of system design, critical reviews that compare predicted to actual values are rare. Yet, recent research indicates that prediction errors can be considerable; in Louie and Dauenhauer (2016) an average error of 305% was reported.

Over-prediction of consumption leads to an over-sized system, which reduces financial viability. Financial viability is widely noted as a critical premise for successful projects (GIZ, 2016; Krithika and Palit, 2013; World Bank, 2012). At best, a short-term problem of low revenues may occur as customer ability to pay can have an upper ceiling. Substitution of electricity with alternative energy sources and low ability to pay limit the ability of the mini-grid owner to raise prices to cope with low revenue generation. Failure of load growth materializing over time represents a major investment risk and can sink the project financially. In Louie and Dauenhauer (2016), it was found that increasing system reliability from 99.7% to 99.9% (in terms of total energy served) drastically reduced the number of systems which could be deployed with an equal investment.

Under-prediction of consumption lowers reliability and availability of the system, potentially leading to serious issues that undermine the technical sustainability of the system, for example, reduction in battery lifespan (Government of New Zealand, 2010; IEEE Working Group for Energy Storage Subsystems, 2007a; IEEE Working Group for Energy Storage Subsystems, 2007b). Under-sized systems can also present sustainability challenges that are non-technical. For customers, reduced availability of the system can be catastrophic when the loss of electrical service at, for example, a maternity ward, may jeopardize lives (World Health Organization and others, 2014). Businesses that cannot depend on the mini-grid will face economic consequences and may employ coping strategies by procuring diesel gen-sets (Rao et al., 2016). Lack of trust in the system to provide reliable electricity may prevent people from purchasing further appliances, thereby decreasing consumption growth.

Standard of practice of energy-use surveying

A survey-approach to energy estimation involves employing a structured questionnaire to: take an inventory of current appliances and likely (near-term) future appliances, determine power ratings and predict daily usage of these appliances (GIZ, 2016; Meier et al., 2010). An average daily energy requirement is then calculated and aggregated for the entire mini-grid. Daily profiles, for example with hourly resolution, can also be formulated if the surveys capture a prediction of what time of day each appliances is used.

This basic structure of energy-use surveys have been present in development programs at least since the 1990s (Ellegård and Nordström, 2001; ESMAP, 1999; Gustavsson, 2007). A comprehensive energy-use survey can be found in World Bank (2003). While this may offer more structure than other simpler surveys, crucially, all depend on respondent-supplied inputs. More recent literature imply that the practice is still prevalent (Alzola et al., 2009; Camblong et al., 2009; Ramchandran et al., 2016).

Challenges with measuring energy demand in low-income households in developing countries have been acknowledged (Brook and Smith, 2000). In their 2000 mini-grid design manual, ESMAP noted that "… *making load projections that reflect reality is frequency a difficult task to accomplish, especially for perspective consumers who have little experience with electrification*" (Inversin, 2000). Unfamiliarity with electricity, changes in behavior before and after installation, and difficulty projecting future energy growth have been noted as specific issues (Ustun, 2016). Furthermore, during design, hourly time-series data for load profiles are typically needed, making the task of estimating this even more difficult through a survey approach (Mandelli et al., 2016a).

Alternative prediction approaches

Alternative energy prediction approaches have been proposed, including use of experts, regression and census data.

The expert approach relies on expert knowledge and judgment to predict consumption using a black-box approach, without specifically collecting individual data (Ghafoor and Munir, 2015). Although simplistic, this approach has the benefit of low data gathering requirements.

Regression can be used to map electricity consumption to explanatory demographic variables, such as the number of people living in the household, and presence of a flush toilet (Fabini et al., 2014; Louw et al., 2008; Zeyringer et al., 2015). Census data of appliance ownership and usage levels can be used to infer likely usage patterns among newly connected customers. This method essentially assumes that the census samples and the target samples are drawn from the same population (Askari and Ameri, 2012; Nouni et al., 2008; Sen and Bhattacharyya, 2014).

Although these approaches produce a prediction of consumption, none attempt to quantify the error. For practical purposes, this leaves a mini-grid designer with few reliable options. These alternative approaches add complexity, additional data and analysis requirements, and costs. Expert-based approaches are the exception, though use of this approach in an actual mini-grid project would be questionable given the subjectivity involved. These obstacles reinforce the practice of energy-use surveying which, in essence, has not changed much in several decades and has not received sufficient critical attention. Consequently, the impact this has on sustainability of the underlying systems, clearly a important issue for furthering electricity access, is not fully known.

Study objectives and paper structure

This research evaluates the accuracy of the widely-employed energy-use survey prediction method. An improved understanding of the boundaries of error will assist mini-grid operators in assessing the risk of consumption prediction error at the design stage. Moreover, it provides them support for weighing the costs and benefits of conducting surveys. The research is based on responses from energy-use surveys for eight Kenyan mini-grids, actual measured consumption over a 31-month period and a follow-up audit of a subset of the customers. Survey, audit and select hourly consumption statistics are included in the Appendix.

Following a description of the systems in Mini-grid description section, we provide the survey methodology in Energy use survey section. Analysis of measured data occurs in Consumption characteristics section. The prediction errors are statistically analyzed in Accuracy of energy-use production section. Results of a follow-up audit are presented in Follow-up audit section. We evaluate an alternative data-driven "proxy" method in Proxy approach of consumption prediction section. An hourly analysis is performed in Hourly load profile analysis section. Discussion section discusses the practical importance of the results. In Conclusions and future work section, we point toward future directions for this research.

Mini-grid description

This work considers eight commercial solar-powered mini-grids installed in Kenya as described in Table 1. The mini-grids are owned and managed by Vulcan, Inc., with on-the-ground operations managed by SteamaCo. Two additional mini-grids were part of the original data set, but due to limited data available on these systems, they were omitted from the analysis. The mini-grids vary in capacity from 1.5 kW to 5.6 kW and are located in different off-grid villages in diverse socio-economic and geographical areas as shown in Fig. 1. The mini-grids serve a variety of customer types, including households, businesses and social service institutions such as churches, among others. The locations were selected based on socio-economic data of nearly 50 communities, balancing potential for commercial success with potential for community development and impact.

Technical description

A generic system diagram for a mini-grid is provided in Fig. 2. The mini-grids are supplied by solar photovoltaic (PV) arrays. Energy is stored in deep-cycle lead acid batteries with capacities ranging from 20.5 to 41 kWh configured in either 24 V or 48 V strings. Charge controllers with maximum power point tracking (MPPT) optimize the energy converted by the photovoltaic arrays and prevent overcharging of the batteries. Inverters supply 230 VAC at 50 Hz to the customers. Customers are nominally provided a 13 A connection and one compact fluorescent light bulb (CFL) upon connection. The cloud-connected meters automatically disconnect customers when their credit balanced drops to zero.

Fig. 1. Location of the mini-grids. The first three letters of each village are used for identification.

Technical data set

The mini-grids were installed at various times from September 2014 to April 2015. It is important to note that not all customers within the same mini-grid were connected on the same date, and that they did not necessarily begin their consumption at the same time. The line to each customer was equipped with a SteamaCo cloud-connected meter that measures and transmits energy consumption with hourly resolution (SteamaCo, 2016). The data set extends through 31 March 2017, spanning up to 31 months for some customers.

As is expected with real-world data, the technical mini-grid data required pre-processing and cleaning. Data corresponding to lines that were installed but never used, had tenant changes, and those with unclear survey responses were censored from the analysis. Only customers with consumption data for at least 93% of the days analyzed were considered.

In general, the mini-grids were designed to have excess capacity in order to reliably supply customer demand. However, in Entesopia on certain days — particularly after a series of rainy days — customer supply was automatically interrupted due to low battery voltage. This artificially lowers consumption. However, the interruptions occurred in the late evening when energy consumption is low. For example, considering those days without interruptions, the consumption between 20:00 and 4:00 averages just 10% of the total daily consumption. To compensate, on days in which interruptions occurred, the measured consumption is increased by an amount commensurate with the consumption that would likely have occurred during the interruptions, based on historical consumption, timing and duration of the interruption.

Any day where more than 30% of the consumption was estimated to have been interrupted was censored from the data set. The average adjustment for days with service interruptions and are included in the data set is just 2.9%, and so the effects of service interruptions does not appreciably detract from the findings of this work.

Tariff structure and implementation

The tariff that each customer pays is structured using economy-of-scale pricing, giving high-consumption customers discounts if their electricity usage reaches a certain threshold. The tariff charged varies between US$1.80/kWh to US$4.00/kWh, depending on the level of consumption.

Energy use survey

Methodology

Data from an energy-use survey carried out at each mini-grid site between June and November 2014 were provided by Vulcan, Inc. The energy-use surveys were conducted by SteamaCo immediately prior to installing the grid. Customers were asked about their current appliances and the appliances that they hoped to acquire once they had electricity access, although no time frame was specified. For each appliance, the customer predicted their typical duration of use each day.

After filtering for completeness, and clarity, energy-use surveys of 176 customers were considered for this work.

Energy use prediction

As is the standard of practice, the results of the energy-use surveys were used to predict average daily consumption according to:

$$\hat{E}_g = \sum_{a=1}^{A} P_a T_a \tag{1}$$

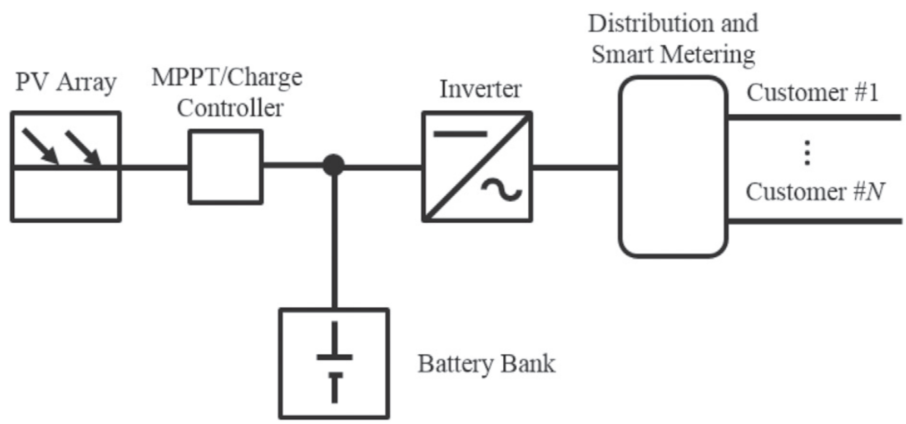

Fig. 2. Representative schematic of a generic N-customer mini-grid.

where $\hat{E}_g$ is the predicted average daily energy consumption of mini-grid customer g in watthours, A is the total number of appliances, and P_a and T_a are the assumed adjusted power rating and predicted hours of use of appliance a, respectively. Note that P_a can be decomposed into $p_a \times K_a$ where p_a is the power rating of the appliance and K_a is the loading percent of the appliance. The loading percent reflects that certain appliances such as refrigerators do not continuously consume rated power. The loading percent is based on estimates or documented typical loading percentages (Energy Information Administration, 2012). The appliances identified by the energy-use surveys, along with the assumed power ratings are provided in the Appendix.

Consumption characteristics

The analysis begins with the statistical characterization of the measured consumption. The analysis is useful in contextualizing the survey results and identifying ways that predictions of consumption can be improved.

