

Effects of Load Estimation Error on Small-Scale Off-Grid Photovoltaic System Design, Cost and Reliability

Henry Louie

Copperbelt University, Kitwe, Zambia

Peter Dauenhauer

University of Strathclyde, Glasgow, Scotland

Abstract

The proliferation of off-grid photovoltaic (PV) systems is rapidly increasing in the least developed countries. The sizing of system components—primarily PV panels and batteries—is critically influenced by the expected daily load. However, accurately estimating incipient electrical load of rural consumers is fraught with challenges. Load estimation error is propagated through the design phase, potentially resulting in a system that is unduly expensive or fails to meet reliability targets. This article investigates the effects of daily load estimation error on system design, cost and reliability. Load and insolation data from seven off-grid systems in Malawi were collected. The systems were redesigned using three different intuitive design approaches considering different levels of load estimation error, ranging from $\pm 90\%$ of the actual measured load. The cost of each design is estimated from in-country prices. The reliability of each design is determined from an hourly simulation using the measured data. The results show that PV array and battery sizing scale proportionately with load estimation error and that the cost of load over-estimation is approximately US\$1.92 to US\$6.02 per watt-hour, whereas under-estimation can precipitously degrade reliability. A cost-versus-reliability analysis shows that for the Malawi systems, on average 46% of the PV and battery costs are used to improve the simulated hourly

Email addresses: hlouie@ieee.org (Henry Louie),
peter.dauenhauer@strath.ac.uk (Peter Dauenhauer)

Preprint submitted to Energy for Sustainable Development

April 10, 2018

reliability from 99% to 100%. Moreover, the results point to the challenges with intuitive design approaches, showing that consideration of average load alone can lead to over- or under-designed systems.

Keywords: Microgrids, solar power, reliability, rural electrification.

1. Introduction

As is articulated by the UN Global Goal Seven, universal energy access is a critical global objective [1]. Sufficient access to clean and sustainable energy underpins many if not all development objectives in least developed countries (LDCs). Achieving universal access is an ongoing struggle for LDCs, where national electrification rates are less than 10 percent, and less than 1 percent in rural areas [2]. Different solutions are being proposed, piloted, and rolled-out by a wide-range of actors: multilateral aid agencies, individual countries, private companies, civil society, universities, communities, and individuals. Progress is being made, since 2000, over 145 million people in Africa alone gained access to electricity [3].

Off-grid renewable energy projects utilizing photovoltaics (PV), wind energy, biomass or hydro deployed in stand-alone systems or mini-grids are a widely promoted solution to universal energy access. The International Energy Agency has estimated that off-grid solutions will provide 59% of first-time access to electricity [4], requiring US\$50B per year invested until 2030. The stakes are high for these projects; failure to address the ongoing sustainability challenges will undermine these investments and impacts they can have.

By definition, off-grid systems have no connection to the national grid. As such, they must be designed to independently balance energy supply with the anticipated load over the short and long term. The PV array must be sized appropriately to satisfy the load (inclusive of losses), and the battery capacity must be sufficient to buffer against periods of decreased insolation, and increased load. Over time, batteries and PV arrays suffer from both aging effects and load increases, putting further pressure on the system to deliver an adequate level of reliability.

Approaches to sizing the PV array, battery and other components vary in their sophistication from simple *ad-hoc* ('rule of thumb') methods to the use of simulation-based computer programs (e.g. HOMER, Hybrid2, PVSyst) which optimize cost or reliability. In the literature, several approaches to PV

system sizing have been established and are classified as intuitive, numerical, and analytical [5, 6].

The intuitive method, as highlighted within this article, involves a simplified set of calculations for the PV sub-systems to reduce the modeling complexity. Simplifications include modeling solar radiation based on the single worst month in the year, ignoring dynamics of the charging cycle, and disregard a reliability optimization. Daily load profiles are commonly reduced to a single value, the average daily energy use, despite research showing the impact different load profiles have on reliability and cost of the designed system [7, 8, 9]. Various iterations on the method have been documented throughout the literature [10, 11] while the specific sizing approach investigated in this paper is found in [12, 13].

Numerical methods involve detailed simulation of energy supply and load over a specified length with the objective of minimizing costs or loss of load probability. Numerical solutions will typically involve modeling stochastic elements such as in [14, 15].

Analytical techniques optimize reliability by adjusting array and battery size using deterministic input data as shown in [16, 17, 18, 19]. These methods seek to model the reliability of the system mathematically. Under certain assumptions, closed form equations can be derived [20, 21]; other approaches rely on fitting parametric equations to simulated data [22], [23].

Both the numerical or analytical techniques will produce more exacting sizing than the intuitive approaches. However, intuitive methods remain widespread among practitioners in LDCs due to the tractability of the method, ability to provide rough but seemingly accurate results, and lack of sufficient data to support more advanced techniques. Intuitive methods in particular are more often applied to small-scale systems—typically less than 2 kW—where the additional effort of numerical or analytic design approaches may not be justified.

Regardless of the sizing approach, the resulting design is intimately tied to and dependent on the estimate of average daily load, which is notoriously difficult to estimate in the context of LDCs [24, 25, 26, 27]. Constant daily loads are often assumed within the optimization techniques for simplicity despite the uncertain and variable nature of newly electrified customers. The problem can be addressed from three angles. First, the load estimate can be assumed based on past experience of similar installations [28]. However, these data are rare, not widely available and perhaps not generalizable. Second, an econometric model with existing consumers that estimates, for example,

the demographic variables correlating to energy consumption, can be used to hypothesize loads for potential future locations [29, 30]. The selection of the predictor variables is non-trivial, for example [31] showed there exists wide variety in consumption patterns even within the same customer class. Third, a bottom-up approach can be used to build up an aggregate load from expected appliance duty cycles and various customer classifications, often employing a field survey as a basis for current and future load.

Surveys of ‘aspirational’ load can be conducted to provide insight into what electric appliances and loads a consumer anticipates on purchasing and how frequently they would be used [25, 32, 33]. The ability of an individual without prior access to electricity and often irregular income to accurately predict their future appliance purchases and usage pattern is questionable. There is little, if any, rigorous research on the accuracy and potential biases of the survey method, but there is at least anecdotal evidence within the practitioner community that surveys can be inaccurate and unreliable [34?]. Loads can be added or removed over time, and human behavior is difficult to predict. For example, it has been reported that on average, a person switches electric devices on and off hundreds of times per day, mostly unaware that they are doing so [35]. Not surprisingly, researchers have called for improved methods for estimating load [24, 26, 27], but a practical and proven approach has yet to emerge.

Errors in the estimation of the average daily load propagate through the design phase, resulting in systems perhaps ill-suited to their application. The practical consequences are considerable and are most tangibly presented in the form of increased or decreased cost or reliability. Acute over-estimation of load may yield extraordinarily reliable systems with larger than needed PV arrays and batteries at an exorbitant cost. Managers of often insufficient rural electrification budgets may prefer to install a greater number of off-grid systems with lower—but still acceptable—reliability than fewer at high reliability. On the other hand, under-estimation of load may lead to inexpensive under-sized systems that fail to meet reliability targets.

The reliability of PV systems has been the subject of several research studies going back to at least the 1970’s. Early work was focused on component reliability [36, 37]. Contemporary research tends to focus on grid-tied rather than off-grid systems. The research often seeks to develop new methods of conceptualizing or calculating reliability, often using probabilistic approaches. A Markov Reward Model was developed in [38] to incorporate reliability into grid-tied PV performance analysis. In [39], the authors present

a new method to incorporate reliability into the levelized cost of energy of grid-tied PV systems. Methods for computing the reliability of off-grid systems using the loss of load probability metric are derived in [20, 21] based on probabilistic models. Off-grid systems are also considered in [40], where the authors combine reliability considerations into the design of such systems. Other existing research considers the reliability of hybrid systems [41] and clustered microgrids [42] using various probabilistic techniques. No existing research directly investigates the the impact of load estimation error on reliability or associates system cost with the error.

This article takes a practical, data-driven approach to investigating the implications of average daily load estimation error on small-scale off-grid PV systems. Hourly insolation and load data from seven real-world systems in Malawi were collected over the course of approximately one year. With the actual average daily load known, several intentional over- and under-estimations of the load were made and used as inputs to hypothetically redesign the systems using various design approaches. The corresponding costs are estimated using a model derived from in-country pricing. The reliability of the redesigned systems are evaluated through a deterministic simulation. The simulation uses the collected data to recreate the real-world insolation and load conditions experienced at the Malawian sites. The use of a deterministic approach rather than probabilistic is a unique aspect of this research. Rather than relying on theoretical models of load and insolation, the use of actual data completely captures the potential complex correlation and dependency structures within and among the energy flows.

The main contribution of this research is the quantification of the sensitivity of cost and reliability to average daily load estimation error. In addition, an opportunistic enquiry into the relationship between cost and reliability independent of load estimation error is made. The research demonstrates the value of accurate load estimation. Broader insight is also gained in how following different design approaches affect system cost and reliability. The results highlight an important disadvantage of intuitive design approaches: considering only average daily load and not its distribution or temporal characteristics can result in over- or under-designed systems.

The remainder of this paper is arranged as follows. The Section 2 provides information on the Malawian systems and analyzes the characteristics of the collected data. The research methodology is presented in Section 3. The re-designed systems are presented and discussed in the Section 4. The cost and reliability of the re-designed systems are computed and discussed in

Sections 5 and 6. An analysis of the relationship of these results is performed in Section 7. Conclusions and future outlook are provided in Section 8.

2. Site and data set descriptions

The data analyzed in this paper are from small-scale off-grid PV systems at seven different sites in Malawi [43]. The systems were installed in educational and health facilities in the nearby Gumbwa and Ndakwera villages using an intuitive *ad-hoc* design approach. Each village has several stand-alone systems—Gumbwa is comprised of two classroom systems, a health post, and a staff room while Ndakwera has one classroom, a maternity ward and an out-patient department (OPD). The systems consisted of the standard elements for an off-grid PV system: PV array, deep cycle lead acid batteries, charge controller, and an inverter.

The systems were outfitted with data acquisition equipment that measured battery voltage and branch currents. The data acquisition system utilized a wireless sensor network to communicate locally and with a central server as described in [44]. System data was captured, stored locally on an SD card, and transmitted over SMS with a GPRS module. The wireless sensor network was implemented with WaspNet on the Zigbee communication protocol [? ?]. The deployment featured in this paper demonstrates the analytical value from the use of advanced data acquisition equipment for off-grid systems in LDCs [?]. The load was post-computed by multiplying the battery voltage by the total load current as measured on the DC side of the inverter. The load values therefore are inclusive of inverter losses. A weather station in Gumbwa measured meteorological quantities including solar irradiance incident to the rooftop PV panels. Although measured in Gumbwa, the meteorological data are generally applicable to Ndakwera as the two villages are nearby.

