LBNL-XXXXX

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Using Public Participation to Improve MELs Energy Data Collection

Margarita Kloss, Iris (Hoi) Cheung, Richard Brown,
Lawrence Berkeley National Laboratory
Alan Meier, Lawrence Berkeley National Laboratory and
University of California, Davis

Environmental Energy Technologies Division

August 2014

Presented at the
2014 ACEEE Summer Study on Energy Efficiency in Buildings
Pacific Grove, CA
August 17, 2014
and to be published in
the Proceedings

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

ACKNOWLEDGEMENT

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Program, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Using Public Participation to Improve MELs Energy Data Collection

Margarita Kloss, Iris (Hoi Ying) Cheung, Richard E. Brown,
Lawrence Berkeley National Laboratory
Alan K. Meier, Lawrence Berkeley National Laboratory and University of California, Davis

ABSTRACT

Miscellaneous and electronic loads (MELs) comprise an increasing share of building energy consumption. Large-scale data collection is needed to inform meaningful energy reduction strategies because of the diversity of MELs and our lack of understanding about how people use them. Traditional methods of data collection, however, usually incur high labor and metering equipment expenses. As an alternative, this paper investigates the feasibility of crowdsourcing data collection to satisfy at least part of the data collection needs with acceptable accuracy.

We assessed the reliability and accuracy of crowd-sourced data by recruiting 18 volunteers and testing our crowdsourcing protocol. The protocol asked volunteers to perform measurement tasks for three MELs devices of increasing complexity – 1) record power meter and MELs product characteristics, 2) identify and measure all power modes available, and 3) report the measured power.

Volunteers performed reasonably well for devices with functionalities with which they were familiar, but many could not correctly identify all available power modes in complex devices. Accuracy may improve when participants measure the power used by familiar devices in their home, or by providing more specific instructions, e.g. videos. Furthermore, crowdsourcing data collection from individual homeowners has the potential to generate valuable information about MELs energy use in homes when integrated with existing programs such as Home Energy Saver and Building America.

Introduction and Background

Miscellaneous and electronic loads (MELs) have proliferated in the last decade, and comprise an increasing share of building energy consumption (DOE 2012). These trends underscore the need for effective efficiency measures. Because of the diversity of MELs and our lack of understanding about how people use them, large-scale data collection could inform meaningful energy reduction strategies. However these studies are expensive, incurring significant labor costs for monitoring system deployment and maintenance, as well as metering equipment expense. MELs are ubiquitous, with 25 to 100 plug-in devices in a typical home. Crowdsourcing – the practice of soliciting volunteers to collect data on the device and power characteristics of MELs – is an appealing alternative to obtain at least part of the data needed to understand and eventually reduce MELs energy consumption.

Crowdsourcing, or citizen science, has long provided important contributions to scientific understanding. Since 1900 people have surveyed their local birds during the Christmas Bird Count (CBC), and the long term perspective of the CBC data has been used to flag imperiled species, pesticide contamination, or as an indicator of climate change (Audubon 2014). Internet and mobile technology have catalyzed large numbers of people for data gathering. Indeed, the

Cornell Lab of Ornithology and National Audubon Society later launched eBird, an online database where citizen scientists can submit data to inform bird abundance and distribution at various spatial and temporal scales. Several scientific papers (Fink et al. 2013 and Kelling et al. 2013) have been published based on the eBird database, demonstrating the value of eBird as an important tool for scientific contribution. The Search for Extra Terrestrial Intelligence @Home uses the Internet-connected computers of over 1.5 million volunteers to analyze radio telescope data (Korpela et al. 2011).