Appliance ownership and aspiration

Fig. 3 shows the total quantity of each appliance reported as being owned or aspired to being owned in the energy-use survey. CFLs and mobile phone chargers are the most common, and the median number of appliances per customer is three.

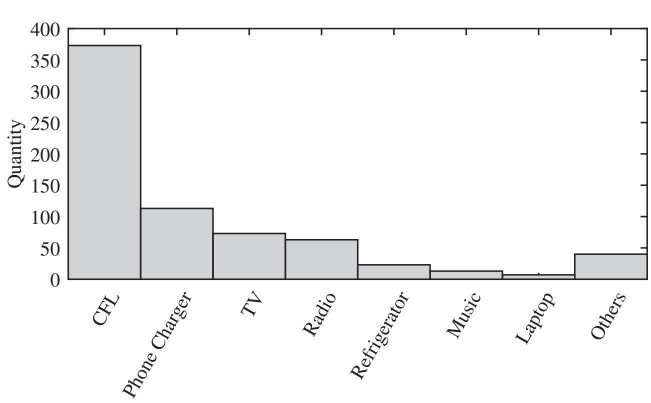

Fig. 3. Count of appliance ownership and aspirational ownership from the energy-use survey.

Consumption trend

The prediction formulated in Eq. (1) is based on average usage behavior and a specific inventory of appliances. It is not well-suited for consumption that varies a-periodically over time, for example as new appliances are added or discarded, or as usage habits or ability to pay for electricity change. The value of even a highly accurate prediction is diminished if it is only valid for a short amount of time.

Fig. 4 shows the average per-customer daily consumption plotted against the number of days since their first use of mini-grid electricity. The average is based on 161 customers that had a consumption record of at least two years, independent of if they responded to the energy-use survey. Linear interpolation was used to estimate the consumption for missing days, if any. The first 15 days are censored as many customers exhibited a short-term spike in consumption immediately following connection.

The overall trend is that of increased consumption as time progresses. Note that because customers did not necessarily begin their consumption on the same day, the abscissa values cannot be mapped to specific calendar days.

Also plotted in Fig. 4 is the best fitting — in a least squares sense — exponential model:

$$\hat{E}_{\text{model}}(t) = \alpha e^{(\beta t)} \tag{2}$$

where $\hat{E}_{\text{model}}(t)$ is the modeled daily consumption, t is time in days and the coefficients $\alpha = 111.7$ and $\beta = 0.00021$. The R^2 value of the model is 0.12, suggesting that other variables influence consumption. The β coefficient corresponds to an 8% increase in consumption

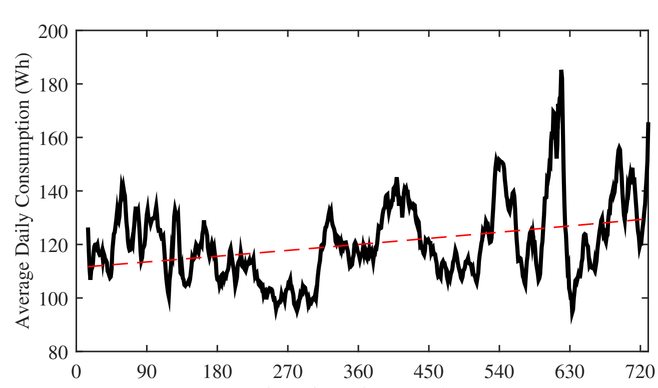

Fig. 4. Trend of average per-customer daily energy consumption since day of first use.

per year based upon this historical data set. However, the increase is not monotonic, and the time-series of consumption does not appear to be stationary, making it challenging for the mini-grid developer to make practical use of a single-point prediction of consumption. It also challenges the evaluation of the prediction's accuracy.

To account for shorter-term fluctuations, in this research the accuracy of the predictions is compared to actual consumption during the year-long period from 30 November 2015 to 29 November 2016. In December 2016, a pilot appliance leasing program began in Entesopia, so it is appropriate to select an end date prior to the program's start. At the start of the period, customers had been connected to electricity for between 7 and 15 months, a reasonable balance of new and established customers.

Village-level consumption

Examined next is the consumption characteristics of each mini-grid. We consider only the 154 customers that have energy-use survey responses and satisfied data-quality requirements for the one-year period ending 29 November 2016. Fig. 5 shows a box plot of the average daily consumption per customer in each mini-grid considering the one-year period.

The center line of the box plot is the 50th percentile (median), the box edges are the 25th and 75th percentiles; and the "+" are outliers. Entesopia has the greatest daily median — 121 Wh, but the rest are similar, ranging from 42 Wh to 81 Wh. The mean of Entesopia — 230 Wh — is inflated by several outliers, and is notably greater than the other mini-grids, which average 84 Wh.

An Analysis of Variation (ANOVA) (Howell, 2010) was conducted on the data. Briefly, ANOVA is a statistical hypothesis test between the means of several sets of samples. The null hypothesis is that the samples are from the same population. In the case of the mini-grid data, the null hypothesis is equivalent to assuming that the customers in, for example, Marti are no different than those in Entesopia in terms of average daily energy consumption, and the variation in averages is due to random chance alone. ANOVA provides information on whether or not the null hypothesis should be rejected for a given confidence level.

Using a confidence level of 0.05, the mean daily consumption of the customers in Entesopia was found to be significantly different from Barsaloi, Opiroi and Marti. All other comparisons among and between the other villages, however, are not statistically significant enough to reject the null hypothesis.

The result of the ANOVA test is surprising, given that the mini-grids are in different locations, with different customer mixes and local economies, yet from a statistical viewpoint, there is no evidence of a difference between them, with the exception of Entesopia. Importantly, it suggests that the mean daily consumption from one grid can be used to accurately estimate the consumption of other grids, with the particular exception of Entesopia. Indeed, Entesopia owes it larger mean to a small number of outliers. This proxy prediction method is explored further in Proxy approach of consumption prediction section.

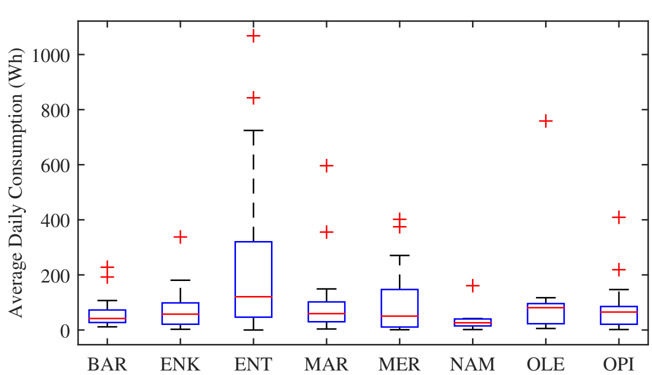

Fig. 5. Box plot of average daily energy consumption per customer, grouped by mini-grid.

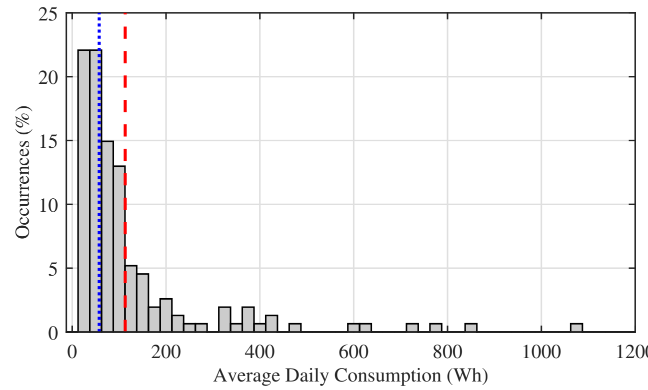

Fig. 6. Histogram of customer average daily energy consumption.

Individual consumption

Narrowing further, we next consider the characteristics of individual customers. The distribution of average daily consumption for each customer is provided in Fig. 6. The dashed line is the mean daily consumption, 113 Wh, whereas the dotted line is the median consumption, 58 Wh. The wide range of average consumption and the skewness of the distribution can make accurate prediction particularly challenging. It is worth noting that the top 10% of consumers — those with average daily consumption exceeding 318 Wh — consume 46% of the total energy. It is therefore especially important to be able to accurately predict the consumption of these high-consumers when compared to other customers.

Averages hide the range of day-to-day consumption of the customers. Twenty-six percent of the 55,767 individual customer-days of considered data showed zero consumption. The median percentage of days with zero consumption when considering individual customers is 13%. In other words, a typical mini-grid customer goes without consumption four days per month. It is unclear exactly why consumption is frequently zero. Possible explanations include: inability to afford or ration energy credits or prolonged travel.

Accuracy of energy-use prediction

We now examine the accuracy of the survey-based energy predictions for the set of 154 customers with energy-use survey and year-long consumption data. The first row of Table 2 provides the actual daily energy consumption statistics, which is important to refer to for context.

Table 1
Mini-grid characteristics.

Location	Installation date	No. of customers
Barsaloi	Oct. 2014	58
Enkoireroi	Sept. 2014	27
Entesopia	Dec. 2014	63
Marti	Apr. 2015	29
Merile	Oct. 2014	27
Namba Koloo	Sept. 2014	19
Olenarau	Oct. 2014	29
Opiroi	Oct. 2014	26
	Total	276

Table 2
Individual customer average daily consumption statistics.

	n	Min. (Wh)	Median (Wh)	Max. (Wh)	Avg. (Wh)	Total (kWh)
Actual	154	0.01	58	1063	113	17.4
Survey	154	16	193	9172	470	72.4

Table 3
Prediction error statistics.

Data set	*n*	Avg. Abs. error (Wh)	Median Abs. error (Wh)	RMSE (Wh)
All days	154	426	151	1079
Zero days censored	154	421	142	1085

Energy use prediction

The bulk statistics of the predictions are provided in second row of Table 2. The prediction error is computed as:

$$e_g = \hat{E}_g - E_g \tag{3}$$

where E_g is the actual average daily consumption of the gth customer and $\hat{E}_g$ is the prediction. Positive values of e_g therefore indicate over-prediction. Note that when considering bulk statistics of individual errors, it is more meaningful to consider $|e_g|$ rather than e_g. Prediction errors are best expressed in watthours, rather than percentages, as the consumption of a few customers is near zero, yielding cumbersomely high percent errors. Instead, the error is given context by noting the average per-customer daily consumption is 113 Wh.

We first examine the distribution of errors, as shown in Fig. 7. The error of the predictions are skewed toward over-prediction. The severity of the over-prediction is most evident in the last column of Table 2, showing the predicted total is more than four times the actual. The error distribution also has a long tail — several errors are greater than 2000 Wh. Unfortunately, this causes wide bounds associated with the predictions: half of the errors lie outside the interval [−32, 287] watthours; and ten percent of the errors are outside [−240, 1830] watthours.

The absolute error statistics are found in the first row of Table 3. The average per customer error is 426 Wh, which is considerable given the average consumption is 113 Wh per day. This is notably similar to the 305% error reported in Louie and Dauenhauer (2016), which suggests wider applicability of the results. The Root Mean Square Error (RMSE), which emphasizes large errors, reflects the long tail of the distribution of error. The median error, which is less influenced by outliers, is lower than the mean; it is 151 Wh. However, it too is larger than the average consumption, showing that error cannot be explained by a few poor predictions.

Of interest is the relative relationship between actual and predicted consumption, which is shown in Fig. 8. Each "+" in the figure corresponds to a specific customer. The zoomed-in inset displays greater detail. For reference, the diagonal corresponds to the trace of error-less prediction and the dashed lines are ±25% prediction error.