2.1. Data processing

Table 1 provides the data set start and end date, as well as the sampling frequency for each site. Like all real-world data sets, the sampled data required processing and cleaning prior to analysis. Missing values occasionally occurred, but the raw data set was more than 90% complete. After converting the data into hourly averages, the data set consisted of 57,222 hourly load values.

Table 1: Data Set Description

Site	Start Date	End Date	Samples Per Min.	Sim. Days
Gumbwa Classroom 1	27/1/2014	31/1/2015	60	291
Gumbwa Classroom 2	27/1/2014	31/1/2015	60	291
Gumbwa Health Post	27/1/2014	31/1/2015	30	289
Gumbwa Staff Room	27/1/2014	31/1/2015	15	291
Ndakwera Classroom	8/6/2014	4/3/2015	15	198
Ndakwera Mat. Ward	9/6/2014	4/3/2015	15	198
Ndakwera OPD	9/6/2014	4/3/2015	15	141
Insolation (all sites)	27/1/2014	31/1/2015	60	—

Table 2: Average Actual and Estimated Daily Load

	Gumbwa				Ndakwera		
	Class-room 1	Class-room 2	Health Post	Staff room	Class-Room	Mat. Ward	OPD
Actual (Wh)	91.5	161.3	308.5	104.9	180.8	388.4	518.6
Estimate (Wh)	758	608	662	455	905	1064	1064
Estimate Error (%)	+728	+277	+115	+334	+400	+174	+105

The load data for each site was grouped into 24-hour subsets, each corresponding to a calendar day. Any 24-hour subset with more than four hours of missing data, consecutive or otherwise, was discarded. The missing values for the remaining 24-hour subsets were synthesized through linear interpolation. As discussed later, the reliability simulation requires synchronized hourly insolation and load data. The number of days with both insolation and load data is shown in the last column of Table 1.

2.2. Load characteristics

The actual average daily load for each site is provided in Table 2. The averages range from 91.5 Wh to 518.6 Wh. For comparison, the estimated load as determined through pre-installation surveying [44], and assuming an inverter efficiency of 90%, is shown in the second row of the Table. The estimation errors are striking—as high as +728%—with an average of +305%. In all cases, the bias was toward over-estimation of load.

Figure 1 provides histograms of the daily load data. Each histogram is

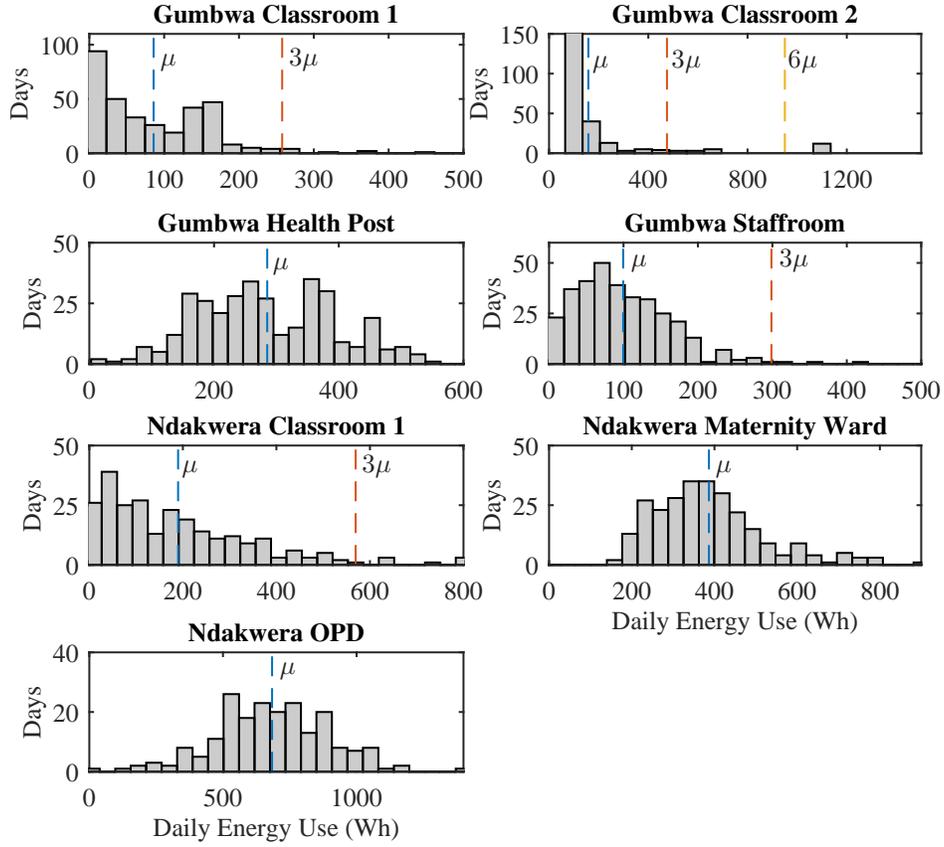


Figure 1: Histogram of daily load with vertical lines indicating the location of the average and select multiples of the average.

unique and it is evident that there is no prototypical load distribution.

Most intuitive design approaches size the usable capacity of the battery by multiplying the estimated average load by a “days of autonomy” factor—the number of days the average load can be supplied without PV input. A system designed with three days of autonomy might suffer an outage on days whose load is greater than three times the average, depending on the insolation and its co-incidence with the load. For reference, the average daily load μ , and three and six times μ are shown in Fig. 1 as the vertical lines, as the abscissa limits allow.

A fundamental problem with designing a system using an intuitive method is that the design is based on the average daily load. The load is characterized as a single parameter, ignoring its distribution and any temporal relation-

ships of the load. Consider for example the Ndakwera OPD in Fig. 1. The distribution appears approximately Gaussian with a relatively small standard deviation. Contrast this with the Gumbwa Classroom 2 where there are several days where the load exceeds 6μ . Additional battery capacity is needed to cope with the days of extreme load.

Fig. 2 shows the daily load of the sites as a time series. The average load, and multiples thereof are shown as the horizontal lines. For Gumbwa Classroom 2, the days in which the load exceeds 6μ occur just several days apart, with higher than average (above 3μ) loads occurring in between them. Again, this necessitates additional battery capacity to achieve high reliability. The effects of load distribution and temporal characteristics are explored more deeply in the Illustrative Critique of the Intuitive Method section.

2.3. Insolation characteristics

The insolation data are used in the reliability simulation of the re-designed systems. The insolation data were processed in the same way as the load data, with details provided in the last row of Table 1, and with the histogram shown in Fig. 3. The average daily insolation by month is provided in Fig. 4. The month with the lowest average daily insolation is July, at $5.6 \text{ kWh/m}^2/\text{day}$. This quantity is used in intuitive methods as the assumed insolation.

3. Methodology overview

The measured load and insolation data are used to investigate how error in the estimation of daily average load influences design, cost and reliability of small-scale off-grid PV systems. Relevant systems to this research are classified within [45] and include decentralized stand-alone systems and micro-grids where solar PV is the sole generation source.

In the context of this research, the “design” of a system explicitly refers to the total size or rating of the PV array (in watts) and battery (in watthours). The sizing of these two components are most directly affected by load estimation and, in the authors’ experience, tend to comprise 40 to 60% of total system component costs, inclusive of balance of system components. Therefore, “cost” refers to the component cost of the PV array and battery, in US Dollars. In the context of this research, the “reliability” is the ability of a system to provide the demanded load each hour—considering only energy balance constraints—and not, for example, equipment failure or malfunction. By focusing in this manner, the analysis remains tractable while still

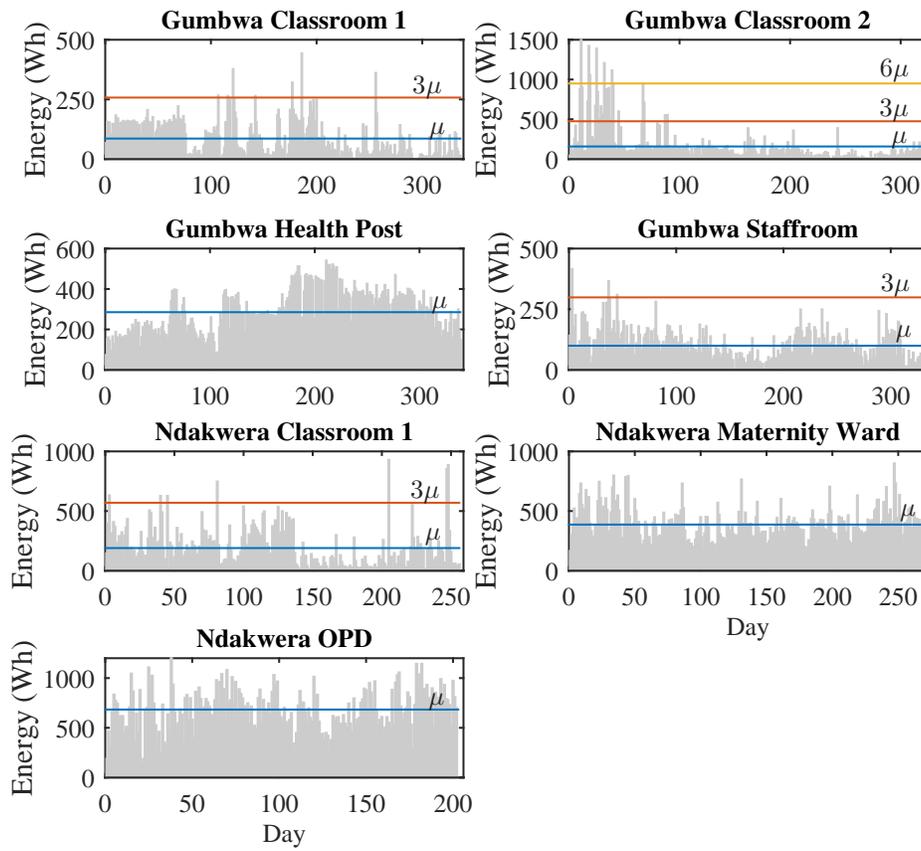


Figure 2: Time-series of daily load with horizontal lines indicating the location of the average and select multiples of the average.