The objectives of this MELs crowdsourcing study were to understand the abilities and limitations of the public to collect and report MELs energy use information, and to test and refine crowdsourcing methodologies. Rather than doing field measurements in homes, we had test subjects measure power use under controlled conditions, on a set of products for which the power modes and levels were well understood by the research team. Due to the short time-setup of the experiment, we asked participants to perform power, rather than energy, measurements on test devices, but the goal was to use these results to refine methodologies for collecting MELs energy data. The study was divided into two phases. In the first phase, we recruited staff from Lawrence Berkeley National Laboratory (LBNL) as volunteers to participate in controlled experiments, with the goal of testing and improving our initial protocol for a second deployment. For the second phase of the study, we deployed the improved protocol and recruited volunteers at the 2012 LBNL Open House, which took place on October 13, 2012.

Controlled Experiment

Eight individuals participated in the controlled experiments in Building 90 at LBNL. These tests allowed the initial definition and refinement of the study protocol, including:

- Selecting devices for testing,
- Developing user instructions,
- Targeting key data for collection,
- Fine-tuning the data entry web-form for user input, and
- Streamlining video recording of experiments to supplement data collection and assist in resolution of post-experiment questions.

Device Selection

The project team considered a number of electronic devices for inclusion in the protocol, including computer monitors, music docking stations, computer speakers, rechargeable power tools, and others. We designed the protocol to begin with a relatively simple device, and to progressively scale up to more complex devices, so that we would be able to explore the abilities and limits of the public to discern and measure the power states. After a series of tests, we selected three devices, which were also used in the Field Experiment set-up as part of the Open House event.

- Level I - Lamp: As the first device presented for metering, the lamp represented a simple case of two power states: on and off.
- Level II - DVD player: The DVD player had four power states: on, standby, play, and tray-out. (Note: “On” refers to when the DVD player is turned on but performing no other function such as “play” or “tray-out”)
- Level III - Digital picture frame: The picture frame represented a more complex power usage scheme, and included a variable brightness setting: off, on-low, and on-high. We included the “low” versus “high” setting to test if participants could recognize that the brightness setting can be adjusted to more than one level.

In addition, we requested that participants separately measure the power consumption of any detachable power supplies (the test picture frame used in this study used a detachable power supply).

Protocol Refinement

The series of controlled experiments afforded the project team the opportunity to refine the study protocol in a stepwise manner. Researcher recommendations and participant feedback were incorporated into the subsequent experimental round. The revisions made to the protocol and experimental tools had been notably reduced by the last two subjects, as the protocol reached its final form:

- Information sheet & cards,
- Study web site,
- Meter selection,
- Participant instructions,
- Web form,
- Experimental checklist,
- Researcher script, and worksheet, etc.

Significant improvements during the controlled experiments included the following:

- Clarified and consolidated user instructions, data input, and exit survey in a single online web form;
- Fine-tuned the experimental setup to its final configuration, as shown in Figure 1 (Section 4 below describes in detail how the setup was used for data collection);
- Integrated human subjects requirements (as specified by the LBNL Human Subjects Protocol); and
- Developed an auditable system to track and record distribution of incentives to study participants.

Field Experiment

Our initial protocol underwent several iterations, based on the feedback we collected from both researchers and participants. We continued to recruit and test volunteers and improve

the protocol until it stabilized prior to the Open House event. We also moved the protocol survey from a Google Form to a commercial survey site, SurveyGizmo, to ensure data recording reliability.

Set-up

A schematic and photograph of the Open House set-up is shown in Figure 1. Each table measured 3'x6', and a digital camcorder was mounted on a tripod at one end to capture the entire experimental station. Video recording was active throughout the experiments, and researchers frequently checked the camera for participant anonymity, view angle, and adequate memory. Two tables were located five feet apart and faced different directions, in order to encourage volunteer concentration and minimize disruption and exchange between the experimental stations (See Figure 2, left).