More than 70% of the predictions over-predicted consumption. The linear correlation coefficient is 0.28, and is statistically significant using a confidence level of 0.05. Although statistically significant, the correlation coefficient itself is small, indicating that predictions of high (low) consumption are not strongly associated with actual high (low) consumption.

The influence of irregular consumption — the high percentage of days with zero consumption — on prediction accuracy is considered by censoring these days from the analysis. The second row in Table 3 shows the results. The error decrease is marginal, and the predictions remain biased toward over-prediction.

Aggregate prediction

Whereas prediction of an individual customer's consumption is useful in deciding whether or not to connect that customer, aggregate prediction is useful in sizing the critical components of the mini-grid. Aggregate prediction is expected to reduce error because individual over- and under-predictions offset each other.

The error statistics of individual and aggregate consumption, grouped by village, are provided in Table 4. The fifth and sixth columns are the aggregate and individual errors divided by the number of customers. The individual errors were discussed in Energy use prediction section. Aggregation offers some improvement in the error. However, the chronic bias toward over-prediction is evident, as large errors remained after aggregation.

Overall, the results support the suspicion that the standard practice energy-use surveys lead to inaccurate and unreliable predictions of average consumption. Over-prediction of consumption is associated with systems with higher than required reliability and increased capital costs (Louie and Dauenhauer, 2016). It is important to contextualize the magnitude of this prediction error. In Louie and Dauenhauer (2016), it was shown that for each watthour of over-prediction, there was a concomitant increase in critical component cost (battery and PV array) of up to US$6.08. For each customer, this translates into a theoretical increase capital cost of US$918 assuming over-prediction by the median error.

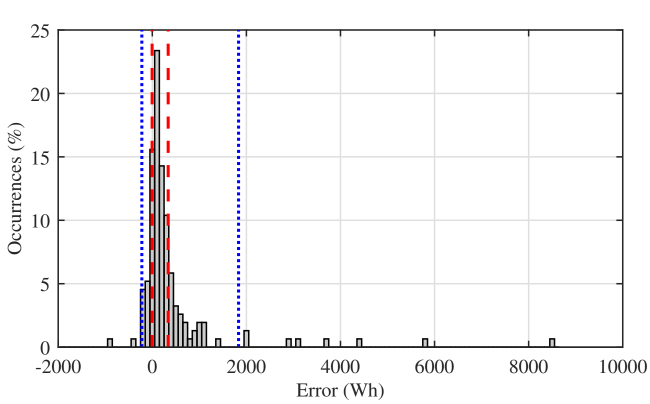

Fig. 7. Distribution of energy-use survey prediction errors. The 25th and 75th percentiles are shown as the dashed lines, and the 5th and 95th percentiles are the dotted lines.

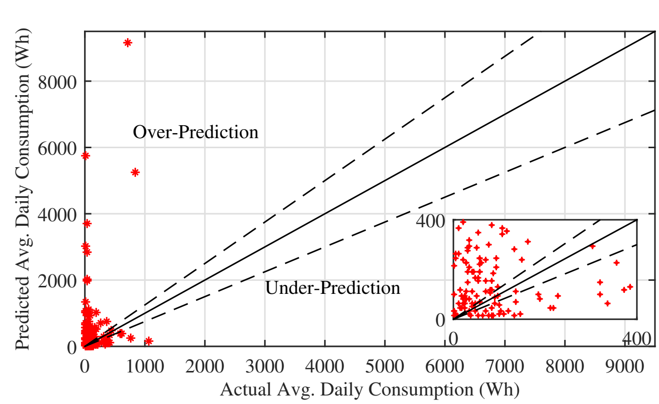

Fig. 8. Predicted average daily consumption versus actual consumption, with zoomed-in view inset.

Table 4
Aggregate prediction error statistics.

Data	n	Actual (Wh)	Survey (Wh)	Agg. error (Wh/Cust.)	Individual error (Wh/Cust.)
Barsaloi	29	1680	12,418	370	391
Olenarau	11	1352	2526	107	211
Enkoireroi	15	1246	7746	433	476
Entesopia	31	7132	19,703	405	581
Marti	24	2326	11,361	376	408
Merile	18	1794	14,676	716	727
Namba Koloo	7	301	1392	156	156
Opiroi	19	1565	2537	51	148
All	154	17,365	72,359	357	426

Table 6
Appliance rating comparison.

Appliance	Energy Survey (W)	Audit Min. (W)	Audit Avg. (W)	Audit Max. (W)
CFL	8	9	11.1	20
DVD player	15	10	20.9	25
Hair dryer	150	900	900.0	900
Iron	1500	1100	1100.0	1100
Hot air gun	375	375	375.0	375
Laptop	80	60	60.0	60
LED light	5	5	6.9	20
Music system	150	25	52.5	80
Radio	10	18	26.2	50
TV	35	36	61.6	92
TV decoder	25	18	19.5	20

Follow-up audit

A possible explanation for errors in energy use surveys is their prognostic nature. Respondents must estimate their future appliance ownership and behavior. To explore the accuracy of these estimations, a follow-up audit of actual appliance ownership was conducted in Entesopia ($n = 42$) and Barsaloi ($n = 44$). A total of 48 customers responded to both the survey and audit. The audits were conducted from 29 Nov. 2016 to 1 Dec. 2016 in Entesopia and from 26–27 Jan. 2017 in Barsaloi. These locations were chosen as they have the greatest amount of customers and they represent mini-grids with high and low average consumption.

The following sub-sections discuss potential sources of error which may have caused the differences between predicted and actual demand. Potential sources of error include: number of appliances, appliance rating, and hours of use.

Appliance inventory

Table 5 compares the appliances from the energy-use survey and the audit. Only the top 15 appliances by number documented in the audit are shown. The total quantity of some appliances in the survey closely matches those in the audit, for example CFLs, TVs and radios. Others show a wide discrepancy. However, the energy-use survey tends to under-predict appliance ownership compared to the audit.

The accuracy of energy-use surveys to predict appliance ownership tends to be low. The final column in Table 5 shows the percent of respondents whose survey responses matched the audit for each appliance. The calculation does not include those that indicated in the energy-use survey they did not or would not own an appliance, otherwise the accuracy for uncommonly owned appliances would be high. Customers were error-prone in particular at predicting ownership of all appliances other than CFLs and TVs. Not a single customer was able to accurately predict their entire appliance inventory. It is notable that most of the high-power devices such as cookers, microwaves and kettles, that customers predicted that they would have, were not found during the audit.

Table 5
Appliance ownership comparison.

Appliance	Survey total	Audit total	Percent correct
Phone charger	17	115	0
CFL	117	93	18.8
LED	0	26	0
TV	26	23	29.4
TV decoder	0	14	0
Radio	12	14	4.0
DVD player	2	12	7.7
Woofer	2	7	0
Music system	4	6	11.1
Hair dryer	2	4	0
Freezer	1	4	0
Blow dryer	0	4	0
Desktop computer	0	3	0
Laptop computer	0	2	0
Fan	0	2	0

Appliance ratings

Energy-use surveys in general rely on assumptions of appliance ratings. Table 6 shows the appliance ratings assumed in the energy-use survey and those documented in the audit. Shown are only those appliances that appeared in the energy-use survey and the audit, and had accessible nameplate ratings during the audit. Assumed and documented ratings of the other appliances are found in the Appendix.

The assumed appliance ratings generally are within the range of the minimum and maximum values documented in the audit. Although the ratings did not match those documented in the audit, the rating mismatch alone does not explain the observed large prediction errors. In fact, replacing the assumed ratings in the energy-use survey with the average documented in the audit increases the error of the prediction.

Hours of use

Another source of error is in the hours of use of each appliance. During the appliance audit, the customers were also asked to estimate their average hours of use of each appliance in a similar fashion as was done in the energy use survey. It must be borne in mind that these are customer estimates of usage, and that comparing survey *predictions* with audit *estimations* explores response consistency and not necessarily accuracy as actual duration of use was not measured.

Comparing the estimates made during the audit with the predictions made during survey again shows a disconnect. Table 7 shows the average number of hours of use of those respondents that indicated at least one appliance of a given type in either the survey or audit. For example, those indicating at least one CFL on the survey and were found to have at least one during the audit originally predicted on average 4.1 h of usage, but estimated 3.0 h of usage during the audit. Note that only one customer indicated a radio in the survey and was found to have one in the audit, so this result might not be generally representative. With the exception of TVs, which matched reasonably well, the survey responses tended to predict increased appliance use compared to the audit.

Table 7
Appliance duration of use comparison.

Appliance	Survey Avg. Hrs.	Audit Mean Hrs	Percent match
CFL	4.1	3.0	15.0
TV	3.5	3.8	0
Radio	8.0	3.0	0
Phone charger	4.7	2.3	33.3

Table 8
Predicted and estimated consumption errors.

Data set	n	Avg. Abs. error (Wh)	Med. Abs. error (Wh)	RMSE (Wh)
Survey prediction	42	565	143	1563
Audit estimate	42	286	155	580

Estimation of energy consumption

We next construct an estimation of average consumption based upon the findings of the appliance audit and the customer estimation of usage. Here, the nuance between prediction and estimation is important. Predictions, as based on the energy-use survey, can be used in pre-implementation design and planning of a mini-grid. The estimates are obtained post-implementation. Nonetheless, the accuracy of the estimates can be considered a near-best case[1] of predictions, since they are based on actual appliance inventories and customers' recollection of recent consumption, which should be more accurate than speculative prediction.

A total of 42 customers responded to both the energy-use survey and the audit and had consumption data that satisfied data quality requirements. The most recent 31 days prior to start of the appliance audit for each grid is considered.

The error statistics of the estimations and predictions for this subset of customers are provided in Table 8. For reference, the average actual daily consumption for this subset of customers over that interval is 182 Wh. As expected, the estimations are more accurate than the predictions, 286 Wh compared to 565 Wh. The improvement was largely due to a reduction in particularly high-error predictions, as the RMSE reduced by over 60%, although the median increased slightly to 155 Wh. Despite the improvements, the estimation error remains high and is larger than the average consumption.

Proxy approach of consumption prediction

The overall results show that prediction errors from energy-use surveys for commercial mini-grids remain unsatisfyingly high, suggesting that alternative approaches be considered. To further probe the potential of alternative prediction methods, we explore a data-driven proxy approach. In this approach, given the results of the ANOVA test discussed previously, the average consumption of one mini-grid should provide a reasonably accurate prediction of another.

For direct comparison, the same 154 customers from Accuracy of energy-use production section are considered. For each mini-grid, the average daily consumption per customer $E_{\mathrm{grid},v}$ is computed as:

$$E_{\mathrm{grid},v} = \frac{1}{C}\sum_{g\in\mathcal{D}} E_g \tag{4}$$

where $\mathcal{D}$ is the set of considered customers in grid v, and C is the number of customers in $\mathcal{D}$. The error when grid j is used to predict grid k is:

$$e_{\mathrm{grid},j,k} = E_{\mathrm{grid},j} - E_{\mathrm{grid},k}. \tag{5}$$

The actual consumption is based on the one-year period ending 29 Nov. 2016. The results are provided in Table 9. The mean error and mean of the absolute errors are provided in the last two columns. This approach offers a clear improvement in accuracy over the energy-use survey approach, with average errors ranging from −146 Wh to 67 Wh per customer. It must be noted that in this exploratory analysis the data sets overlap in time, which would be impossible in a real-world application. Nonetheless, the results offer insight into the potential of data-driven proxy consumption prediction approaches.