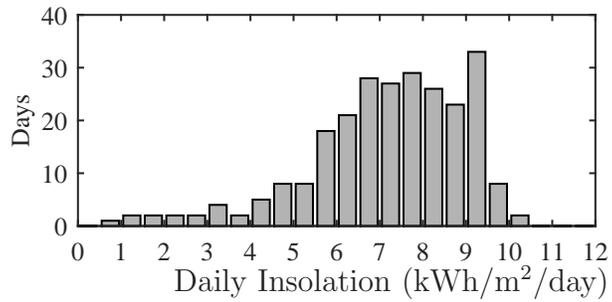


Figure 3: Histogram of daily insolation.

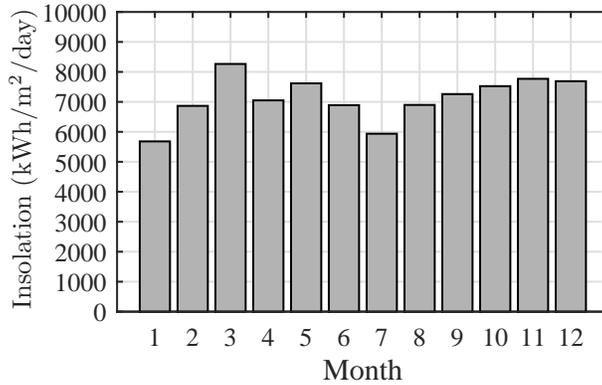


Figure 4: Daily average insolation by calendar month.

providing meaningful insight into how the two most critical and costly system components are affected by load estimation error.

The overall approach in this research is to produce re-designed systems for each of the seven Malawian sites based on intentional over- and under-estimations of average daily load, then estimate the corresponding cost and compute the reliability through deterministic simulation based on the measured data.

3.1. Load estimation

For each site, the average daily load was computed from the measured data as was shown in Table 2. Average daily load “estimations” are made that range between $\pm 90\%$ of the actual measured load at nominally 10% intervals. A Base Case with zero load estimation error is considered as an important reference case.

3.2. System design

There are several methods to sizing the PV array and battery of an off-grid system based on an estimate of average daily load, as described in the Introduction. This article focus on the intuitive method. To acknowledge the variability of specific intuitive approaches, and to guard against biases that might be inherent within a particular approach, three hypothetical designs were produced for each site and load estimate. An *ad-hoc* approach was applied to produce one set of designs, whereas IEEE standards [12, 13] were applied to produce the other two sets of designs.

Several *ad-hoc* or ‘rule-of-thumb’ design approaches are common and have been documented in various case studies. For example, one approach estimated a simple daily load curve with no weekend or seasonal variation and had no accounting for panel and battery aging effects [46]. Further case studies incorporate different levels of design sophistication and demonstrate the variability of the approach [47, 48]. The *ad-hoc* approach applied in this research is the one followed by the designers of the Malawi systems, which is further detailed in the Appendix.

IEEE standards 1562-2007 (PV sizing) and IEEE 1013-2007 (Battery sizing) are among those applicable to the design of off-grid PV systems [12, 13, 49]. These two IEEE standards are used in this research. They are jointly applied to produce each complete design. The IEEE standards require information which is generally not known *a priori* and not captured in the data set, such as the efficiency of specific components and the effects of aging on energy production. Estimates of these parameters are needed, and in some cases the range of reasonable values is broad—for example 5–54% for total losses. To reflect the range of reasonable values, one set of designs was produced as a result of optimistic assumptions and another as a result of pessimistic assumptions. These designs provide upper and lower bounds on the variety of PV and battery sizes that could be reasonably produced by following the IEEE standards. Specific details of the standards-based design approach are provided in the Appendix.

Each of the three intuitive approaches requires an estimate of the daily insolation during the month with the lowest average. To retain a focus on energy load estimation error, the actual measured lowest-month average daily insolation value of 5.6 kWh/m²/day is used in all designs.

3.3. Cost

It has been reported that the cost of PV systems in LDCs can vary widely and are generally more expensive than the world average [50]. In the following, a cost model based on Malawian price data is developed.

The cost to realize a given design with PV array rating R_{PV} and battery rating R_B depends on the quantity and rating of the individual panels and batteries used. The per-panel and per-battery price model is derived from 27 price quotes of PV panels and 15 price quotes of batteries of various ratings, as shown in Fig. 5 and Fig. 6. No distinction was made between manufacturer, warranty periods or other technical specifications beside rated capacity, which explains some of the variability in quoted price for a given

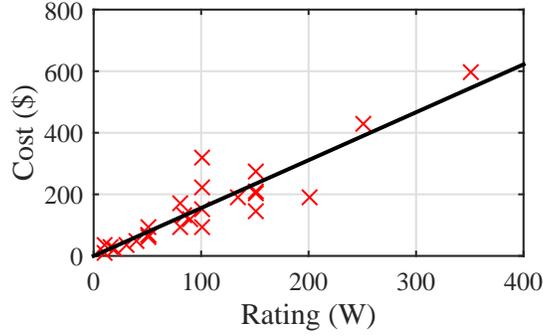


Figure 5: Linear cost curve fitted to PV panel price data.

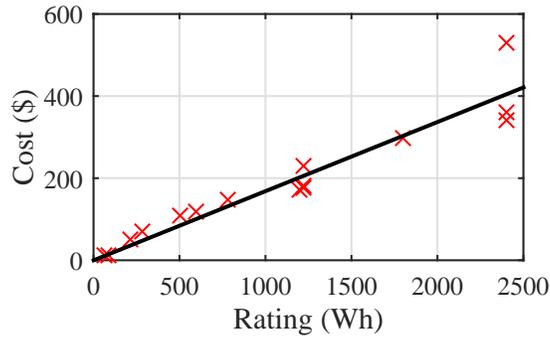


Figure 6: Linear cost curve fitted to battery price data.

rating. Balance of system components and installation were not included in the quotations. Undoubtedly, costs will vary between and within countries depending on exchange rates, duties, competition and other factors.

Linear cost models were fit to the data:

$$c(R) = \beta R \tag{1}$$

where c is the panel or battery cost, β is the cost coefficients and R is the equipment rating in watts for PV panels and watthours for batteries. The coefficients were determined from a least squares fit and are provided in Table 3, along with the unadjusted coefficient of determination r^2 . Inspection of Fig. 5 and Fig. 6 show that the models fit the data reasonably well. The cost model is specific to the Malawian case, but is in general agreement with cost models reported elsewhere [50].

Table 3: Equipment Cost Coefficients

Equipment	β	r^2
Battery	0.168	0.901
PV Panel	1.555	0.853

3.4. Reliability

The reliability of each design is computed through deterministic simulation as described hereafter. There are several measures of reliability. The analyses in this paper considers metrics based on the Loss of Power Supply Probability and to a lesser extent, the Expected Energy Not Served. These metrics are commonly applied to off-grid PV systems [20, 21, 51, 52, 53].

While these two measures often have probabilistic interpretations, because they are computed from a deterministic simulation, a statistical interpretation is more appropriate in the context of this research. The more appropriate designations of Loss of Power Supply Percentage (LPSP) and Energy Not Served (ENS) are adopted hereafter to highlight this distinction.

The temporal resolution of the simulation is one hour, so that the LPSP is interpreted as the percent of total hours when some or all of the load was not served. For convenience, the reliability is presented as $100 - LPSP$ so that a larger number indicates greater reliability. The ENS is the total value of the energy not served during the simulation time period [53]. Because the number of hours and days simulated varies among the sites as was shown in Table 1, the ENS per hour or ENS per hour with an outage are reported for a more appropriate comparison.

Conceptually, the simulation exposes the redesigned systems to same insolation and load conditions recorded at the Malawian sites to see how they would have performed. For each site, the load and insolation data are synchronized in time so that any correlation, autocorrelation or dependency structure between or within the load and insolation are represented. The use of synchronized measured data makes the simulation methodology unique, as these potentially complex relationships are not captured in probabilistic or synthetic time-series approaches to reliability evaluation.

Each design is simulated independently, using the specified PV array and battery ratings. The parameters and assumptions used in each simulation are identical and are summarized in Table 4. The assumptions are such that they are neither at the optimistic or pessimistic extremes. It is important

Table 4: Simulation Parameters

Parameter	Value
End of Life Rating, A_{PV} , A_B	80%
Battery Charging Efficiency, $\eta_{B,in}$	92%
Battery Discharging Efficiency, $\eta_{B,out}$	92%
Battery Depth of Discharge, DOD	50%
Combined Efficiency, η_{SYS}	85%
Initial Battery State of Charge	100%

to note that the End of Life parameters are set to 80%, implying that the components have a functional capacity that is 80% of the designed rating due to age-related effects. The simulation therefore can be conceptualized as occurring near the end of the life of the equipment.

Let $h[t]$ and $l[t]$ be the measured insolation and load for a given site at time t , respectively. The simulation progresses sequentially, one hour at a time, so that t is interpreted as the hour index. As described previously, days missing more than four hours of either load or insolation data are removed from the data set, so that on some occasions non-consecutive days are simulated consecutively.

The simulation begins with the first hour of the first day of the processed and synchronized data set under the assumption that the battery is initially fully charged. The insolation is first converted to energy production:

$$p[t] = \frac{h[t]}{G_{STC}} \times R_{PV} \quad (2)$$

where R_{PV} is the PV panel rating, G_{STC} is the irradiance at which the PV panels are rated—typically 1000 W/m^2 —and $p[t]$ is the idealized energy production from the PV panels during hour t . Next, the effects of component aging and losses are included:

$$p'[t] = p[t] \times (A_{PV}) \times \eta_{SYS} \quad (3)$$

where A_{PV} is the PV end of life rating adjustment, and η_{SYS} is the combined efficiency of the components other than the battery and inverter, including losses caused by dust, shading, PV panel mismatch and cable resistance. The variable $p'[t]$ is interpreted as the available energy from the PV array

during hour t that can be applied for useful purposes—either supplying load or charging the battery.

The net energy $n[t]$ is found by subtracting the load:

$$n[t] = p'[t] - l[t]. \quad (4)$$

If $n[t]$ is positive, then the battery is charged during hour t ; whereas a negative value indicates the battery is discharged. The change in battery energy level after accounting for losses associated with charging or discharging is:

$$\Delta b[t] = \begin{cases} n[t] \times \eta_{B,\text{in}} & n[t] > 0 \\ n[t]/\eta_{B,\text{out}} & n[t] \leq 0 \end{cases} \quad (5)$$

where $\eta_{B,\text{in}}$ and $\eta_{B,\text{out}}$ are the battery charging and discharging efficiencies, respectively. The battery energy level is increased or decreased according to:

$$b'[t] = b[t - 1] + \Delta b[t] \quad (6)$$

where $b[t - 1]$ is the energy in the battery at the end of hour $t - 1$,

The battery capacity limits are then applied:

$$b[t] = \begin{cases} R_B \times A_B & b'[t] > R_B \times A_B \\ R_B \times DOD & b'[t] < R_B \times DOD \\ b'[t] & \text{otherwise} \end{cases} \quad (7)$$

where R_B is the battery rating, A_B is the end of life adjustment for the battery, and DOD is the battery depth-of-discharge limit. If $b'[t]$ is less than $b[t]$, then some or all of the load cannot be served. Practically, the inverter or battery charge controller would disconnect the load due to low battery voltage, in order to protect the battery from a damaging deep discharge. The energy not served during hour t is

$$e[t] = \max(b[t] - b'[t], 0). \quad (8)$$

The index h is then incremented and the next hour is simulated. The process continues until all hours have been considered.