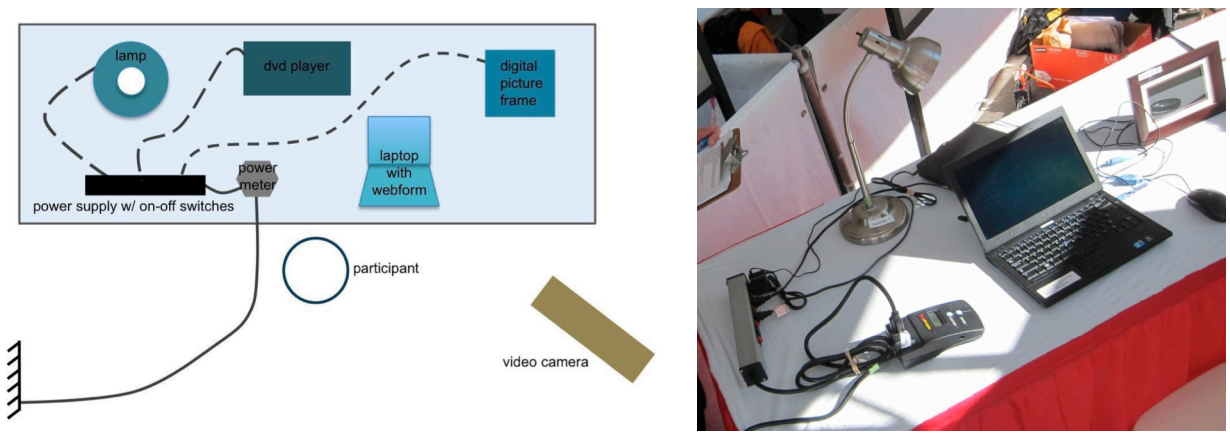

Figure 1. Experimental Set-up: schematic (left), photo (right). *Source*: LBNL

Each experimental station used a power strip with four outlets; each outlet had a separate hard on/off switch. The power strip was located at one end of the table and plugged into the power meter. Three test devices were plugged into the outlets on the power strip. When the participant was asked to perform measurements on a particular test device, the researcher switched on the device at the power strip, so that only the test device under evaluation received power, while the other two devices remained off. In effect, the power meter measured only the power strip and the test device. The central element of the test set-up was the laptop, which the participant used to access the web form. The web form had instructions to guide the participant, served as the portal for data entry, and included the exit survey. Volunteers were provided a magnifying glass to ease reading device nameplate information, which is frequently very small type.

A different power meter was used at each table - the "Belkin Conserve Insight"[1] and the "Watts up? PRO ES"[2] (Wattsup Pro) - to draw insight on the impact of power meter design on measurement results. Both meters are user-friendly and designed for use by the average

[1] Belkin Conserve Insight - http://www.belkin.com/us/F7C005-Belkin/p/P-F7C005;jsessionid=A8EE0A5FC1E43A04D60FB154763A10F0.
[2] Wattsup Pro - https://www.wattsupmeters.com/

consumer; the Wattsup Pro is also widely used in the plug-load research community because of its data logging capabilities. By advancing through a few screens, the user can view readouts of the test device's power consumption in watts, accumulated energy use in kilowatt-hours, and the annual energy cost of operating the device.

The research team designed the protocol to begin with a relatively simple device, and to progressively scale up to more complex devices. This stepped approach allowed researchers to explore the abilities and limitations of the public in discerning and measuring power states in the test devices.

Figure 2. Two experimental stations (left); recruiting table (right). *Source*: LBNL

Recruiting. The crowdsourcing experimental set-ups were adjacent to the Standby Power Exhibit, which invites participants to play with common household plug loads and read power measurements for different operational and power modes. Power modes of a device are distinguished by its power draw depending on the different operational settings (e.g. on, off, and standby mode). The exhibit often attracts a crowd; recruiting benefited from synergies by co-locating the crowdsourcing experiment adjacent to the Standby Power exhibit.

A recruiting table was located at the entrance to the tent that contained the experimental set-ups (see Figure 2, right). We displayed two common power meters that could be used at home to measure the energy use of plug-load devices and determine annual energy consumption. A researcher staffed the table throughout the day to recruit volunteers, and followed the approved LBNL Human Subjects Protocol (written for the purpose of this experiment). Researchers hosted a continuous flow of participants and both experimental stations were occupied throughout the entire Open House event.