There are important implications of this result. If data sets are widely available, these results suggest that mini-grid developers can use them to better predict consumption than the common survey approach.

This approach can likely be improved upon, for example, by weighting the results based on the customer make-up — number of businesses, households or other establishments connected to the grid. However, this is beyond the scope of this work.

Hourly load profile analysis

While previous sections have addressed total daily usage at the individual customer and aggregate mini-grid levels, it is equally valuable to consider the impact of energy-use-surveying as an input to constructing hourly consumption profiles.

Computer-aided mini-grid design programs such as HOMER require predictions of load profiles on an hour-by-hour basis as well as parameters indicating the variability of consumption. Constructing such a load profile should be more challenging and error-prone than predicting the daily average alone because the timing of the consumption must be specified. Methods and tools have been proposed to construct load profiles based on appliance inventory and ratings (Boait et al., 2015; Mandelli et al., 2016b). In these methods, random variables are used to model the likelihood that an appliance will be used during any given time throughout the day. Reasonable results can be obtained if the predictions of the appliance inventory and power ratings are accurate, which can be a questionable assumption, as shown in this paper. In addition, the designer must have some knowledge of the distribution of the random variables used in the methods. Nonetheless, these methods are able to construct load profiles which contain information related to peak power demand and time-of-use consumption, which is missing from a blanket daily average estimation.

Prototypical profiles

The hourly load profiles for 42 customers in Entesopia and Barsaloi were constructed based on the consumption from the most recent 31 days prior to the appliance audit. The load profiles were classified into three groups based on time of consumption. In order to meaningfully combine load profiles from customers whose average daily consumption are different, the hourly data are normalized for each user by dividing the hourly consumption by that user's average total daily consumption across all 31 days, that is:

$$E^{*}_{\mathrm{hr},g}[h] = E_{\mathrm{hr},g}[h]\frac{\sum_{d=1}^{D} E_{\mathrm{day},g}[d]}{D} \tag{6}$$

where $E^{*}_{\mathrm{hr},g}[h]$ is the normalized hourly consumption of the gth customer during hour h, $E_{\mathrm{hr},g}[h]$ is the customer's hourly consumption on the hth hour, $E_{\mathrm{day},g}[d]$ is the customer's daily consumption on the dth day, and D is the number of days considered. The normalized values for a given group and hour are treated as samples from the same population and their statistics are computed for that hour, resulting in a load profile. The statistical moments and select quantiles of the load profiles for all groups are provided in the Appendix.

The first group, "Night Users", consumed more than 75% of their daily energy on average from 18:00 to 6:00. This corresponds to the

[1] Assumptions were needed for appliances without nameplate ratings, and for loading percentages, as detailed in the Appendix. These assumption introduce error aside from estimations of duration of use.

Table 9
Per customer error of proxy method.

	Predictor error (Wh/customer)								Avg.	Avg.
	BAR	ENK	ENT	MAR	MER	NAM	OLE	OPI	e_{grid}	$\lvert e_{grid}\rvert$
BAR	–	25	172	39	42	−15	65	24	50	55
ENK	−25	–	147	14	17	−40	40	−1	22	40
ENT	−172	−147	–	−133	−130	−187	−107	−148	−146	146
MAR	−39	−14	133	–	3	−54	26	−15	6	40
MER	−42	−17	130	−3	–	−57	23	−17	3	41
NAM	15	40	187	54	57	–	80	39	67	67
OLE	−65	−40	107	−26	−23	−80	–	−41	−24	55
OPI	−24	1	148	15	17	−39	41	–	22	41

hours in which solar input is generally not available, and so is significant from a design perspective, as batteries are among the most costly components of a mini-grid. A total of 18 (43%) customers are in this group.

The corresponding load profile is shown in Fig. 9. A modified box plot is generated for each hour, where the thick green line is the mean value, the thin red line is the median value, the "+" correspond to the maximum and minimum values, and the whiskers extend to the 5th and 95th quantiles. A prototypical Night User consumes the majority of their energy in the six-hour window from 18:00 to 23:00, but there is wide variation of usage during these hours.

The next group of users consumed more than 50% of their energy between 6:00 and 18:00, and are referred to as "Day Users." The load profile for this group is in Fig. 10. Only seven (17%) customers are in this group. Many in this group were businesses or mixed business/household customers. Day Users spread their consumption over many hours, from 8:00 to 18:00. The peak occurs around 12:00. Note that the ordinate has been truncated at 120% of average daily consumption to better show the detail of the mean values. The maximum values can be found in the Appendix. During several daytime hours, the maximum measured values exceeded 200% of the daily average. This is indicative of intermittent energy-intensive activities that can occur in businesses.

The remaining 17 (40%) customers are "Mixed Users", with associated load profile shown in Fig. 11. Like Night Users, Mixed Users have a concentration of usage in the few hours following 18:00, but unlike Night Users, they also consume during the day.

The load profiles could be used a starting point for a data-driven load profile synthesis program. Although beyond the scope of the present work, the statistical characteristics for each hour could be used to model the distribution of consumption for each hour of the day. A Monte Carlo approach could be used to draw samples for each hour of interest, creating a time series of values. The values would be multiplied by the an estimate or prediction of the customer's average daily consumption to produce a time series whose units are watts. Such a method would be a compromise between top-down and bottom-up approaches as detailed knowledge of appliances, ratings and windows and probabilities of usage need not be explicitly modeled. Rather, each customer would need to be classified as belonging to one of the three groups and their average daily consumption predicted either by a proxy approach or surveys. Of course, this does not escape the error in predicting the average daily consumption, but it is a method for converting daily predictions into hourly values, which can then be used in mini-grid design programs such as HOMER.

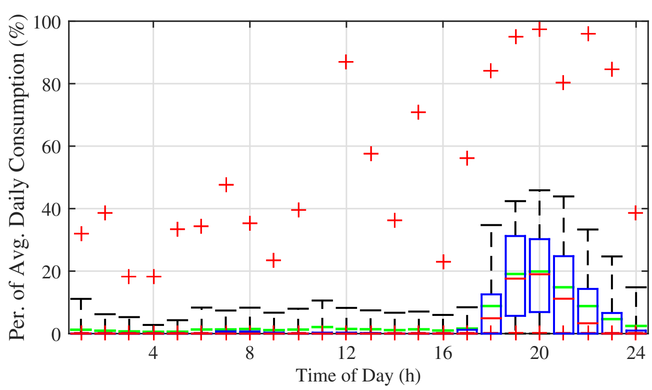

Fig. 9. Prototypical load profile for Night User group. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

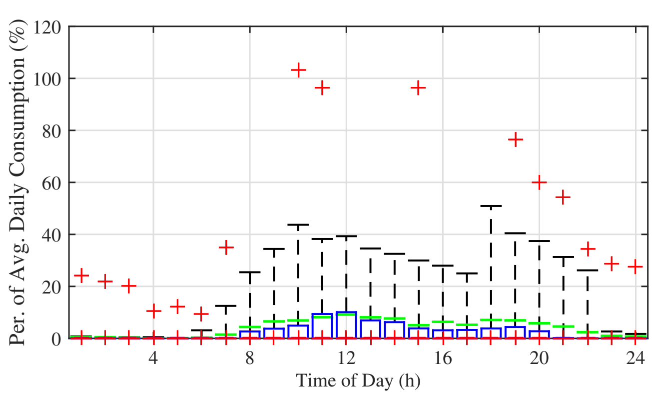

Fig. 10. Prototypical load profile for Day User group.

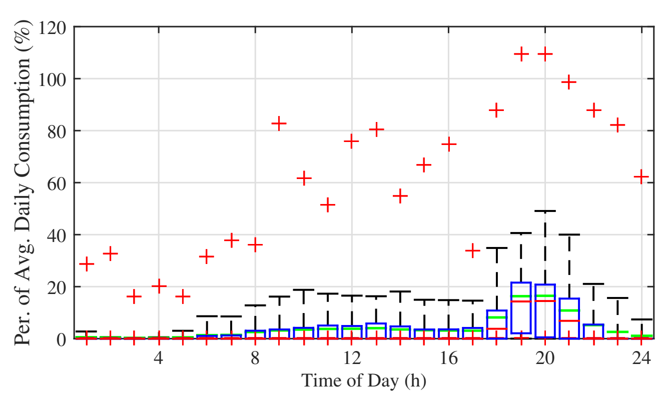

Fig. 11. Prototypical load profile for Mixed User group.

Table 10
Case studies audit data.

Appliance	Quantity	Rating (W)	Duration (h/day)	Est. (Wh/day)	Actual (Wh/day)
Case I				400	266
TV	1	80	5		
Case II				3658	1995
CFL	3	11	3		
Mobile phone	4	3.7	4		
Hair dryer	1	900	2		
Blow dryer	1	1700	1		
Case III				81	391
CFL	2	11	3		
Mobile phone	2	3.7	2		

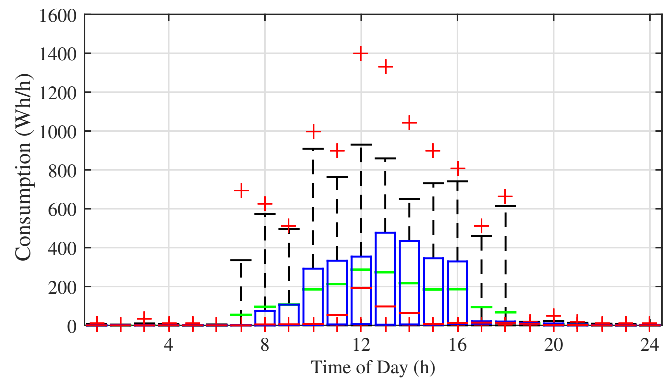

Fig. 13. Load profile of Case II Day User.

Case studies

Although proto-typical load profiles are useful understanding the hourly consumption of a group of similar users, insight is also gained from examining individual users as is done in the following three brief case studies. The results from the appliance audit for these cases is provided in Table 10. These case studies provide a micro-level glance at some of the errors which may contribute to the inaccuracy of load estimation.

Case I

The first case is a customer in Entesopia that is classified as a Night User, with load profile in Fig. 12. The appliance audit shows that only a television is powered by the mini-grid (a separate solar home system is used for lighting). It is notable that the consumption often exceeded the television's rated power from 18:00 to 20:00. There are several possible explanations for this: allowing neighbors to plug in additional appliances, inaccurate rating or variable power consumption of the television, and appliances not reported during the audit or a change in appliances owned. Hourly consumption in excess of the sum of a customer's appliance ratings was observed in several customers. The inability of a survey or audit to capture this characteristic contributes to prediction errors, and should be included in bottom-up load profile construction methods.

The customer's estimate of hours of usage reflects the *window* in which the TV was primarily used, from 17:00 to 21:00, but not the actual average hours of usage. If customers tend to estimate windows of use, instead of duration of use, then it would explain the tendency to overestimate consumption. Finally, even with this single appliance, the range of hourly consumption varied considerably during their peak hours of usage.