The LPSP is computed by counting the number of hours where $e[t] \neq 0$ and dividing by the total number of hours simulated, whereas the *ENS* is the sum of $e[t]$ over all hours simulated.

3.5. Considerations

The results must be interpreted within the bounds of the methodology and is subject to three main caveats. First, the method explicitly considers PV array and battery design only, for reasons stated previously. Other components are also affected by load estimation error. However, the sizing of charge controllers, breakers, and other balance of system components also depends on DC bus voltage level, system topography and physical dimensions, which if considered would cause the analyses to lose tractability. Inverters are another critical component, but they are sized based on estimated power demand which does not uniquely map to load.

Since other components will be effected by load estimation error, the cost-related results presented in this article are attenuated as compared to what would be observed in practice. It is expected that over- or under-estimation of load would have a somewhat greater affect on cost than the results presented.

Second, it is stressed again that in the context of this article, “reliability” refers to the “energy reliability” of the system. Many factors influence the ability of an off-grid system to serve the load. Outages due to equipment malfunction, system mis-operation and other causes are not considered. The true reliability after accounting for these factors will be somewhat less than those reported here.

Lastly, the models and design approaches are limited to decentralized stand-alone systems and micro-grids where solar PV is the sole generation source, approximately 2 kW or less. As PV array capacity increases, design approaches often transition from the intuitive methods considered in this work to numerical or analytical methods. Larger systems are also more likely to use maximum power point trackers, higher DC voltages and perhaps other factors that are not considered in this research.

4. Design results

The component sizes corresponding to the three approaches for the zero estimation error (Base Case) are provided in Table 5 for each site. The Table shows that the component ratings are related to the average daily load—large ratings correspond to sites with large average daily load, such as Ndakwera OPD, whereas smaller component ratings correspond to sites with small average daily load, such as the Gumbwa Classroom 1. This relationship can be more precisely defined.

Table 5: Design Results–Base Case (0% Error)

Site	PV Size (W)			Battery Size (Wh)		
	<i>Ad-hoc</i>	Opt.	Pes.	<i>Ad-hoc</i>	Opt.	Pes.
Gumbwa Classroom 1	51	33	78	576	1266	2518
Gumbwa Classroom 2	89	58	137	1014	2231	4437
Gumbwa Staff Room	58	38	89	660	1452	2887
Gumbwa Health Post	171	110	262	1940	4268	8487
Ndakwera Classroom 1	100	65	154	1137	2502	4975
Ndakwera Mat. Ward	215	139	330	2442	5373	10685
Ndakwera OPD	287	185	440	3261	7175	14268

4.1. *Component rating is proportional to estimated load*

When using intuitive approaches, the mathematical relationship between battery rating and daily load is linear: $R_B = Z_B \times L$, where L is the average daily load the system is designed for and Z_B is the variable of proportionality. In a single value, Z_B encapsulates all of the factors and assumptions used in a given design approach—loss percentage, and days of autonomy, among others. After applying each design approach using the assumptions detailed in the Appendix, the resulting Z_B were calculated and are provided in Table 6. The PV rating depends on the battery rating, the insolation and PV-related losses and de-rating. It too can be related to average daily load by a variable of proportionality Z_{PV} , as provided in Table 6.

4.2. *Component ratings scale proportionally with estimate error from base case*

A consequence of the linear relationships between component rating and estimated load is that the PV array and battery ratings scale from the Base Case in proportion with the load estimation error. For example, the design of a system with a +25% load estimation error requires a PV array and battery that are 25% larger than the Base Case. In other words, there is a one-to-one relationship between the error in average daily load estimation and the increase or decrease in the rating of the PV array and battery from the Base Case.

4.3. *Different design approaches yield different component ratings and ratios*

Inspection of Table 5 and Table 6 highlight two important findings. First, for a given site, the three approaches resulted in a wide range of component

Table 6: Proportionality Variable Values

Design Approach	Z_B	Z_{PV}	Z_B/Z_{PV}
<i>Ad-Hoc</i>	6.2	0.55	11.3
Optimistic	13.8	0.35	39.4
Pessimistic	27.5	0.85	32.4

ratings. This is true even when the same set of IEEE standards are applied but with different assumptions. This points to the independent question of which design approach is most appropriate for a given application, but that investigation is beyond the scope of the present analysis.

Second, not only do the ratings exhibit a wide range, but so does the ratio of battery rating to PV rating, as shown in the last column of Table 6. This ratio strongly influences the cost-effectiveness of a design. For example, a small ratio is generally better suited to sites where the load is co-incident with the solar production because less energy is needed to be supplied by the battery overnight.

5. Cost results

The cost associated with each design is shown in Fig. 7. The results show a linear trend between estimation error and cost. The explanation for this trend is as follows: any estimation error results in a proportional change in component rating from the Base Case. Due to the linear PV array and battery cost models, the proportional increase in ratings leads to a proportional increase in cost from the Base Case. In other words, for every one percent error in load estimation, there is a one percent change in cost.

5.1. Cost is influenced by design approach

The results highlight the economic consequences of design approach selection. The pessimistic designs are the most costly, costing twice as much as the optimistic designs and three times as much as the *ad-hoc* designs. These results are more meaningful when expressed in absolute monetary terms. At 90% over-estimation, the cost of the PV array and battery for the Ndakwera OPD for the *ad-hoc* design is \$1892. For the pessimistic design, the cost balloons to \$6331. From an economic perspective, the choice in design approach can critically influence the number of systems installed or even whether or not an off-grid electrification intervention is economically justified.

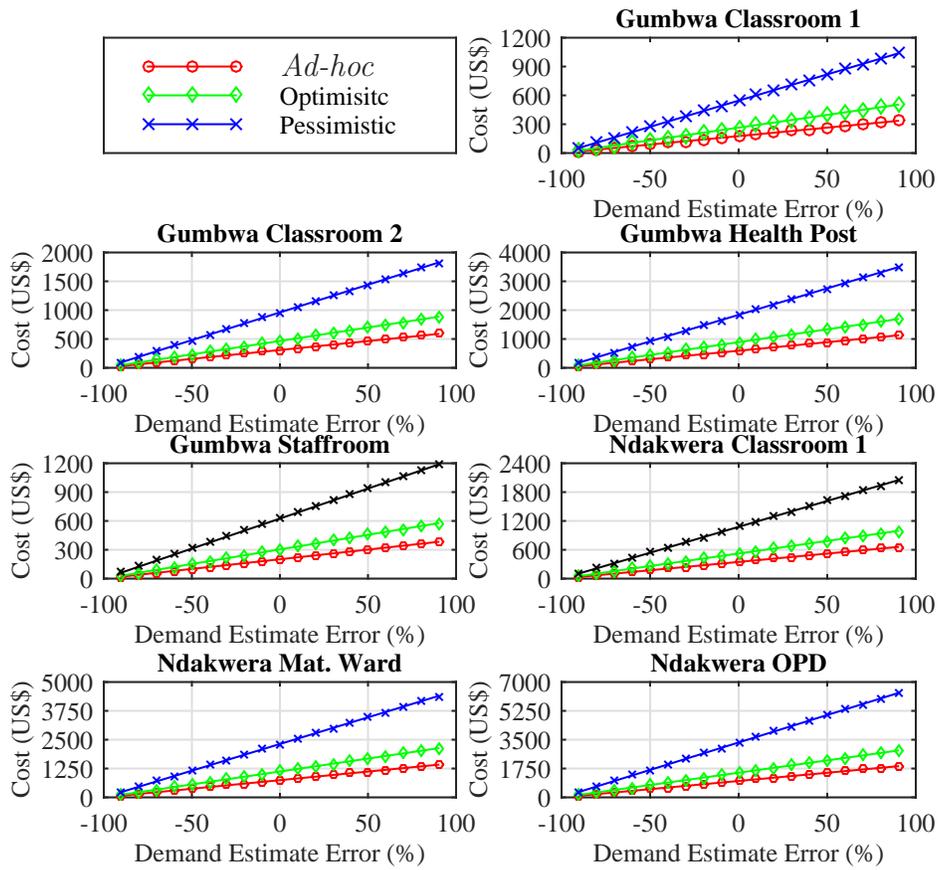


Figure 7: Cost of PV array and battery for each site using the optimistic design.

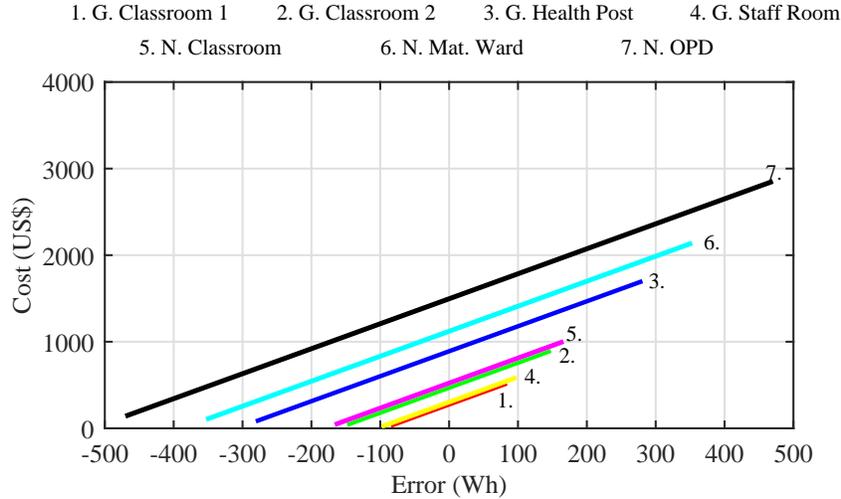


Figure 8: Cost of PV array and battery for each site and design.

5.2. Value of accurate load estimation

The results can be presented another way to determine the value of an accurate load estimate. Fig. 8 shows the costs for each site using the optimistic design approach as a function of estimate error expressed in watthours. The slope of each trace is the cost per watthour of load estimation error. The traces are nearly parallel, showing the cost of estimation error to be relatively constant, and on average \$2.89 per watthour of error. The *ad-hoc* and pessimistic approaches are similar but with slopes of \$1.92/Wh and \$6.02/Wh, respectively.