Experiment. After the volunteer consented to participate in the experiment, they were assigned to one of the test set-ups. Researchers briefly explained the purpose of the test, and introduced the volunteer to the different elements of the experimental set-up, including the web form, power meter, and test devices. The researcher at each table encouraged the participant to follow the instructions on the web form as closely as possible, enter the data they collected, and ask any questions that might arise. (See Figure 3) As participants proceeded with the experiment, researchers timed each step, for example the elapsed time to identify product information and to meter each device. Researchers also noted any observations or problems encountered by the volunteers.

Figure 3. Researcher observing experiment. *Source*: LBNL

Web Form. The web form prompted the participant to fill in a unique identification number assigned by the recruiting staff, which linked the web form data, video recording, and incentive award. The survey questions stepped the participant through the protocol and guided volunteers to identify the meter manufacturer name and model number; product metadata, including type, model number, rated power, and available power settings; and subsequently to measure the power consumption in each power mode. If a device included a power adaptor, the web form asked the volunteer to measure its energy use alone. The web form concluded with an exit survey, which elicited participant feedback and anonymous demographic information.

Results and Discussion

Below we summarize our preliminary findings from the Open House experiment and describe how they can inform MELs crowdsourcing data collection in homes in the future.

Meter Identification. We asked the volunteers to identify the manufacturer name and model number of the power meter. The Belkin Insight meter information is only available near the meter plug, and one would need to unplug the meter in order to read the information. To avoid confusion, we attached a fictitious label near the meter screen that read “Belkin xyz123,” which made it easy for those users to identify meter manufacturer and model number. Three of the eight

volunteers (38%) listed "Electronic Educational Device, Inc." as the meter manufacturer for the Wattsup Pro, while the rest filled in the model name in this field. There appeared to be confusion between manufacturer and model names, especially with the Wattsup Pro as the model number was clearly listed on the meter, and the manufacturer name could only be found on the nameplate located on the back.

Product Identification. As indicated previously, the three test devices were a lamp, DVD player, and digital picture frame (including an external power adaptor). Seventy-eight percent, or 14 of the 18 volunteers recorded the model number on the DVD player correctly (one recorded a partial model number, two identified the serial number as the model number, and one volunteer could not locate the model number). All but one of the 18 volunteers (94%) recorded the model number on the digital picture frame correctly. Finally, 16, 14, and 6 of the 18 volunteers correctly identified and recorded the rated power for the lamp, DVD player, and digital picture frame, respectively (89%, 78%, and 33%). Many appliances list their rated voltage and current (amps) rather than power, when both sets of information were present; this was the case for the digital picture frame. Since volunteers were less familiar with electrical characteristics expressed in volts and amps, many of them recorded the volts and omitted amps.

Mode Identification. The web form asked volunteers to independently identify up to four available power settings for each device. The results are plotted in Figure 4. For the simplest device – the lamp, almost all volunteers were able to identify both the *on* and *off* modes (100% in *on* mode, 94% in *off* mode). A majority of the volunteers also identified the *on* and *standby* modes of the DVD player correctly (83% and 89%), but only a third detected that the DVD *tray-out* was a separate power mode. When investigating the digital picture frame, most volunteers recognized the *on* mode (89%), and the power adaptor being unplugged as a separate mode (83%). Several volunteers discovered the brightness adjustment button, which was located on the side of the digital picture frame, but only one recognized it as a different power state to be measured separately (6%).