Case II

The second case is a Day User in Entesopia with load profile in Fig. 13. This customer primarily uses electricity for high-power appliances in their hair salon business. These appliances are likely used intermittently throughout the daytime. As a result, the daytime hourly distributions are non-Gaussian, exhibiting strong positive skewness. Indeed, the median value is near zero even during late morning and late afternoon hours, presumably when the business is open. Non-Gaussian characteristics were observed for most hours for most customers, independent of group.

This case highlights the special challenge of predicting consumption of service-based businesses in which appliance use is dependent on an intermittent, irregular flow of customers. Predicting or estimating the cumulative daily duration of use of an appliance that is used intermittently is likely more error-prone than of an appliance that is only turned on and off once per day such as a light or television. Models of hourly consumption would need to based customer flow, which requires additional modeling and introduces error. Methods that construct hourly load profiles using a bottom up approach would need to incorporate in their formulations the exclusions or dependencies of appliance use. For example, is it possible for the hair dryer and blow dryer to be used simultaneously? Are the hair dryer and blow dryer used one after another as part of a process?

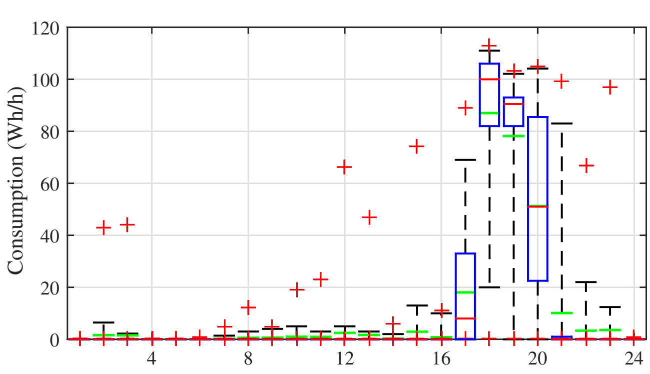

Fig. 12. Load profile of Case I Night User.

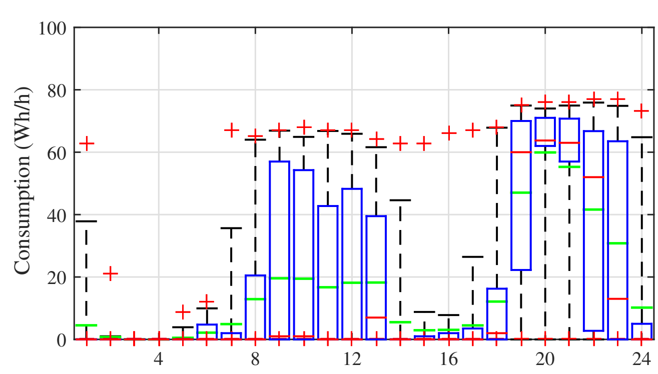

Fig. 14. Load profile of Case III Mixed User.

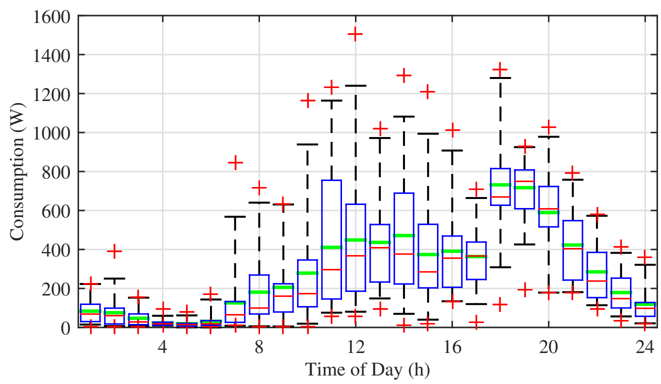

Fig. 15. Aggregate hourly load profile for 19 customers in Entesopia.

Case III

The third case is a Mixed User in Barsaloi. From Table 10, the audit showed only two CFLs and two mobile phones, yet their actual consumption is five times what is estimated from the appliance audit. There is no obvious explanation. It serves to highlight that for some customers a survey or audit will not be reliable input data for predicting average daily consumption or hourly load profiles.

This customer's profile, shown in Fig. 14, deviates from the prototypical Mixed User. The customer's mini-grid service is used to power their business and home, which can explain the presence of day and evening consumption. While the shapes of many customers' individual load profiles generally correspond to the prototypical profiles, some do not, as shown with this customer.

Aggregate profiles

When many customers are served by the same mini-grid, the diversity of Night, Day and Mixed Users should reduce the observed peak consumption and tend to flatten the aggregate load profile. The aggregate profile for Entesopia and Barsaloi, are shown in Figs. 15 and 16. The profiles are based on the last 31 days of consumption for the same 42 customers considered previously. Some days within the 31-day window were censored because of missing hourly data from one or more customers during this period. In Entesopia, 22 days are considered; in Barsaloi, 25 days are considered. Note that the consumption data has not been normalized, as it is for the prototypical load profiles.

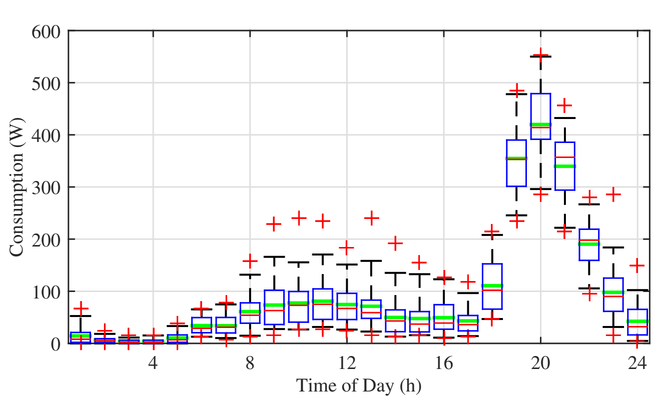

Fig. 16. Aggregate hourly load profile for 23 customers in Barsaloi.

The benefits of load diversification are apparent as the peak loads from individual customers were not co-incidental. The sum of the maximum customer hourly consumption for Entesopia is 3822 Wh/h, but the maximum aggregate demand was 1507 Wh/h. In Barsaloi, the diversification effect is less, but still meaningful, as the sum of customer hourly maximums was 975 W/h but the maximum hourly aggregate was 553 Wh/h.

Discussion

The energy-use surveying approach, from which the data were drawn, followed the standard practice employed prior to installing mini-grids in developing countries. Our primary conclusion is that the prediction error derived from energy-use survey is unacceptably high — alternative approaches are preferable such as the proxy-method demonstrated herein or others mentioned in Alternative prediction approaches section.

The preceding Sections showed that respondents over-estimated their actual usage, with an average error of 3.3 times the actual consumption. By querying historical use of appliances, the audit removed the aspirational elements of the estimation – both for ownership of appliances and in most cases, the power ratings of the appliances – leaving only the time-of-use estimation uncertain. Despite these advantages, audit-based estimations still erred by 286 Wh. The remaining error therefore seems likely to be heavily influenced by the survey process itself and the respondent time of use estimation. The hourly analysis suggested that respondents might be fairly accurate at estimating the window of hours an appliance might be used, but not the average hours of actual use in that window. Further, estimating the average use of appliances used intermittently was error-prone.

A proxy-method which assumes that the sample of customers from one mini-grid can be drawn from the same population of the customers of a mini-grid in planning yielded better results. Average unweighted absolute error of per customer prediction at one village to another was 60.6 Wh. Exclusion of the one outlier mini-grid, Entesopia, where a few large customers drove up per customer usage, reduces the prediction error to 48.4 Wh. This method could perhaps be combined with a proto-typical load profile to construct an hourly time series of consumption.

Potential causes of error in energy use surveying

Sources of error within the survey process is well covered in the literature. Specific examples are: interviewer bias (Freeman and Butler, 1976; Kish, 1962), satisfying behavior (Krosnick, 1991), respondent cognitive burden (Bradburn, 1978), evasive answer bias (Warner, 1965), future estimation bias, cultural issues (Schwarz, 2003) and other forms of response bias. It is beyond the scope of this paper to estimate the individual impacts these effects may have had on energy use surveying. However, as respondent estimation of time of use of appliances seems to be particularly erroneous, it is worthwhile to start the discussion on how biases can arise.

In the audit, there appeared to be a desire of the customer to overstate their usage to accentuate the impression of a high cost burden in hopes that this might influence the operator to reduce prices. When asked during the follow-up audit the degree to which the price of electricity impacted usage, 68.4% and 51.2% of respondents at the Barsaloi and Entesopia, respectively, responded that it had a high or very high impact. It can be argued that at the time of the energy-use survey, customers may have had difficulty when factoring in the

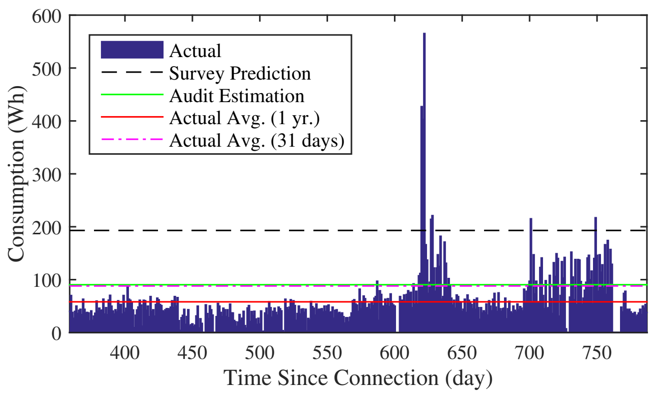

Fig. 17. Consumption of an Entesopia customer with accurate 31-day estimation, but poor long-term prediction of consumption.

price of energy in the estimate — effectively requiring an implicit knowledge of their demand curve.

The time-perception issue when predicting appliance energy use and ownership seemed problematic. In forward-looking surveys, aspirational usage is meant to capture likely growth such that the design will not immediately be overcome by a growing demand. A yearly 8% growth over the first two years was shown for the mini-grids in this study. Yet with a point estimate of future use, even if perfectly accurate, will only be so for a short period in the future. Thus, a design that planned for 16% on top of immediate consumption will be over-designed for the first two years and under-designed thereafter.

The audit demonstrated how some respondents grappled with the time issue. Fig. 17 shows a customer who was quite accurate in estimating their last 31-days of average consumption, but had relatively poor prediction of consumption.

The energy prediction in (1), though conceptually simple and widely employed, has several drawbacks. Appliance ratings can vary widely among the same type of appliance. For example one study found that depending on the specific model, DVD players consume between 5 and 17 W (Lawrence Berkeley National Laboratory, 2017). If the survey respondent does not already own the appliance, it is challenging to know which value to select. It is also problematic to assign a loading percent to certain appliances. For example, the power consumed by music systems at least partially depends on the volume and even the type of music played (JBL Corp., 2017), and for other appliances there is no documented typical loading percent. Stand-by consumption of appliances — the power they consumed when turned off — further complicate and potentially add error to this formulation (Lawrence Berkeley National Laboratory, 2017). The hourly profiles showed that power consumption in excess of the total power ratings were common, suggesting that the appliance power draw is variable or inaccurate or that unreported appliances were used.

Implications for future practices

Mini-grid design requires a reliable prediction of consumption in order to size its critical components to be both reliable and economically viable. The use of energy-use surveys to predict consumption has been shown to be prone to significant errors from within the survey process itself, suspected to be from response and interviewer biases.