The results show that the *value* of an improved estimate of daily load is as high as \$6.02/Wh. This is a useful rule-of-thumb for organizations allocating resources to estimate the load for a site. For example, if conducting a detailed survey improves the estimation of daily load by 100 Wh, and costs less than \$602 to conduct, it may be worthwhile to do so. These results only meaningfully apply to load over-estimation. Under-estimation reduces component costs, but has consequences related to the reliability of the system.

6. Reliability results

The reliability of each design was computed via the described simulation method. The results are provided in Fig. 9. Before exploring the reliability

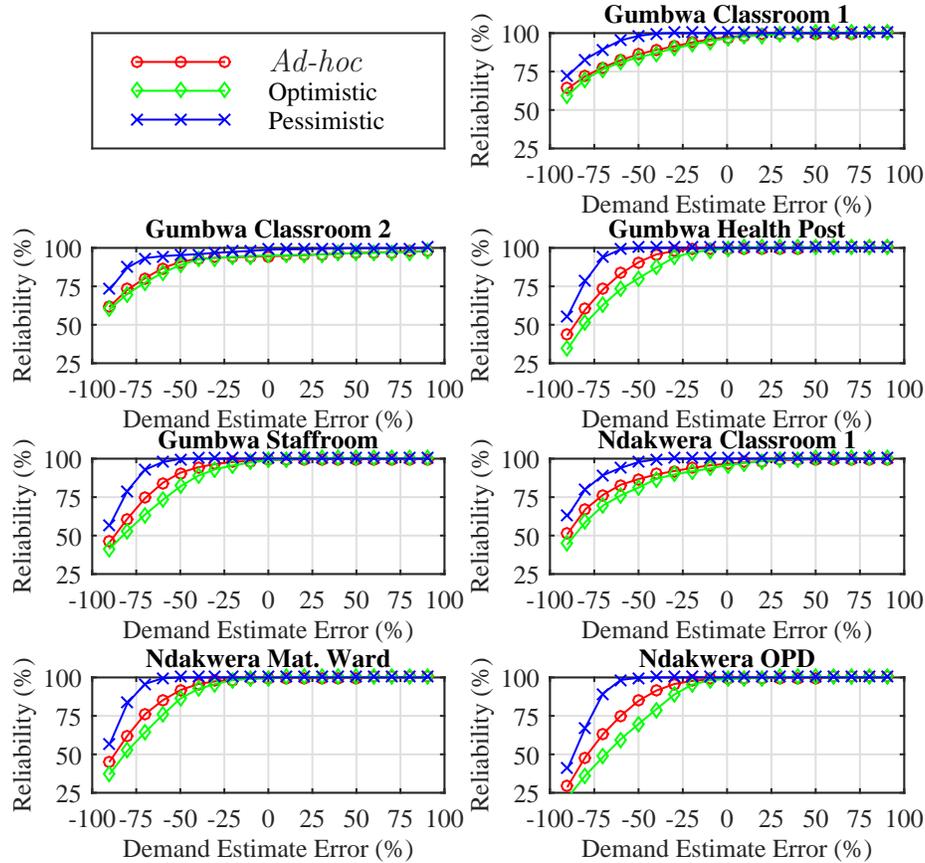


Figure 9: Reliability results expressed as a function of load estimation error percentage.

curves in total, it is insightful to first consider the Base Case (perfect load estimation).

6.1. Perfect load estimation does not always yield perfect reliability

Inspection of Fig. 9 shows a somewhat surprising result: even with zero estimate error, none of the designs obtained 100% simulated reliability for Gumbwa Classroom 2. In other words, even with perfect estimation of average daily load and making pessimistic assumptions, the system designed using the IEEE standards resulted in the occasional loss of power supply. The reason for the low reliability for Gumbwa Classroom 2 is the relatively large number of days with extremely high load that occurred consecutively, as was shown in Fig. 1 and Fig. 2. An important conclusion is that follow-

ing accepted standards for sizing system components, even with conservative assumptions, does not guarantee perfect reliability. Loads exhibiting or expecting to exhibit “fat” tails in their distribution and days with large loads that occur consecutively or nearly consecutively require special consideration in the design phase.

The performance of the designs with zero load estimate error are explored further in Table 7. The reliability, ENS per hour and ENS per outage hour for the base case of zero-error load estimation are provided. In general the *ad-hoc* designs were the least reliable. The design approach appears to be poorly suited for the Malawian sites. In practice, the implemented systems achieved high reliability serendipitously, due to the extreme over-estimation of load and under-estimation of the insolation, as was shown in Table 2.

The ENS-derived metrics show a relatively low impact of the outages across all sites and design approaches, with an overall average loss of energy of just 0.4 Wh per hour. When adjusted to only consider hours when outages occurred, the average becomes 21 Wh per hour, which is greater than the overall average hourly load of 10.4 Wh. This suggests that outages tend to occur when the hourly load is high.

6.2. Sensitivity to load estimation error

As shown in Fig. 9, the relationship between average daily load estimation error and reliability is non-linear and sensitive to design approach and load characteristics. The general trend is similar for all sites and design approaches. At 90% under-estimation the reliability is low, generally between 20 and 60%. As the estimate improves, there is a rapid increase in reliability, albeit with some variation between sites and design approach. The reliability on average increases by 1.4 percentage points for every one percentage point improvement in load estimate for estimates between -90 and -70%. The knee of the curves occur at error levels near -50%. As the estimation error approaches zero, the reliability approaches 100% and becomes less sensitive to improvements in estimation. Here the average reliability increases by just 0.04 percentage points for every one percentage point improvement in estimation.

6.3. Reliability is influenced by design approach

The different design approaches yield different levels of reliability for a given estimation error. For all sites, the pessimistic designs had the greatest reliability for a given estimate error. This is due to the large Z_B and Z_{PV}

Table 7: Base Case (0% Error) Reliability Results

	Site	Approach	100-LPSP (%)	ENS/hr (Wh/hr)	ENS/outage hr (Wh/hr)
Gumbwa	Classroom 1	<i>Ad-hoc</i>	97.32	0.28	10.32
		Optimistic	96.64	0.40	12.03
		Pessimistic	100.00	0.00	0.00
	Classroom 2	<i>Ad-hoc</i>	94.55	1.80	32.97
		Optimistic	94.97	1.87	37.28
		Pessimistic	98.85	0.46	40.21
	Health Post	<i>Ad-hoc</i>	99.80	0.06	27.77
		Optimistic	99.80	0.06	31.70
		Pessimistic	100.00	0.00	0.00
	Staff Room	<i>Ad-hoc</i>	99.04	0.10	10.66
		Optimistic	98.97	0.07	6.94
		Pessimistic	100.00	0.00	0.00
Ndakwera	Classroom	<i>Ad-hoc</i>	96.68	0.71	21.41
		Optimistic	95.75	0.78	18.45
		Pessimistic	100.00	0.00	0.00
	Mat. Ward	<i>Ad-hoc</i>	99.60	0.17	42.88
		Optimistic	99.77	0.10	43.12
		Pessimistic	100.00	0.00	0.00
	OPD	<i>Ad-hoc</i>	99.29	0.42	58.61
		Optimistic	99.00	0.44	43.92
		Pessimistic	100.00	0.00	0.00

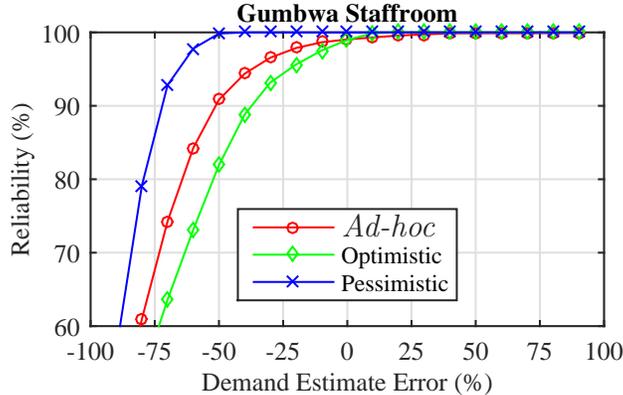


Figure 10: Zoomed-in reliability results of the Gumbwa Staffroom.

values associated with this approach. The *ad-hoc* designs achieve greater reliability than the optimistic designs for estimates to left of the knee of the curve. This can be explained by comparing the Z_B and Z_{PV} values associated with the two methods in Table 6. The values are such that the optimistic designs have larger batteries but smaller PV arrays than the *ad-hoc* designs for a given estimate. When the load is greatly under-estimated, having a large battery does little to improve reliability if the PV array is not commensurately sized. However, as the estimate error passes the knee of the curve, the reliability of the optimistic designs tend to surpass that of the *ad-hoc* designs. The reliability plot for Gumbwa Staffroom is enlarged in Fig. 10 to better illustrate this general point. This suggests that for the Malawian sites, obtaining high reliability requires relatively large batteries. Large batteries buffer against the rare occurrences of extreme load, especially when these occur in the evening.

6.4. Perfect assumptions do not always yield perfect reliability

Next, the simulation parameters are changed so that they more closely match the design assumptions in each approach, rather than being the same for all approaches. The Base Case is considered. Conceptually, this is akin to the designer knowing beforehand the average load, the system efficiencies, depth of discharge and insolation of the lowest month. The simulation parameters for each approach are provided in Table 8, as derived from the design assumption provided in the Appendix. Note that the IEEE Standards bundle aging effects and battery efficiencies (losses) into the total system

Table 8: Simulation Parameters

Parameter	<i>Ad-Hoc</i>	Optimistic	Pessimistic
Aging Derating, A_{PV} , A_B	100%	100%	100%
Battery Charging Efficiency, $\eta_{B,in}$	100%	100%	100%
Battery Discharging Efficiency, $\eta_{B,out}$	100%	100%	100%
Battery Depth of Discharge, DOD	50%	50%	50%
Total System Loss, $(100 - \eta_{SYS})$	35%	5%	54%
Initial Battery State of Charge	100%	100%	100%

Table 9: Reliability Using As-Designed Simulation Parameters Base Case (0% Error)

	Gumbwa				Ndakwera		
	Class-room 1 (%)	Class-room 2 (%)	Health Post (%)	Staff room (%)	Class-Room (%)	Mat. Ward (%)	OPD (%)
<i>Ad hoc</i>	99.56	95.86	100.00	99.90	99.43	100.00	99.94
Optimistic	100.00	96.81	100.00	100.00	100.00	100.00	99.94
Pessimistic	100.00	97.87	100.00	100.00	100.00	100.00	100.00

loss. To avoid double counting, these parameters are set to 100% in the simulation. Note that to be consistent with the design assumptions in the Appendix, Table 8 converts the Combined Efficiency to Total System Loss. The results are shown in Table 9.