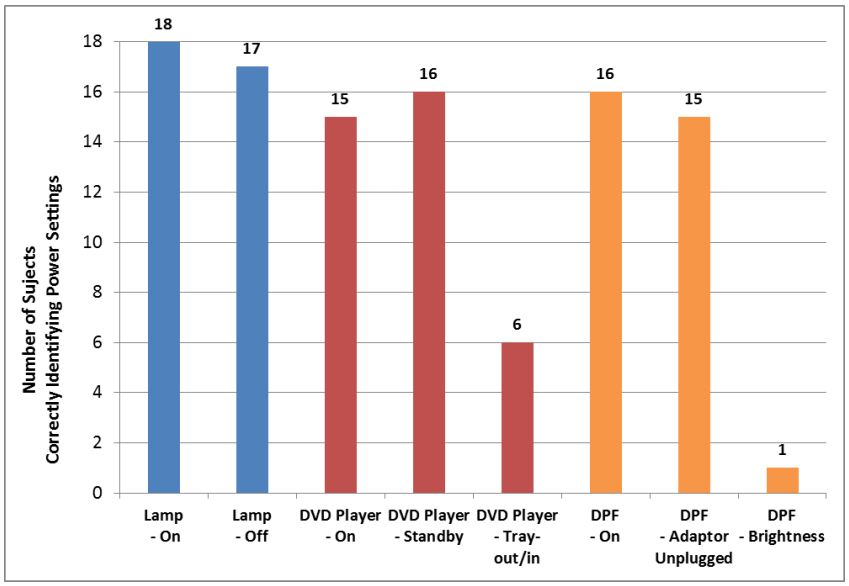

Figure 4. Success in identifying available power modes of MELs device (18 volunteers). *Source*: LBNL

Power Measurements by Mode. Along with the ability to identify available power modes, we assessed the ability of volunteers to accurately perform the corresponding power measurements. Figures 5 and 6 display the measured power of selected operational modes reported by volunteers for the three test devices. As shown in Figure 5 left, the measured on-mode power for the lamp ranges between 15 and 17W. Some of the variation resulted from the two different light bulbs with slightly different power ratings in each of the lamps. In addition, power measurements may be sensitive to when the measurement is taken, due to bulb warm-up. Variations in power measurements (possibly caused by heating, filtering, and other processes) should be anticipated during development of large-scale crowdsourcing protocols, to ensure data quality and usability. For example, the protocol should specify how long the participant should meter different types of products, and whether they should average the readings taken over time.

The digital picture frame features included variable brightness. The data shown in the right panel of Figure 5 reflect power measurements at low brightness (~1.7W), and full brightness (~4W). Some volunteers were not actually aware that the brightness could be adjusted. In an actual crowdsourcing deployment, operational and power modes could be difficult to detect. These and similar issues should be considered when developing a large-scale crowdsourcing protocol.

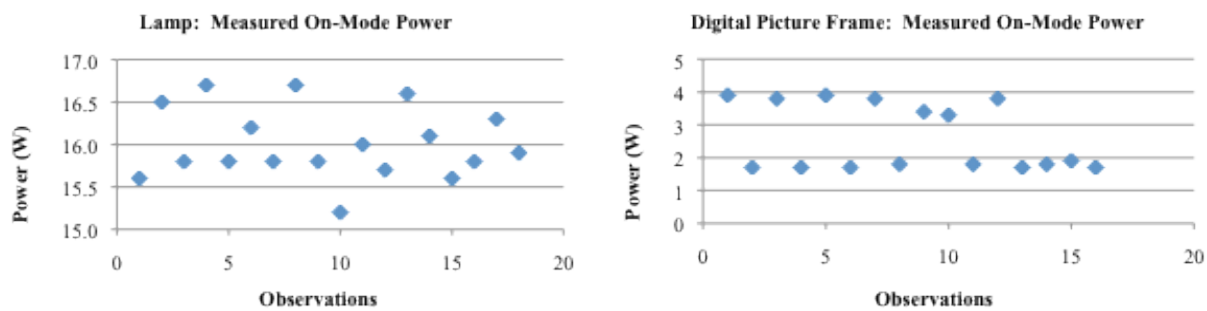

Figure 5. Lamp: *on* mode power (left); digital picture frame: *on* mode power (right). *Source*: LBNL

Volunteers reported consistent *on* mode power measurements for the DVD player, (Figure 6, left), except for one person who reversed *on* and *off* power measurements. Figure 6, right, shows the reported power measurements when the DVD player was in *tray-out* mode. Only six volunteers (33%) successfully identified and measured this power mode. Five of the six volunteers reported measured power between 4.5 and 5.2W, one volunteer reported 8W (the volunteer did not take the measurement until after the DVD tray closed again, therefore the lower measurement reflected *on* mode power).