Use of the proxy-method is a simple and preferable approach where mini-grid load demand data sets can shared. When surveys are used for design purposes, this research suggests added scrutiny over the energy-use survey process itself and caution over the results which, on an average consumption of 113 Wh, can error by ±426 Wh. Ensuring reliability of the system with these error bounds suggest a roughly four-fold increase in sizes of the major components which is unlikely to be economically possible.

Improvements to the energy-use survey are needed to increase its accuracy. Until more data sources are made widely available of mini-grid operational data, energy-use surveying is likely to remain the prevailing practice. Researchers and practitioners need to revisit this approach to be more aware of the cognitive challenge for respondents of such a survey. Efforts to reduce the burden of responses are needed, for example by clearly defining the terms of use in the future such as likely expense levels and availability of the system. Another approach would be to sensitize potential customers to the functionality, limitations and costs of appliance ownership prior to design. Furthermore, responses aimed at influencing the mini-grid operator's decisions may play a larger role in introducing error than previously thought. Finally, identification and further vetting of customers who are predicting relatively heavy consumption is needed to reduce their error in prediction.

From a research perspective, and with sufficient scope to conduct the analysis, this paper has argued that increased accuracy can be achieved through the proxy-method. Practically, this is not always possible for a project to undertake. Mini-grid designs which must depend on energy-use surveys as primary load data can, as much as possible, design for flexibility and modularity when up-/down-scaling is needed after consumption is known.

Conclusions and future work

This paper evaluated the accuracy of the standard practice of conducting energy-use surveys prior to implementation of mini-grid projects in developing countries. These surveys provide a key input to mini-grid design (a profile of demand) constructed by taking an inventory of each household's current and aspirational appliances: the type, power rating, and expected time of use. Operational data from eight mini-grids in Kenya covering over two years of data was used to assess the prediction errors.

Predictions were poor, with error arguably most influenced by duration of use estimations and the general survey approach. Building a mini-grid system fit for the estimated demand, which had a mean absolute error of 426 Wh on a mean consumption of 113 Wh per day per customer, would nominally cost four-times the amount when compared to a system built to meet actual demand. This excess spend significantly affects the viability of the business model.

The use of a data-driven proxy village method, which used mean customer consumption from each mini-grid to predict customer consumption from other mini-grids, was a more accurate approach, reducing mean absolute error to 42.4 Wh. This discovery highlights the importance of sector-wide sharing and aggregation of mini-grid consumption data. Construction and analyses of hourly consumption profiles, and comparison to predictions and appliance ownership provided further insights: hourly consumption exceeding the total rating of the appliances was common; and, for some customers, their load profile could not be reconciled with their appliance audit or estimated hours of use. Most, but not all, users exhibited one of three prototypical load profiles, which could be used in conjunction with a prediction of daily average consumption to formulate a load profile. The aggregate hourly profiles showed the impact of diversification in reducing peak demand. This research has indicated that caution is warranted when using energy-use survey results for design activities. Furthermore, the energy-use errors reported support the need for a renewed attention to reducing error from energy-use surveying methods, as well as research into alternative load prediction methods.

Appendix A

Appliance-specific power ratings and assumed loading percentage for the appliances are provided in Table 11. The ratings shown under the Survey heading are the assumed ratings inclusive of the loading percent adjustment, and were provided to authors along with the survey results. The audit ratings are the appliance nameplate rating. Note that an "*" indicates that the nameplate power rating was not found for all appliances for a given type; in these cases, the power rating was estimated. For some appliances, the loading percent is based on additional information provided by the customer and so the values shown might not be generally representative. The appliances listed are those that appeared in the energy-use survey responses of the 154 customers considered in Accuracy of energy-use production section or in the audit of the 42 customers considered in Estimation of energy consumption section.

Table 11
Appliance ratings.

Appliance	Survey (W)	Audit Min. (W)	Audit Avg. (W)	Audit Max. (W)	Audit Load Per. (%)
Air pump	–	–	180*	–	10
Blow dryer	–	375	1369	1700	100
CFL	8	9	11	20	100
Cables	30	–	–	–	–
Cooker	5000	–	–	–	–
Desktop computer	–	225	225	225	33
DVD player	15	10	21	25	100
Fan	–	–	20*	–	100
Florescent tube light	–	18	18	18	100
Freezer	–	–	19	–	100
Hair dryer	900	900	900	900	100
Hot air	375	375	375	375	100
Lamp	9	–	–	–	–
Kettle	2000	–	–	–	–
Laptop	80	60	60	60	40
LED light	5	5	7	20	100
Microwave	900	–	–	–	100
Music system	150	25	52.5	80	100
Phone charger	5	–	3.7*	–	100
PA system w / mixer	–	–	300*	–	100
Printer	150	–	65	–	100

Table 12
Appliance ratings continued.

Appliance	Survey (W)	Audit Min. (W)	Audit Mean (W)	Audit Max. (W)	Load Per. (%)
Shaver	10	–	–	–	–
Soldering iron	–	40	40	40	100
Radio	10	18	26	50	100
Refrigerator	40	–	–	–	–
Rolling kit (for hair)	–	12	12	12	100
TV	35	36	61	92	100
TV decoder	25	18	19.5	20	100
Welding machine	–	–	3120	–	10
Woofer	50	15	22	30	100

The actual average daily consumption for the one year period 30 November 2016 to 29 November 2017, energy-use survey predicted and audit-estimated consumption are provided in Tables 13–15. Only customers with energy-use survey responses are shown.

Table 13
Data summary.

ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)	ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)
1	BAR	192	530	308	27	BAR	52	96	213
2	BAR	53	169	148	28	BAR	69	1130	62
3	BAR	34	24	416	29	BAR	11	690	–
4	BAR	87	185	580	30	ENK	79	610	–
5	BAR	34	32	165	31	ENK	80	107	–
6	BAR	37	16	72	32	ENK	338	63	–
7	BAR	83	16	116	33	ENK	181	1023	–
8	BAR	35	1992	–	34	ENK	163	314	–
9	BAR	83	380	104	35	ENK	45	1013	–
10	BAR	27	80	88	36	ENK	18	48	–
11	BAR	34	80	61	37	ENK	101	58	–
12	BAR	103	32	349	38	ENK	23	1058	–
13	BAR	49	285	81	39	ENK	91	88	–
14	BAR	65	1070	–	40	ENK	14	1005	–
15	BAR	16	96	96	41	ENK	58	188	–
16	BAR	228	96	444	42	ENK	35	2008	–
17	BAR	26	225	141	43	ENK	3	100	–
18	BAR	46	64	–	44	ENK	20	63	–
19	BAR	43	2845	125	45	ENT	29	114	–
20	BAR	32	290	120	46	ENT	1068	139	808
21	BAR	27	80	–	47	ENT	121	32	586
22	BAR	42	192	339	48	ENT	373	118	1127
23	BAR	21	80	177	49	ENT	319	148	400
24	BAR	13	364	–	50	ENT	98	133	–
25	BAR	107	344	693	51	ENT	321	99.5	476
26	BAR	27	935	76	52	ENT	89	66	371

Table 14
Data summary continued.

ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)	ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)
53	ENT	843	5242	602	79	MAR	19	1360	–
54	ENT	387	128	3658	80	MAR	4	244	–
55	ENT	608	375	521	81	MAR	56	244	–
56	ENT	40	51	–	82	MAR	79	300	–
57	ENT	135	67	–	83	MAR	7	588	–
58	ENT	92	210	191	84	MAR	34	295	–
59	ENT	132	16	665	85	MAR	58	244	–
60	ENT	186	98	–	86	MAR	94	244	–
61	ENT	71	156	–	87	MAR	71	446	–
62	ENT	0	32	26	88	MAR	98	94	–
63	ENT	84	157	–	89	MAR	34	320	–
64	ENT	211	48	–	90	MAR	132	280	–
65	ENT	188	83	970	91	MAR	149	240	–
66	ENT	724	9172	632	92	MAR	30	146	–
67	ENT	304	296	547	93	MAR	30	3712	–
68	ENT	32	276	260	94	MAR	55	728	–
69	ENT	31	193	29	95	MAR	61	225	–
70	ENT	43	193	210	96	MAR	29	225	–
71	ENT	58	193	90	97	MAR	355	225	–
72	ENT	37	555	362	98	MAR	596	374	–
73	ENT	0	424	29	99	MAR	106	366	–
74	ENT	59	424	–	100	MER	162	558	–
75	ENT	450	464	875	101	MER	10	558	–
76	MAR	22	241	–	102	MER	147	496	–
77	MAR	135	110	–	103	MER	47	32	–
78	MAR	72	110	–	104	MER	102	512	–

Table 15
Data summary continued 2.

ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)	ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)
105	MER	25	48	–	130	OLE	99	21	–
106	MER	271	689	–	131	OLE	81	121	–
107	MER	8	482	–	132	OLE	759	261	–
108	MER	58	69	–	133	OLE	54	156	–
109	MER	54	16	–	134	OLE	5	261	–
110	MER	375	767	–	135	OLE	83	121	–
111	MER	41	528	–	136	OPI	83	48	–
112	MER	55	759	–	137	OPI	147	24	–
113	MER	19	5740	–	138	OPI	219	48	–
114	MER	1	32	–	139	OPI	58	220	–

Table 15 (*continued*)

ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)	ID	Grid	Actual (Wh)	Survey (Wh)	Audit (Wh)
115	MER	11	3022	–	140	OPI	77	16	–
116	MER	8	16	–	141	OPI	111	24	–
117	MER	402	352	–	142	OPI	70	32	–
118	NAM	24	64	–	143	OPI	11	64	–
119	NAM	26	64	–	144	OPI	409	126	–
120	NAM	161	514	–	145	OPI	40	114	–
121	NAM	2	214	–	146	OPI	71	238	–
122	NAM	42	64	–	147	OPI	65	16	–
123	NAM	35	208	–	148	OPI	86	16	–
124	NAM	12	264		149	OPI	20	391	–
125	OLE	117	356	–	150	OPI	55	348	–
126	OLE	5	21	–	151	OPI	2	16	–
127	OLE	47	166	–	152	OPI	16	46	–
128	OLE	86	921	–	153	OPI	2	656	–
129	OLE	15	121	–	154	OPI	24	94	–

The mean (μ), standard deviation (σ), skewness (γ) and kurtosis (β) and quantiles for the load profiles are provided in Table 16 to Table 20. The quantiles, denoted $Q(\cdot)$, are computed by first constructing an empirical inverse cumulative distribution function, $F^{-1}(x)$, from the data set and then evaluating $F^{-1}(0.05)$, $F^{-1}(0.25)$, $F^{-1}(0.50)$ (median), $F^{-1}(0.75)$, and $F^{-1}(0.95)$. The Night, Day and Mixed User data are expressed in percent of average daily load.

Table 16
Hourly distribution statistics — night users ($n = 18$).