The use of as-designed assumptions as simulation parameters somewhat improved the reliability of the *ad hoc* and optimistic designs, but reduced it for the pessimistic designs, reflecting the selection of moderate simulation parameters used in the previous sections. It is interesting that when the simulation parameters are set to match the as-designed assumptions, the simulated reliability is less than 100 percent for many of the sites. The failure of the designs to attain 100% simulated reliability when even the design assumptions are correct, points to a fundamental drawback of the intuitive approach—load distribution and temporal characteristics are ignored.

6.5. Illustrative critique of the intuitive design approach

The advantages of the intuitive design approaches, like those considered in this article, is that they require limited input information—primarily average

load and worst month insolation—and that the design approach is straightforward, which allows them to be readily implemented. There are several disadvantages. Selection of parameters such as the effects of aging and system losses can vary widely and are difficult to know *a priori*. The differences between the cost and reliability of the optimistic and pessimistic designs shown in this article highlight this point. Although intuitive approaches require limited input information, estimating average load in particular is often challenging as shown in this article. Further, the intuitive approaches do not account for the distribution of daily load, nor the temporal relationship between occurrences of high load. This last point is further explored in the following.

Consider now Gumbwa Classroom 2 and Ndakwera OPD. The average load for the OPD is 3.2 times as large as Gumbwa Classroom 2. Using an intuitive approach results in the OPD system PV panels and batteries being 3.2 times larger than Gumbwa Classroom 2. Consequentially, the cost is also 3.2 times greater. However, because the distribution of the OPD load has fewer extreme occurrences of high load as seen in Fig. 1 and that these occurrences are spread in time, as seen in Fig. 2, the battery capacity can in fact be sized similarly to Gumbwa Classroom 2.

Fig. 11 shows the results of an hourly simulation for the Gumbwa Classroom 2 Base Case using the pessimistic design with simulation parameters in Table 4. The battery capacity is 4437 Wh, which is derated to 3550 Wh due to the effects of aging. A five-day period is shown. The top plot in Fig. 11 is the hourly load. The load as expressed as a multiple of the overall average μ is provided at the top of the plot for each day. During the shown period, there are three successive days of load exceeding 5μ . The load was more than 22 times the average load occurred during this five-day period.

The daily PV production, shown in the second plot, is insufficient to replenish the battery after the high load in day two. Consequently, the battery discharges, as shown in the third plot. The bottom plot shows the energy stored in the battery. The dashed line is the depth of discharge cutoff level (50 percent of the rating). The system is able to withstand the extremely high load of day two, which was 5.46μ . However, the occurrence of an even greater load the following day, at 9.29μ , results in a loss of power supply. Despite the pessimistic design approach being based on 10.5 days of autonomy, two consecutive days of high load resulted in a loss of power supply.

This can be contrasted with the Ndakwera OPD. The pessimistic design

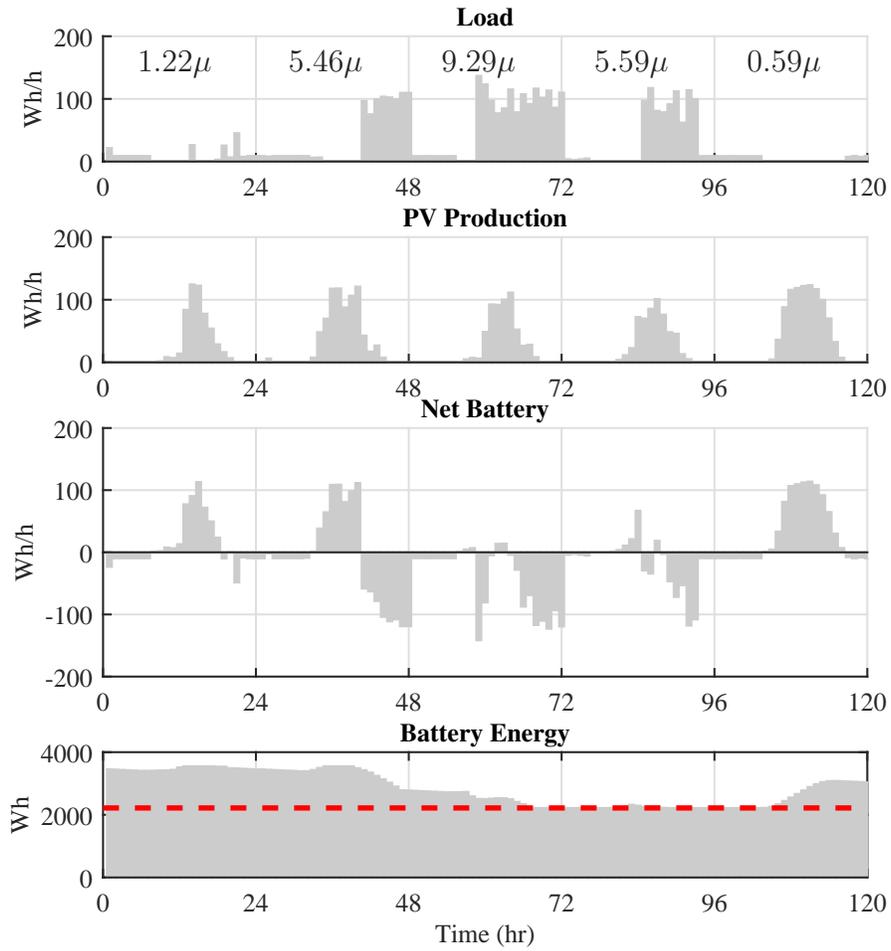


Figure 11: Time-series of Gumbwa Classroom 2 showing hourly load, PV production, net power to the battery and battery energy level.

is considered, however instead of using the 14268 Wh battery called for in the design, let the smaller 4437 Wh battery called for in Gumbwa Classroom 2 design be used. This equates to just 3.3 days of autonomy. The simulation was re-run and the five days with the greatest average load are shown in Fig. 12. During this period, the load is above average during the entire period, but not as severe as occurred with Gumbwa Classroom 2. Despite the 3.2 times greater average load in the OPD than Gumbwa Classroom 2, the distribution and temporal characteristics of the load mean that the battery rating at the Ndakwera OPD can be the same size as the Gumbwa Classroom 2 and achieve similar overall reliability. This encapsulates a fundamental disadvantage of the intuitive method: although it can be readily implemented, it relies on incomplete load information, obscuring the cost versus reliability tradeoff in the system design.

7. Reliability versus cost

This section explores the important but not well understood reliability-versus-cost tradeoff in small-scale off-grid PV systems. The reliability-versus-cost curves are shown in Fig. 13. The curves are developed using the same methodology to compute cost and reliability as previously described (using the simulation parameters in Table 4), but the range of considered estimation errors is extended to +600% at 1% intervals to achieve high resolution curves. The region of the curve of greatest interest is near 100% reliability as this is targeted by most system designers. The ordinate is logarithmic to better distinguish the sensitivity of reliability to cost.

7.1. Cost effectiveness is influenced by the Z_B/Z_{PV} ratio

The *ad-hoc* approach is generally the most cost-effective for obtaining reliability up to 99.9%. For example, for the Gumbwa Classroom 1, the *ad-hoc* design achieves 99.9% reliability at a cost of approximately \$280, whereas the next lowest-cost design costs \$350. The cost-effectiveness of a design is closely related to the ratio of Z_B/Z_{PV} . The insolation and load characteristics of each site dictate which ratio is more cost effective for a given level of reliability. For reliability up to 99.9% in the Malawian sites, the lower ratio of the *ad-hoc* designs (11.3) is most cost-effective. It follows that the pessimistic designs are the next most cost-effective as their ratio is the next lowest (32.4). However, past 99.9%, the most cost-effective Z_B/Z_{PV}

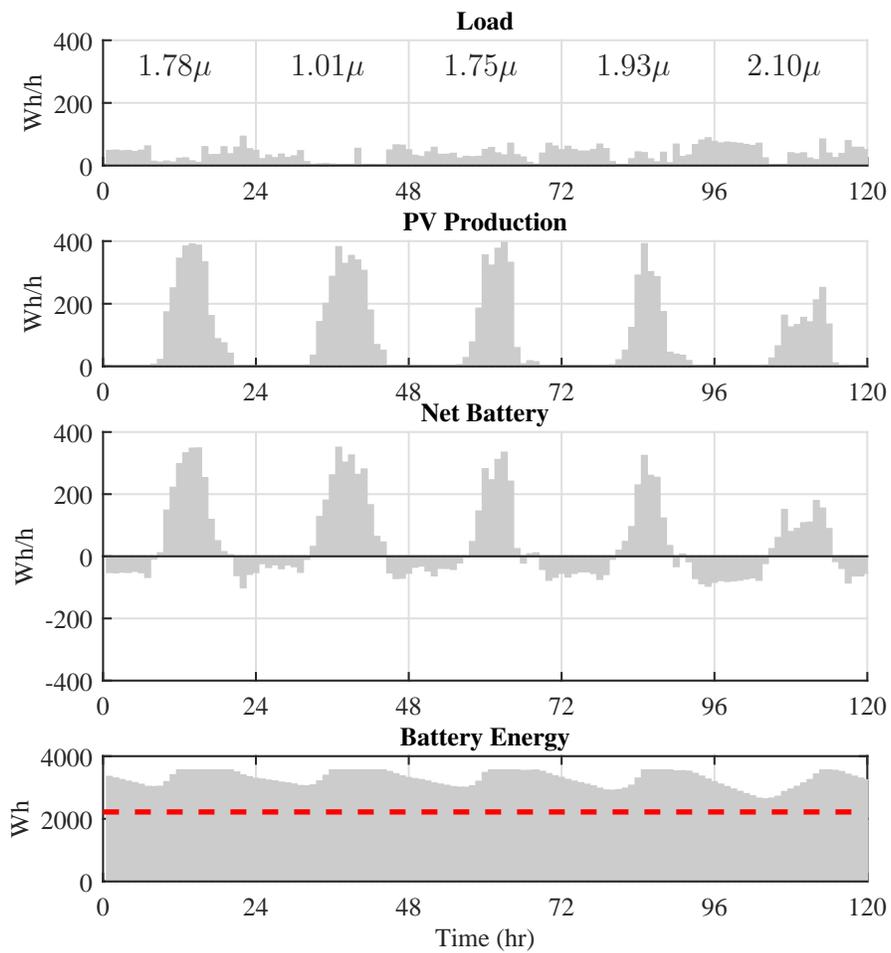


Figure 12: Time-series of Ndakwera OPD showing hourly load, PV production, net power to the battery and battery energy level.