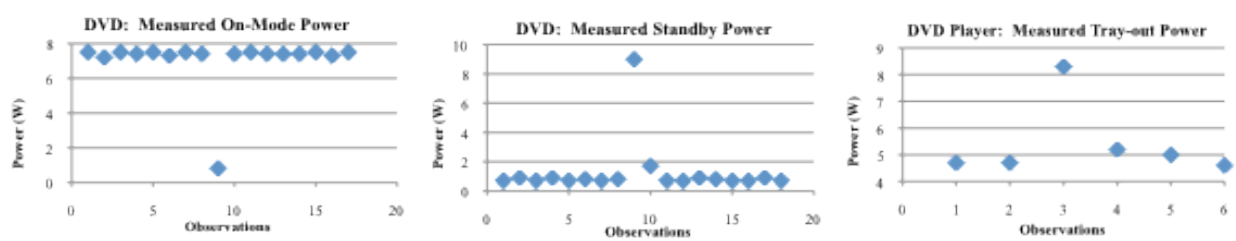

Figure 6. DVD player: *on* mode power (left); *standby* power (center); *Tray-out* power (right). *Source*: LBNL

Measurement Time for Each Test Device. Researchers recorded the elapsed time required for volunteers to measure each power mode, and identify meter and device information. Table 1 summarizes the length of time volunteers spent measuring each device, including collecting device characteristics (manufacturer name, model number, and rated power), identifying available power modes, and measuring corresponding power levels.

Although the volunteers recruited at the Open House were not expected to be a representative sample, the data provide valuable insight on data collection times. The participants required, on average, about the same amount of time to measure each device (see Table 1). However, there was a wide range between the minimum and maximum measured times across all three devices, demonstrating a spectrum of abilities among volunteers. The lamp was the simplest device, but since it was the first device to test and meter, volunteers likely spent more time familiarizing themselves with the power meter and the test set-up. Crowdsourcing participants will probably need less time to measure energy use in their own homes because they are likely more familiar with the device features.

Table 1. Time needed to collect device data, per volunteer (minutes)

	Lamp	DVD player	Digital picture frame
Mean	5.0	5.4	5.5
Median	4.9	5.5	4.8
Min.	2.8	2.8	3.5
Max.	7.8	8.1	11.3

Source: LBNL

Conclusions and Recommendations

In this section we discuss the quality of crowdsourced data collection, and recommendations and next steps. Due to the small size of the study sample, limited by the number of volunteers we were able to recruit over the one-day Open House event, the results discussed below are based on our observations from the collected data.

Quality of Data Collection

Crowdsourcing as a data collection strategy will not be viable if the citizen measurers (CMs) provide erroneous data filled with gaps. A major focus of this research aimed to assess the quality of data collected by CMs. Eighteen CMs participated at the 2012 LBNL Open House; each volunteer recorded nameplate data of the power meter and test devices, identified the power modes of three devices, and measured and reported the power consumption for each mode. In this experiment, we evaluated the abilities of the CMs to accurately record product characteristics data; reliably detect different power/operational modes; and accurately measure and record power levels.

In addition, we explored the impact of increased familiarity with equipment and tasks. The above skills are needed to assemble data to understand appliance and plug-load energy use and to put the measurements in a broader context. Note that the sample size of 18 participants in the experiment is relatively small, and not representative of the general population. Nonetheless, these results could inform the focus and protocols for future crowdsourcing data collection efforts.

Product Characteristics . The CMs were able to accurately record the product type (i.e. lamp, DVD player, and digital picture frame). They were less successful in identifying the product model number; the overall error rate was 14% (22% for the DVD player and 6% for the digital picture frame). When recording the product's rated power, CMs encountered the electrical characteristics of some devices that were expressed in terms of volts and amps instead of watts. This confused some CMs, and the success rates were approximately 90%, 80%, and 30% for the lamp, DVD player, and digital picture frame, respectively.