Hr	Min. (%)	Q(5) (%)	Q(25) (%)	Q(50) (%)	Q(75) (%)	Q(95) (%)	Max (%)	μ (%)	σ (%)	γ	β
1	0.0	0.0	0.0	0.0	0.0	11.1	32.0	1.2	3.6	3.81	20.1
2	0.0	0.0	0.0	0.0	0.0	6.2	38.7	0.9	3.1	5.61	49.1
3	0.0	0.0	0.0	0.0	0.0	5.3	18.2	0.7	2.5	4.50	23.9
4	0.0	0.0	0.0	0.0	0.0	2.8	18.1	0.6	2.3	5.13	30.2
5	0.0	0.0	0.0	0.0	0.0	4.3	33.4	0.6	2.5	7.65	81.9
6	0.0	0.0	0.0	0.0	0.0	8.3	34.6	1.3	4.3	4.77	28.2
7	0.0	0.0	0.0	0.0	0.7	7.4	47.7	1.3	3.9	6.08	54.9
8	0.0	0.0	0.0	0.0	0.7	8.3	35.4	1.5	4.2	4.89	32.3
9	0.0	0.0	0.0	0.0	0.3	6.7	23.5	1.1	2.9	3.73	19.8
10	0.0	0.0	0.0	0.0	0.0	8.0	39.6	1.3	3.9	5.11	37.0
11	0.0	0.0	0.0	0.0	0.3	10.6	189.1	2.1	10.0	13.58	234.4
12	0.0	0.0	0.0	0.0	0.3	8.2	86.9	1.5	5.2	9.60	136.8
13	0.0	0.0	0.0	0.0	0.2	7.4	57.5	1.4	4.3	6.70	67.9
14	0.0	0.0	0.0	0.0	0.0	6.7	36.1	1.1	3.3	5.04	38.2
15	0.0	0.0	0.0	0.0	0.0	7.1	70.9	1.4	5.0	8.49	97.3
16	0.0	0.0	0.0	0.0	0.0	6.0	22.9	1.0	2.7	4.12	24.7
17	0.0	0.0	0.0	0.0	1.2	8.4	56.1	1.6	4.6	6.06	54.2
18	0.0	0.0	0.0	4.9	12.6	34.7	84.2	8.8	11.8	2.21	9.6
19	0.0	0.0	5.7	17.6	31.2	42.4	94.9	19.1	16.1	1.04	5.1
20	0.0	0.0	6.9	19.0	30.2	45.9	97.6	19.9	16.6	1.30	6.4
21	0.0	0.0	0.0	11.2	24.8	43.9	80.5	14.8	14.7	0.98	3.7
22	0.0	0.0	0.0	3.3	14.3	33.3	96.0	8.8	12.1	2.00	9.0
23	0.0	0.0	0.0	0.0	6.6	24.7	84.8	4.6	9.0	3.24	19.3
24	0.0	0.0	0.0	0.0	1.0	14.8	38.8	2.4	5.6	2.94	12.9

Table 17
Hourly distribution statistics — day users ($n = 7$).

Hr	Min. (%)	Q(5) (%)	Q(25) (%)	Q(50) (%)	Q(75) (%)	Q(95) (%)	Max (%)	μ (%)	σ (%)	γ	β
1	0.0	0.0	0.0	0.0	0.0	0.8	24.0	0.6	3.3	6.07	39.1
2	0.0	0.0	0.0	0.0	0.0	0.3	21.7	0.5	2.9	6.13	39.9
3	0.0	0.0	0.0	0.0	0.0	0.5	20.4	0.3	2.1	7.37	59.9
4	0.0	0.0	0.0	0.0	0.0	0.5	10.8	0.2	0.9	8.90	93.1
5	0.0	0.0	0.0	0.0	0.0	0.1	12.5	0.1	0.9	12.26	163.4
6	0.0	0.0	0.0	0.0	0.0	3.1	9.4	0.3	1.1	4.85	30.1
7	0.0	0.0	0.0	0.0	0.0	12.5	34.9	1.5	4.5	4.07	22.4
8	0.0	0.0	0.0	0.0	2.7	25.5	220.2	4.4	16.8	10.32	130.0
9	0.0	0.0	0.0	0.0	3.7	34.4	213.3	6.6	19.3	6.70	65.1
10	0.0	0.0	0.0	0.0	4.9	43.7	103.2	6.9	15.3	3.08	13.9
11	0.0	0.0	0.0	0.0	9.4	38.2	96.3	8.2	16.4	2.93	12.6
12	0.0	0.0	0.0	0.0	10.1	39.3	195.4	9.2	23.0	5.39	39.6
13	0.0	0.0	0.0	0.0	6.9	34.6	168.3	8.1	18.9	4.69	32.9
14	0.0	0.0	0.0	0.0	6.2	32.5	251.5	7.7	22.6	7.33	72.0
15	0.0	0.0	0.0	0.0	3.9	30.0	96.3	5.1	12.8	4.16	24.5

(*continued on next page*)

Table 17 (*continued*)

Hr	Min. (%)	Q(5) (%)	Q(25) (%)	Q(50) (%)	Q(75) (%)	Q(95) (%)	Max (%)	μ (%)	σ (%)	γ	β
16	0.0	0.0	0.0	0.0	3.1	28.0	264.9	6.4	25.3	7.60	68.8
17	0.0	0.0	0.0	0.0	3.2	25.0	230.5	5.2	18.9	8.74	98.1
18	0.0	0.0	0.0	0.0	3.8	50.9	182.4	7.1	20.3	4.88	33.2
19	0.0	0.0	0.0	0.0	4.4	40.5	76.8	7.0	15.9	2.90	11.0
20	0.0	0.0	0.0	0.0	2.8	37.4	59.9	5.8	12.7	2.31	7.5
21	0.0	0.0	0.0	0.0	0.2	31.3	54.2	4.6	11.1	2.44	8.0
22	0.0	0.0	0.0	0.0	0.0	26.2	34.7	2.4	7.4	3.13	11.4
23	0.0	0.0	0.0	0.0	0.0	2.7	28.6	1.0	4.7	5.16	28.5
24	0.0	0.0	0.0	0.0	0.0	1.7	27.6	0.7	4.0	5.87	36.8

Table 18
Hourly distribution statistics — mixed users ($n = 17$).

Hr	Min. (%)	Q(5) (%)	Q(25) (%)	Q(50) (%)	Q(75) (%)	Q(95) (%)	Max (%)	μ (%)	σ (%)	γ	β
1	0.0	0.0	0.0	0.0	0.0	2.8	28.6	0.5	2.6	6.22	47.9
2	0.0	0.0	0.0	0.0	0.0	0.6	32.7	0.4	2.5	7.99	79.4
3	0.0	0.0	0.0	0.0	0.0	0.0	16.4	0.3	1.8	6.99	54.4
4	0.0	0.0	0.0	0.0	0.0	0.5	20.4	0.3	1.9	8.03	75.2
5	0.0	0.0	0.0	0.0	0.0	3.0	16.4	0.5	2.0	5.60	37.7
6	0.0	0.0	0.0	0.0	1.1	8.6	31.6	1.3	3.2	3.65	22.4
7	0.0	0.0	0.0	0.0	1.2	8.5	37.9	1.4	3.5	4.48	33.3
8	0.0	0.0	0.0	0.0	3.1	12.8	36.3	2.5	4.9	3.01	14.0
9	0.0	0.0	0.0	0.0	3.5	16.2	82.8	3.2	6.5	4.96	48.2
10	0.0	0.0	0.0	0.0	4.2	18.8	61.6	3.5	6.7	3.11	17.3
11	0.0	0.0	0.0	0.0	5.0	17.3	51.5	3.7	6.9	2.83	13.5
12	0.0	0.0	0.0	0.0	4.9	16.5	76.1	3.8	7.3	3.71	25.3
13	0.0	0.0	0.0	0.0	5.8	16.3	80.6	4.0	7.7	4.27	31.6
14	0.0	0.0	0.0	0.0	4.7	18.1	54.8	3.5	6.6	3.00	14.7
15	0.0	0.0	0.0	0.0	3.5	15.0	66.9	3.3	6.9	4.27	28.6
16	0.0	0.0	0.0	0.0	3.6	14.8	74.6	3.1	6.7	4.43	33.6
17	0.0	0.0	0.0	0.0	4.1	14.6	34.1	3.1	5.3	2.43	10.1
18	0.0	0.0	0.0	3.8	10.8	34.9	87.7	8.2	12.2	2.63	11.7
19	0.0	0.0	2.1	14.3	21.5	40.6	109.6	16.3	16.8	2.37	11.9
20	0.0	0.0	0.5	14.5	20.8	49.1	109.6	16.5	17.8	2.37	11.1
21	0.0	0.0	0.0	6.8	15.4	40.0	98.6	10.8	15.5	2.71	12.2
22	0.0	0.0	0.0	0.0	5.4	21.1	87.7	5.2	12.2	4.30	25.2
23	0.0	0.0	0.0	0.0	0.0	15.6	82.1	2.6	8.8	6.11	48.8
24	0.0	0.0	0.0	0.0	0.0	7.4	62.4	1.1	4.9	7.31	73.3

Table 19
Hourly aggregate distribution statistics — Entesopia.

Hr	Min. (Wh)	Q(5) (Wh)	Q(25) (Wh)	Q(50) (Wh)	Q(75) (Wh)	Q(95) (Wh)	Max (Wh)	μ (Wh)	σ (Wh)	γ	β
1	7.0	14.8	30.0	68.5	119.0	223.0	223.0	83.6	66.0	0.89	2.7
2	0.0	7.2	18.0	60.5	99.0	250.6	388.0	75.5	84.0	2.44	9.8
3	4.0	7.0	14.0	28.3	69.3	152.6	158.0	47.2	45.4	1.33	3.7
4	9.0	9.6	11.0	19.0	27.0	59.8	97.0	22.9	18.7	2.92	12.3
5	0.0	0.6	7.0	15.0	21.0	61.8	78.0	19.1	18.1	1.81	6.3
6	0.0	0.0	11.0	18.5	35.0	143.4	174.0	31.1	41.0	2.52	8.7
7	8.0	8.6	26.0	65.0	133.0	567.8	848.0	126.0	185.5	2.94	11.7
8	3.0	6.0	69.0	99.5	269.0	640.2	717.0	180.9	190.8	1.53	4.5
9	2.0	5.0	79.3	160.5	224.0	630.8	635.0	205.7	193.6	1.27	3.4
10	6.0	19.2	106.0	173.5	346.0	938.8	1162.0	278.7	294.2	1.62	4.9
11	57.0	75.6	146.0	296.5	755.0	1163.8	1231.0	410.8	344.2	1.08	3.0
12	59.0	80.6	186.0	367.5	632.0	1240.0	1507.0	448.7	361.4	1.42	4.6
13	93.0	148.8	233.0	409.0	528.0	971.8	1018.0	436.3	268.7	1.01	2.8
14	11.0	69.8	223.0	376.5	689.0	1081.6	1294.0	471.2	324.4	0.79	2.9
15	17.0	39.8	203.0	285.0	529.0	993.8	1211.0	374.0	285.7	1.31	4.5
16	129.0	134.4	206.0	355.5	469.0	907.6	1015.0	391.1	233.9	1.04	3.6
17	25.0	119.8	246.0	366.5	438.0	664.0	706.0	363.8	154.6	0.20	3.2
18	115.0	308.8	627.0	669.5	815.0	1279.8	1326.0	731.5	263.4	0.40	4.0
19	190.0	425.8	609.0	749.5	808.0	925.0	928.0	717.0	162.3	−1.44	6.0
20	175.0	178.6	516.0	608.5	723.0	978.2	1031.0	589.5	237.5	−0.19	2.4
21	181.0	181.0	243.0	404.0	548.0	757.8	792.0	423.0	192.5	0.30	1.9
22	96.0	114.0	153.0	238.0	385.0	572.8	583.0	285.1	156.8	0.67	2.1
23	35.0	56.6	100.0	148.0	253.0	382.6	412.0	179.0	105.0	0.73	2.4
24	17.0	21.2	57.0	98.0	128.0	321.4	358.0	119.3	90.7	1.40	4.1

Table 20
Hourly aggregate distribution statistics — Barsaloi.