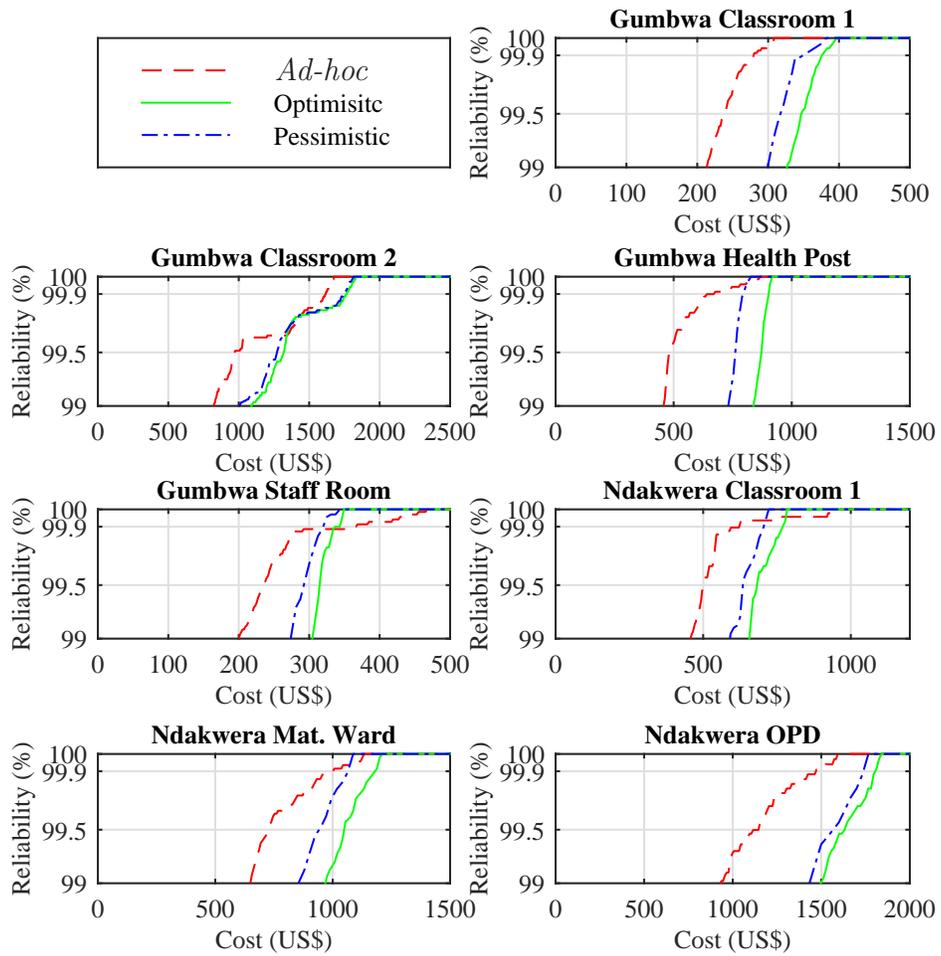


Figure 13: Reliability-versus-cost curves.

ratio tends to increase, and in some cases the pessimistic designs are the most cost effective.

A note of caution is warranted here. Although cost-effective, recall that the *ad-hoc* approach when applied with zero estimation error produced the least reliable designs. For some sites the reliability would likely be unacceptable, as was seen in Table 7. The *ad-hoc* approach could be modified to improve the reliability at zero estimation error without sacrificing its cost-effectiveness by changing its assumptions so that Z_B and Z_{PV} are increased while maintaining the same Z_B/Z_{PV} ratio.

7.2. High reliability comes at high cost

In general, the incremental cost of improving reliability greatly increases as the reliability approaches 100%. On average, 46% of the total cost of a 100% reliable system is spent on improving reliability from 99% to 100%.

Over the course of a year, a system with 99.7% reliability will experience 26 hours of outage. For the Malawian sites this will cause an average of 546 Wh of energy to not be served. For a system with 99.9% reliability, the figures are nine outage hours and 189 Wh of energy not served per year—an additional 357 Wh of energy served. The average cost to improve reliability from 99.7% to 99.9% is \$197. As a simple calculation, assuming a 10-year lifespan for PV array and battery, the cost to serve the additional 357 Wh each year is approximately \$55.2 per kWh. This extraordinarily large value highlights the need for broader research and discussion on reliability expectations of small off-grid systems in LDCs, knowing that perhaps twice as many systems can be installed or the electricity supplied at a less expensive tariff if target reliability levels were relaxed.

8. Conclusions and future outlook

This research investigated the impact of average daily load estimation error on system design, cost and reliability using hourly load and insolation data from seven off-grid PV systems in Malawi. The basic relationship between cost and reliability was also analyzed.

The results showed that estimation error has a proportional affect on PV array and battery ratings, which in turn has a nearly proportional affect on the cost. The cost of over-estimation depends upon what design approach is applied, but generally ranges from \$1.92 to \$6.02 per watthour of error. Reliability, on the other hand, is non-linearly influenced by load estimation

error and is generally less sensitive to estimation error as the error approaches zero (perfect estimation).

With some design approaches, even perfect load estimation did not yield perfect reliability. This points to the need to better consider the distribution of daily load in the design process, not just the mean. A comparison of reliability and cost showed that the incremental cost of improving reliability sharply increases as the reliability approaches 100%. For the sites considered in this article, it is questionable whether or not the premium paid for additional PV and battery capacity to achieve perfect reliability is justifiable.

The analyses have real and practical consequences. For example, recall that the Ndakwera maternity ward’s average daily load was estimated to be 1182 Wh compared to an actual average daily load of 388.4 Wh. Had the IEEE standards been applied using pessimistic assumptions, the resulting PV array and battery costs are estimated to have been \$7032. This design has a 100% simulated reliability. Contrast this if the average load had been perfectly estimated. The PV array and battery costs would be \$2313 and the simulated remain at 100%. The savings would be enough to electrify two additional maternity wards of the same size.

This work raises broader questions regarding the design off-grid PV systems. The economic merit of more accurate average daily load estimates has been shown, but a methodology for improving the estimates is lacking. The weakness in using average daily load as a primary input to intuitive design approaches was called into question, as the distribution of the daily load—particularly the outliers—and when these outliers occur was shown to be important. Design approaches incorporating additional load characteristics may be useful. Finally, given the high cost of high reliability, greater discussion among the research and practitioner community regarding target reliability standards for off-grid systems is needed.

- [1] Project Everyone. The global goals for sustainable development. [Online]. Available: <http://www.globalgoals.org/>
- [2] International Energy Agency, “World energy outlook 2015,” 2015.
- [3] —, “Africa energy outlook 2014,” 2014.
- [4] —, “World energy outlook 2011,” 2011.

- [5] T. Khatib, A. Mohamed, and K. Sopian, "A review of photovoltaic systems size optimization techniques," *Renewable and Sustainable Energy Reviews*, vol. 22, pp. 454 – 465, 2013.
- [6] R. Posadillo and R. L. Luque, "Approaches for developing a sizing method for stand-alone PV systems with variable demand," *Renewable Energy*, vol. 33, no. 5, pp. 1037 – 1048, 2008.
- [7] A. Celik, "Effect of different load profiles on the loss-of-load probability of stand-alone photovoltaic systems," *Renewable Energy*, vol. 32, no. 12, pp. 2096 – 2115, 2007.
- [8] S. Treado, "The effect of electric load profiles on the performance of off-grid residential hybrid renewable energy systems," *Energies*, vol. 8, no. 10, p. 11120, 2015.
- [9] E. Osaghae, E. Ibhado, and C. A. Ikonwa, "Impact of different load profile patterns on intermittent renewable energy systems for an off grid power system," *International Journal of Engineering Research in Africa*, vol. 17, pp. 82–93, 8 2015.
- [10] R. Chapman, "Sizing handbook for stand-alone photovoltaic/storage systems," Sandia National Labs., Albuquerque, NM (USA), Tech. Rep., 1987.
- [11] C. Salame, G. Khoury, M. Aillerie, M. M. E. Ali, and S. K. Salih, "A Visual Basic-based tool for design of stand-alone solar power systems," *Energy Procedia*, vol. 36, pp. 1255 – 1264, 2013.
- [12] "IEEE recommended practice for sizing lead-acid batteries for stand-alone photovoltaic (PV) systems," *IEEE Std. 1013-2007*, 2007.
- [13] "IEEE guide for array and battery sizing in stand-alone photovoltaic (PV) systems," *IEEE Std. 1562-2007*, 2007.
- [14] C. V. T. Cabral, D. O. Filho, A. S. A. C. Diniz, J. H. Martins, O. M. Toledo, and L. de Vilhena B. Machado Neto, "A stochastic method for stand-alone photovoltaic system sizing," *Solar Energy*, vol. 84, no. 9, pp. 1628 – 1636, 2010.

- [15] R. Posadillo and R. L. Luque, “A sizing method for stand-alone PV installations with variable demand,” *Renewable Energy*, vol. 33, no. 5, pp. 1049 – 1055, 2008.
- [16] A. Fragaki and T. Markvart, “Stand-alone PV system design: Results using a new sizing approach,” *Renewable Energy*, vol. 33, no. 1, pp. 162 – 167, 2008.
- [17] S. Labed and E. Lorenzo, “The impact of solar radiation variability and data discrepancies on the design of PV systems,” *Renewable Energy*, vol. 29, no. 7, pp. 1007 – 1022, 2004.
- [18] A. N. Celik, T. Muneer, and P. Clarke, “Optimal sizing and life cycle assessment of residential photovoltaic energy systems with battery storage,” *Progress in Photovoltaics: Research and Applications*, vol. 16, no. 1, pp. 69–85, 2008.
- [19] A. Q. Jakhrani, A.-K. Othman, A. R. H. Rigit, S. R. Samo, and S. A. Kamboh, “A novel analytical model for optimal sizing of standalone photovoltaic systems,” *Energy*, vol. 46, no. 1, pp. 675 – 682, 2012.
- [20] I. Abouzahr and R. Ramakumar, “Loss of power supply probability of stand-alone photovoltaic systems: A closed form approach,” *IEEE Transactions Energy Conversion*, vol. 6, no. 1, pp. 1–11, Mar. 1991.
- [21] L. Bucciarelli, “Estimating loss-of-power probabilities of stand-alone photovoltaic solar energy systems,” *Solar Energy*, vol. 32, no. 2, pp. 205–209, 1984.
- [22] L. Barra, S. Catalanotti, F. Fontana, and F. Lavorante, “An analytical method to determine the optimal size of a photovoltaic plant,” *Solar Energy*, vol. 33, pp. 509–514, 1984.
- [23] T. Khatib, A. Mohamed, K. Sopian, and M. Mahmoud, “A new approach for optimal sizing of standalone photovoltaic systems,” *International Journal of Photoenergy*, vol. 2012, pp. 1–7, 2012.
- [24] P. Díaz, R. Pena, J. Munoz, C. Arias, and D. Sandoval, “Field analysis of solar PV-based collective systems for rural electrification,” *Energy*, vol. 36, pp. 2509–2516, 2011.