Mode Detection. A high number of CMs correctly identified the *on*, *standby*, and *off* modes, ranging from 80-100% (see Figure 4). Errors increased as the devices became more complex, and modes required taking special actions, such as the DVD *tray-out* mode (success rate of 33%), and the brightness setting of the digital picture frame (success rate of 6%).

Power Consumption. The CMs were able to measure power consumption reliably if they could identify the operational modes correctly. For example, in the case of the DVD player, only 1 of 18 measurements was erroneous because one CM transposed the standby and active power measurements.

Product Familiarity. We observed that the CMs were able to measure and record product data more rapidly as they progressed through the tasks, showing that they were learning and becoming more familiar with the power meter and the measurement protocol. We also noticed that CMs took a significant amount of time to study the products to understand the labeling and power modes. We suspect that measurement times and error rates will decline when CMs measure devices they use regularly in their own home. To characterize the energy use of complex products, more specific instructions or perhaps a video would improve participant responses and guide CMs through product functionalities and power modes.

Recommendations and Future Work

Based on the findings of this project, we conclude that crowdsourcing data collection can be a useful and inexpensive way to gather MELs data for simpler devices but may not produce highly accurate field measurements. Our assessment focused on power measurements, but energy data would be of greater interest to inform energy efficiency policies. Measuring energy use over a period of time (as opposed to a measurement of power draw) offers simplifications and new complexities. For example, identification and measurement of specific modes would not be a concern. On the other hand, the CMs will need to engage for longer periods, and deployed meters will need to be able to store data or transmit data to a central location. Finally, we need a deeper understanding of how CM behavior may impact collected data quality.

The crowdsourcing protocols should carefully consider the logistics of incentive distribution to CMs, distribution of meters, and protection of participant privacy.

This research explored the reliability and accuracy of MELs data collection through crowdsourcing and described protocol recommendations. Crowdsourcing could be extended to test ongoing DOE Building America projects in homes. This could further strengthen the data collection protocol developed as part of this project – add video instructions, test energy measurements, and others. Following a version of our crowdsourcing protocol, a data submission module could be added to the Home Energy Saver™ (HES) tool, allowing volunteers to collect and upload MELs energy data for the devices used in their homes. The collected data could be used to improve MELs energy estimates of the HES tool, and could in turn inform purchasing guidance for MELs devices. Both of these methods would explore data collection for energy instead of power measurements alone, and usage patterns including the time devices spent in each available mode. Finally, wireless power meters that enable device control through mobile and remote applications, such as the Modlet, could also be tested.

By exploiting digital and mobile technologies, citizen science and crowdsourcing give volunteers the ability to participate at the local level while transforming individual efforts into a collective endeavor with wide implications and possibilities.

References

Audubon Society. 2014. http://birds.audubon.org/history-christmas-bird-count; http://birds.audubon.org/how-christmas-bird-count-helps-birds [accessed March 14, 2014].

Fink, D., Damoulas, T., Dave, J. 2013. "Adaptive Spatio-Temporal Exploratory Models: Hemisphere-wide species distributions from massively crowdsourced eBird data." Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13) July 14–18, 2013 in Bellevue, Washington, USA.

Kelling, S., Lagoze, C., W. K. Wong, J. Yu, T. Damoulas, J. Gerbracht, D. Fink, C. Gomes, 2013. "eBird: A Human/Computer Learning Network to Improve Biodiversity Conservation and Research", Artificial Intelligence Magazine, vol. 34(2), 10-20.

Korpela, E.J. et al. 2011. "Status of the UC-Berkeley SETI Efforts." http://setiathome.berkeley.edu/sah_papers/status_of_ucb_seti_efforts_2011.pdf.[accessed March 14, 2014].

US DOE. Annual energy outlook 2012. http://www.eia.gov/forecasts/archive/aeo12/index.cfm [accessed Sept 29, 2013].