Hr	Min. (Wh)	Q(5) (Wh)	Q(25) (Wh)	Q(50) (Wh)	Q(75) (Wh)	Q(95) (Wh)	Max (Wh)	μ (Wh)	σ (Wh)	γ	β
1	0.0	0.0	2.0	8.0	21.0	52.5	66.0	14.4	17.0	1.60	4.9
2	0.0	0.0	1.5	5.0	9.0	18.3	25.0	6.3	6.1	1.23	4.5
3	0.0	0.0	0.0	2.0	6.3	11.3	15.0	3.7	4.5	0.89	2.7
4	0.0	0.0	0.0	2.0	6.3	15.3	16.0	3.9	4.9	1.26	3.5
5	0.0	0.0	2.0	7.0	16.3	33.3	37.0	10.3	11.0	1.18	3.2
6	12.0	12.8	20.5	29.0	49.5	65.3	66.0	34.2	17.1	0.51	1.9
7	8.0	10.3	20.0	31.0	49.8	74.8	77.0	33.9	19.9	0.78	2.6
8	14.0	14.8	38.8	54.0	77.7	131.7	158.0	61.0	34.9	0.90	3.7
9	17.0	27.5	36.3	63.0	102.0	166.0	229.0	73.6	47.3	1.50	5.6
10	26.0	26.8	40.8	73.0	99.5	155.5	241.0	77.5	45.3	1.83	7.8
11	27.0	31.5	46.0	75.0	104.5	170.5	235.0	81.0	45.1	1.58	6.5
12	25.0	26.5	45.8	67.0	96.0	151.3	182.0	74.6	36.9	1.07	4.1
13	16.0	22.8	48.0	59.0	83.0	158.3	240.0	71.2	45.0	2.24	9.0
14	13.0	13.0	22.5	43.0	64.2	135.3	193.0	49.8	38.9	2.16	8.5
15	15.0	15.8	22.0	37.0	61.0	132.8	156.0	47.7	36.2	1.54	4.7
16	11.0	11.0	27.3	39.0	74.2	123.0	126.0	49.4	32.6	0.90	3.0
17	14.0	14.0	24.0	36.0	53.5	96.5	119.0	43.2	26.1	1.19	4.0
18	46.0	46.8	65.8	102.0	152.5	208.0	214.0	110.7	51.2	0.58	2.2
19	234.0	245.6	301.3	353.0	390.0	478.0	484.0	354.6	68.3	0.12	2.4
20	287.0	296.0	391.5	414.0	479.0	550.0	553.0	419.8	71.3	0.03	2.5
21	215.0	221.8	294.0	357.0	385.8	432.2	457.0	339.7	65.4	−0.24	2.2
22	95.0	105.5	159.3	198.0	219.0	266.8	281.0	190.3	48.4	−0.19	2.4
23	15.0	31.5	61.5	90.0	125.3	184.0	286.0	97.9	52.7	1.70	7.6
24	4.0	4.8	16.3	32.0	65.3	102.3	148.0	42.4	33.5	1.28	4.9

References

Action P. Poor people's energy outlook 2016. Rugby, UK: Practical Action Publishing; 2016.

Alzola J, Vechiu I, Camblong H, Santos M, Sall M, Sow G. Microgrids project, Part 2: design of an electrification kit with high content of renewable energy sources in Senegal. Renew Energy 2009;34:2151–9.

Askari IB, Ameri M. Techno-economic feasibility analysis of stand-alone renewable energy systems (PV/bat, Wind/bat and Hybrid PV/wind/bat) in Kerman, Iran. Energy Sources Part B 2012;7:45–60.

Azimoh CL, Klintenberg P, Mbohwa C, Wallin F. Replicability and scalability of mini–grid solution to rural electrification programs in sub-Saharan Africa. Renew Energy 2017;106:222–31.

Bekele G, Tadesse G. Feasibility study of small Hydro/PV/Wind hybrid system for off–grid rural electrification in Ethiopia. Appl Energy 2012;97:5–15. Energy Solutions for a Sustainable World — Proceedings of the Third International Conference on Applied Energy, May 16–18, 2011- Perugia, Italy.

Birol F. World energy outlook.2015.

Birol F. World energy outlook.2016.

Blanchard R, Little M. Developing an open access monitoring device for off-grid renewables. Development in the in Renewable Energy Technology (ICDRET), 2016 4th International Conference on the, IEEE. 2016. p. 1–6.

Boait P, Advani V, Gammon R. Estimation of demand diversity and daily demand profile for off-grid electrification in developing countries. Energy Sustain Dev 2015;29:135–41.

Bradburn N. Respondent burden. Proceedings of the Survey Research Methods Section of the American Statistical Association. vol. 35. 1978. p. 40.

Brook PJ, Smith S. Energy and development report 2000: energy services for the world's poor. World Bank; 2000.

Camblong H, Sarr J, Niang A, Curea O, Alzola J, Sylla E, Santos M. Micro-grids project, Part 1: analysis of rural electrification with high content of renewable energy sources in Senegal. Renew Energy 2009;34:2141–50.

Díaz P, Peñ R, Muñoz J, Arias C, Sandoval D. Field analysis of solar pv-based collective systems for rural electrification. Energy 2011;36:2509–16.

Ellegård A, Nordström M. Rural energy service companies — experiences from Zambia. Stockholm: Stockholm Environment Institute, SEI, 2001,52.

Administration Energy Information. Buildings energy data book.2012.

ESMAP, Global lighting services for the poor Phase II: test marketing of small "solar" batteries for rural electrification purposes. World Bank; 1999.

Fabini D, Baridó DPL, Omu A, Taneja J. Mapping induced residential demand for electricity in Kenya. Proceedings of the Fifth ACM Symposium on Computing for Development. 2014. p. 43–52.

Freeman J, Butler EW. Some sources of interviewer variance in surveys. Public Opin Q 1976;40:79.

Ghafoor A, Munir A. Design and economics analysis of an off-grid PV system for household electrification. Renew Sustain Energy Rev 2015;42:496–502.

GIZ, What size shall it be? A guide to mini-grid sizing and demand forecasting. 2016.

Government of New Zealand, Stand-alone power systems — part 2: system design. Technical Report AS/NZS4509.2:2010, Standards New Zealand.; 2010.

Gustavsson M. With time comes increased loads — an analysis of solar home system use in Lundazi, Zambia. Renew Energy 2007;32:796–813.

Howell DC. Statistical methods for psychology. 7th ed. ed., Wadsworth; 2010.

IEEE Working Group for Energy Storage Subsystems, IEEE guide for array and battery sizing in stand-alone photovoltaic (PV) systems. IEEE Std. 1562-2007, 2007a.

IEEE Working Group for Energy Storage Subsystems, IEEE recommended practice for sizing lead-acid batteries for stand-alone photovoltaic (PV) systems. IEEE Std. 1013-2007, 2007b.

Inversin AR. Mini-grid design manual. World Bank; 2000.

Corp JBL. Speaker power requirements. Tech. Note rev 7-90, 2017.Accessed: Mar. 2017.

Kish L. Studies of interviewer variance for attitudinal variables. J Am Stat Assoc 1962;57:92–115.

Kolhe ML, Ranaweera KIU, Gunawardana AS. Techno-economic sizing of off-grid hybrid renewable energy system for rural electrification in Sri Lanka. Sustainable Energy Technol Assess 2015;11:53–64.

Krithika P, Palit D. Participatory business models for off-grid electrification. Rural electrification through decentralised off-grid systems in developing countries. Springer.2013. p. 187–225.

Krosnick JA. Response strategies for coping with the cognitive demands of attitude measures in surveys. Appl Cogn Psychol 1991;5:213–36.

Lawrence Berkeley National Laboratory, Stand-by power.2017, http://standby.lbl.gov/summary-table.html. Accessed: Mar. 2017.

Louie H, Dauenhauer P. Effects of load estimation error on small-scale off–grid photovoltaic system design, cost and reliability. Energy Sustainable Dev 2016;34:30–43.

Louie H, Goldsmith G, Dauenhauer P, Almeida RH. Issues and applications of real-time data from off-grid electrical systems. PowerAfrica, 2016 IEEE PES. 2016. p. 88–92.

Louw K, Conradie B, Howells M, Dekenah M. Determinants of electricity demand for newly electrified low-income African households. Energy Policy 2008;36:2812–8.

Mandelli S, Brivio C, Colombo E, Merlo M. Effect of load profile uncertainty on the optimum sizing of off-grid PV systems for rural electrification. Sustainable Energy Technol Assess 2016a;18:34–47.

Mandelli S, Merlo M, Colombo E. Novel procedure to formulate load profiles for off-grid rural areas. Energy Sustainable Dev 2016b;31:130–42.

Meier P, Voravate T, Barnes DF, Bogach SV, Daniel F. Peru — national survey of rural household energy use. World Bank; 2010.

Nerini FF, Broad O, Mentis D, Welsch M, Bazilian M, Howells M. A cost comparison of technology approaches for improving access to electricity services. Energy 2016;95:255–65.

Nouni M, Mullick S, Kandpal T. Providing electricity access to remote areas in India: an approach towards identifying potential areas for decentralized electricity supply. Renew Sustainable Energy Rev 2008;12:1187–220.

Olatomiwa L, Mekhilef S, Huda A, Ohunakin OS. Economic evaluation of hybrid energy systems for rural electrification in six geo-political zones of Nigeria. Renew Energy 2015;83:435–46.

Rajbongshi R, Borgohain D, Mahapatra S. Optimization of PV-biomass-diesel and grid base hybrid energy systems for rural electrification by using HOMER. Energy 2017;126:461–74.

Ramchandran N, Pai R, Parihar AKS. Feasibility assessment of Anchor-Business–Community model for off-grid rural electrification in India. Renew Energy 2016;97:197–209.

Rao N, Agarwal A, Wood D. Impacts of small-scale electricity systems: a study of rural communities in India and Nepal. Technical Report, 2016.

Schwarz N. Self-reports in consumer research: the challenge of comparing cohorts and cultures. J Consumer Res 2003;29:588.

Sen R, Bhattacharyya SC. Off-grid electricity generation with renewable energy technologies in India: an application of HOMER. Renew Energy 2014;62:388–98.
SteamaCo, About-SteamaCo.2016, http://steama.co/about/. Accessed: Mar. 2017.
Ustun TS. Computer-based decision making systems for improved energy access in sub-Saharan Africa. J Autom Control Eng 2016;4:247–51.
Warner SL. Randomized response: a survey technique for eliminating evasive answer bias. J Am Stat Assoc 1965;60:63–9.
World Bank, Monitoring and evaluation in rural electrification projects: a demand-oriented approach. World Bank; 2003.
World Bank, Addressing the electricity access gap. Report, World Bank.; 2012.
World Health Organization and others, Access to modern energy services for health facilities in resource-constrained settings: a review of status, significance, challenges and measurement. Technical Report, 2014.
Zeyringer M, Pachauri S, Schmid E, Schmidt J, Worrell E, Morawetz UB. Analyzing grid extension and stand-alone photovoltaic systems for the cost-effective electrification of Kenya. Energy Sustainable Dev 2015;25:75–86.