- [25] S. Mandelli, M. Merlo, and E. Colombo, “Novel procedure to formulate load profiles for off-grid rural areas,” *Energy for Sustainable Development*, vol. 31, pp. 130–142, 2016.
- [26] N. Cross and C. Gaunt, “Application of rural residential hourly load curves in energy modelling,” in *Proceedings of IEEE Power Tech Conference*, vol. 3, Bologna, Italy, 2003, pp. 1–4.
- [27] M. Howells, T. Alfstad, N. Cross, L. Jeftha, and G. Goldstein, “Rural energy modeling,” Oct. 2002. [Online]. Available: http://www.erc.uct.ac.za/Research/publications-pre2004/02Howells-et-al_Rural_Energy_Modelling.pdf
- [28] R. Sen and S. C. Bhattacharyya, “Off-grid electricity generation with renewable energy technologies in India: An application of HOMER,” *Renewable Energy*, vol. 62, pp. 388 – 398, 2014.
- [29] S. Pachauri, “An analysis of cross-sectional variations in total household energy requirements in india using micro survey data,” *Energy Policy*, vol. 32, no. 15, pp. 1723 – 1735, 2004.
- [30] M. Zeyringer, S. Pachauri, E. Schmid, J. Schmidt, E. Worrell, and U. B. Morawetz, “Analyzing grid extension and stand-alone photovoltaic systems for the cost-effective electrification of kenya,” *Energy for Sustainable Development*, vol. 25, pp. 75 – 86, 2015.
- [31] P. Dauenhauer and H. Louie, “System usage trends for off-grid renewable energy users in developing communities,” in *Proceedings of the 4th Symposium on Small PV Applications*, Muchen, Germany, Jun. 2015.
- [32] O. Adeoti, B. Oyewole, and T. Adegboyega, “Solar photovoltaic-based home electrification system for rural development in Nigeria: domestic load assessment,” *Renewable Energy*, vol. 24, no. 1, pp. 155 – 161, 2001.
- [33] P. Boait, V. Advani, and R. Gammon, “Estimation of demand diversity and daily demand profile for off-grid electrification in developing countries,” *Energy for Sustainable Development*, vol. 29, pp. 135 – 141, 2015.
- [34] C. Blodgett, personal communication, May 2016.

- [35] A. Meier, “Saving electricity in a hurry,” 2005. [Online]. Available: <https://www.iea.org/publications/freepublications/publication/savingelec.pdf>
- [36] L. M. Stember, W. R. Huss and M. S. Bridgman, “A methodology for photovoltaic system reliability & economic analysis,” *IEEE Transactions on Reliability*, vol. 31, no. 3, pp. 296–303, Aug. 1982.
- [37] P. Longrigg, “System design and economic analysis of solar photovoltaic power supply,” in *Proceedings of IEEE Intelec International Telephone Energy Conference*, Dec. 1978, pp. 114–120.
- [38] S. Dhople and A. Domínguez-García, “Estimation of Photovoltaic System Reliability and Performance Metrics,” *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 554–563, Feb. 2012.
- [39] S. Shimura, R. Herrero, M. K. Zuo and J. A. B. Grimoni, “Production costs estimation in photovoltaic power plants using reliability,” *Solar Energy*, vol. 133, pp. 294–304, 2016.
- [40] H. A. M. Maghraby, M. H. Shwehdi and G. K. Al-Bassam, “Probabilistic Assessment of Photovoltaic (PV) Generation Systems,” *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 205–208, 2002.
- [41] P. Paliwal, N. P. Patidar and R. K. Nema, “A novel method for reliability assessment of autonomous PV-wind-storage system using probabilistic storage model,” *Electrical Power and Energy Systems*, vol. 55, pp. 292–703, 2014.
- [42] N. Nikmehr and S. N. Ravadaneg, “Reliability evaluation of multi-microgrids considering optimal operation of small scale energy zones under load-generation uncertainties,” *Electrical Power and Energy Systems*, vol. 78, pp. 80–87, 2016.
- [43] M. Mafuta, P. Dauenhauer and T. Chadza, “Remote Monitoring System Effectiveness: Small Scale Rollout of PV systems in Chikhwawa district, Malawi,” WASHTED, Tech. Rep., Jul. 2015. [Online]. Available: https://pure.strath.ac.uk/portal/files/54225726/Remote_Monitoring_System_Effectiveness_in_Malawi.pdf

- [44] K. Tembo, M. Mafuta and P. Dauenhauer, “Solar PV Strategic Energy Project in Southern Malawi: Development, Design and Impact,” WASHTED, Tech. Rep., Jul. 2015. [Online]. Available: https://pure.strath.ac.uk/portal/files/54225719/MREAP_Solar_PV_Strategic_Energy_Project.pdf
- [45] S. Mandelli, J. Barbieri, R. Mereu, and E. Colombo, “Off-grid systems for rural electrification in developing countries: Definitions, classification and a comprehensive literature review,” *Renewable and Sustainable Energy Reviews*, vol. 58, pp. 1621 – 1646, 2016.
- [46] M. Ishaq, U. H. Ibrahim and H. Abubakar, “Design of an off grid photovoltaic system: A case study of Government Technical College, Wudil, Kano State,” *International Journal of Scientific and Technology Research*, vol. 2, no. 12, 2013.
- [47] H. Guda and U. Aliyu, “Design of a stand-alone photovoltaic system for a residence in Bauchi,” *International Journal of Engineering and Technology*, vol. 5, no. 1, pp. 34–44, 2015.
- [48] S. Sakiliba, A. S. Hassan, J. Wu, E. S. Sanneh, and S. Ademi, “Assessment of stand-alone residential solar photovoltaic application in sub-Saharan Africa: A case study of The Gambia,” *Journal of Renewable Energy*, vol. 2015, pp. 1–10, 2015.
- [49] Government of New Zealand, “Stand-alone power systems - part 2: System design,” Standards New Zealand, Tech. Rep. AS/NZS4509.2:2010, 2010.
- [50] S. Baurzhan and G. P. Jenkins, “Off-grid solar PV: Is it an affordable or appropriate solution for rural electrification in Sub-Saharan African countries?” *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1405–1418, 2016.
- [51] J. Wang and F. Yang, “Optimal capacity allocation of standalone wind/solar/battery hybrid power system based on improved particle swarm optimisation algorithm,” *Renewable Power Generation*, vol. 7, no. 5, pp. 443–448, 2013.

- [52] J. Lucio, R. Valdés, and L. Rodríguez, “Loss-of-load probability model for stand alone photovoltaic systems in Europe,” *Solar Energy*, vol. 6, no. 1, pp. 2515–2535, Mar. 2012.
- [53] S. Rios, V. Vidal, and D. Kiguel, “Bus-based reliability indices and associated costs in the bulk power system,” *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 719–724, 1995.
- [54] M. Suri, N. Suriova, T. Cebecauer, A. Skoczek, J. Betak, and B. Schnierer, “Solar resource mapping in Malawi : solar modeling report,” World Bank Group, Tech. Rep., 2015.

A. Appendix

The principal calculations within the IEEE Standard method are noted in this appendix for completeness. The fully defined procedure is specified in the Standards documents [12, 13]. The parameters used in each design are found in Table 10.

The unadjusted battery capacity required for the system is given by:

$$C_{\text{unadj}} = D \times (L \times (1 + k_p)) \quad (9)$$

where D is the desired days of autonomy, L is the total daily average load in Amp-hours and k_p is the parasitic losses. The parasitic losses are due to the charge controller and are modeled as a percentage of L while inverter losses are not included because in this work the load is estimated and measured on the dc side of the inverter. Final required battery capacity, in Amphours, is determined by:

$$C_{\text{adj}} = \frac{C_{\text{unadj}} \times k_t \times k_m}{DOD} \quad (10)$$

where the temperature correction factor is k_t , design margin is given by k_m , and maximum daily depth of discharge is DOD .

The number of panels, P_N , and final PV array size, P_{WP} , are given by:

$$P_N = \frac{L \times A}{I_{\text{pp}} \times S_H \times (1 - K_L)} \quad (11)$$

$$P_{\text{WP}} = P_N \times P_{\text{STC}} \quad (12)$$

where the desired PV module is characterized by peak power current, I_{pp} , rated capacity at standard test conditions, P_{STC} , and array to load ratio, A . Peak sun hours, S_h , is determined by external reference, such as [54]. It should be noted that for reproduction of results, the simulation in the paper used $S_h = 5.6$ based on the lowest-month average daily insolation value from the data set. The individual elements of the system losses, K_L , are found in Table 10. Note that the losses are summed in accordance with the IEEE Standard, not multiplied. As such, they are expressed as a percentage of system load. Number of PV panels, P_N , in practice is rounded up to the nearest integer. The result is that for both the sizing of the battery and PV panel, the designed ratings scale linearly with L , the estimated load.

The design assumptions of each method are provided in Table 10. As shown, the *ad-hoc* parameters were not fully specified; in place of individual component loss parameters, a total system loss value was used in the design.

Table 10: Design Assumptions

Parameter	IEEE (Optimistic)	IEEE (Pessimistic)	<i>Ad-hoc</i>
Battery Sizing Parameters			
Days of Autonomy	6	10.5	3
Inverter losses	10%	10%	10%
Parasitic losses	1%	1%	-
Maximum Daily Depth of Discharge	50%	50%	50%
Temp. Correction Factor	1.048	1.048	-
Design Margin	1.1	1.25	1
Panel Array Sizing Parameters			
Panel Rated Capacity	85	85	85
Panel Peak Power Current	4.89	4.89	4.89
Peak Sun Hours	5.6	5.6	5.6
Array to Load Ratio	1.3	1.5	1.38
Coulombic Battery Losses	5%	25%	-
Wire Losses	0%	5%	-
Module Mismatch Losses	0%	5%	-
Module Aging Losses	0%	8%	-
Dust/Dirt Losses	0%	11%	-
Total System Losses	5%	54%	35%