

SunDial Integration of Building Load Management, Solar PV, and Energy Storage to Support the Electric Grid: Lessons from a Field Pilot

*Matthew Kromer, USA Center for Manufacturing Innovation
Kurt Roth, Fraunhofer, USA Center for Manufacturing Innovation
Jim Boch, IPKeys Power Partners*

ABSTRACT

We developed and piloted SunDial, a virtual power plant platform that enables integration of high-penetrations of solar PV within a distribution grid. A vendor-agnostic control platform (the “Global Scheduler,” GS) optimally shaped the net load for a virtual portfolio of non-co-located distributed energy resources (DERs) based on user-defined policy objectives, thereby mitigating challenges associated with deploying solar PV on renewables-dominant distribution networks while providing grid-facing services that can support both the bulk and local power systems. During the pilot, the GS dispatched ~200kW of flexible loads and a 0.5MW/1MWh energy storage system (ESS) based on forward-looking predictions of solar production (1.5MW, peak) and facility demand to address use cases including peak shaving, energy cost optimization, peak-power dispatch, and power firming, while minimizing rapid changes in solar production.

Field testing during the 15-month pilot on a National Grid distribution feeder found that integrating highly granular time-series predictions effectively increased storage capacity relative to a non-predictive baseline by upwards of 20%. Furthermore, the results highlight the need to manage prediction uncertainty, e.g., by maintaining sufficient ESS reserve capacity. Flexible loads were dispatched repeatedly to supplement battery energy storage, comprising the equivalent of ~10% of additional storage capacity. We discuss several challenges impeding scaling of integrated systems, focusing on practical difficulties and solutions for efficiently integrating and characterizing flexible facility loads into a virtual power-plant platform.

Introduction

Background

In response to decreasing costs coupled with policy incentives, the penetration of solar PV has increased rapidly over the past two decades. Most analysts expect this growth to continue, with some projecting solar PV as the lowest cost electric resource in all U.S. states by 2030 (Merchant 2021). The intermittency of PV generation can create several well-known challenges for electric distribution grids as PV penetration increases. On shorter (seconds to minutes) timescales, large fluctuations in PV generation due to changes in cloud cover can create problems including voltage instability, reverse power flow (“backfeed”), increased wear on distribution equipment (e.g., load tap changers [Ari 2001]), and single-phase imbalances. On a multi-hour timescale, PV generation may exceed power consumption, creating the potential for sustained power reverse power flow that most of the grid is ill-equipped to handle, as often illustrated by the know-famous (or infamous) “duck curve.” In addition, high power ramp rates, particularly later in the afternoon when PV generation wanes and electric demand waxes, require a coordinated effort to bring large generation resources on-line. An evolving regulatory

landscape, exemplified by, e.g., California Rule 21 and IEEE1547a, allow for solar inverters to help mitigate voltage issues and/or cascading power failures through the injection of reactive power (i.e., VARs). However, ultimately, the only solution for mismatched supply and demand is the ability to inject or absorb real power in response to real-time conditions. While this supply/demand mismatch can be addressed at the bulk power system level, tools and technologies that enable *localized* load-shaping at the level of the distribution system have numerous advantages if they can: 1) mitigate or eliminate congestion on both the transmission and distribution systems; 2) reduce system peaks at both the local- and system level; 3) limit the need for infrastructure upgrades; 4) reduce line losses; and 5) enable higher penetrations of renewables on local networks.

At present, the primary mechanisms to address the localized impacts of potential supply/demand mismatches due to distribution-sited solar PV are either: (1) the denial of interconnection applications in congested areas, or (2) requiring the applying customer to undertake costly upgrades on distribution equipment, which in most cases effectively prevents the project from moving forward. Strategic deployment of bulk energy storage systems (ESS) on distribution networks can help address these challenges, but adoption has been limited due to high upfront cost, and limited ability to fully monetize the benefits of multiple “stacked” value streams. Difficulty monetizing the benefits of distributed storage reflects both an immature market for localized distribution-level services; and technical challenges stemming from a lack of toolsets that can reliably identify an optimal dispatch strategy (given uncertain current- and future system states) and coordinate this strategy across a portfolio of multiple DERs.

Although markets exist to address the real-time value of power at the transmission grid scale, e.g., demand response, capacity, ancillary services, etc., solutions for the *distribution* grid scale remain elusive. Due to the greater range of distribution grid load shapes, problems can arise on the distribution grid *prior* to affecting transmission. Furthermore, commercial and industrial (C&I) customers can currently participate in transactive energy markets but, excepting larger facilities, most do not to a significant extent due to the cost and challenge of implementing the necessary infrastructure. Finally, using demand-side resources to address PV-related grid challenges differs from capacity-focused markets in several important ways (see Table 1).

Table 1: Differences between traditional DR and high-penetration PV use cases.

Attribute	Traditional DR	Load Shaping to support high-penetration PV
Frequency of Calls	~10-20 times/year	Majority of days/year
Key Goal	Load shedding	Load sinking and shedding to increase PV and load coincidence
Duration of Load Management	1 to 4 hours	Up to 8 hours/day

To address these challenges, Fraunhofer USA developed SunDial, a virtual power plant platform that provides integrated control of battery energy storage systems (ESS) and commercial and industrial (C&I) facility loads to mitigate distribution grid challenges caused by high PV penetration. We conducted a 15-month pilot of the system on a National Grid distribution feeder in Shirley, MA. The vendor-agnostic control platform (the “Global Scheduler”) optimally shaped the net load for a virtual portfolio of non-co-located DERs based on user-defined policy objectives. The demonstration comprised a 0.5MW/1.0MWh electric energy storage system (ESS) adjacent to two National Grid PV systems totaling 1.5MW, and

~3.8MW (peak) of customer loads spread across three commercial and industrial (C&I) facilities: a school; a food-processing facility; and a food-production facility.¹

The rest of the paper presents the SunDial system and its components, the participating facilities, the results of field testing, and conclusions and recommendations to enable integration of building load management at scale.

The SunDial System

Overview and Architecture

SunDial is a virtual power plant (VPP) framework for optimizing power flow on distribution systems in high-penetration solar environments by coordinating the behavior of multiple distributed energy resources (DERs), regardless of whether these assets are owned and operated by different entities, and regardless of whether resources are co-located behind a single point of common coupling. Notably, the SunDial architecture separates the complexity of facility energy management, where facilities can have widely varying load profiles, load management capabilities, and tariff structures, from distribution grid power optimization.

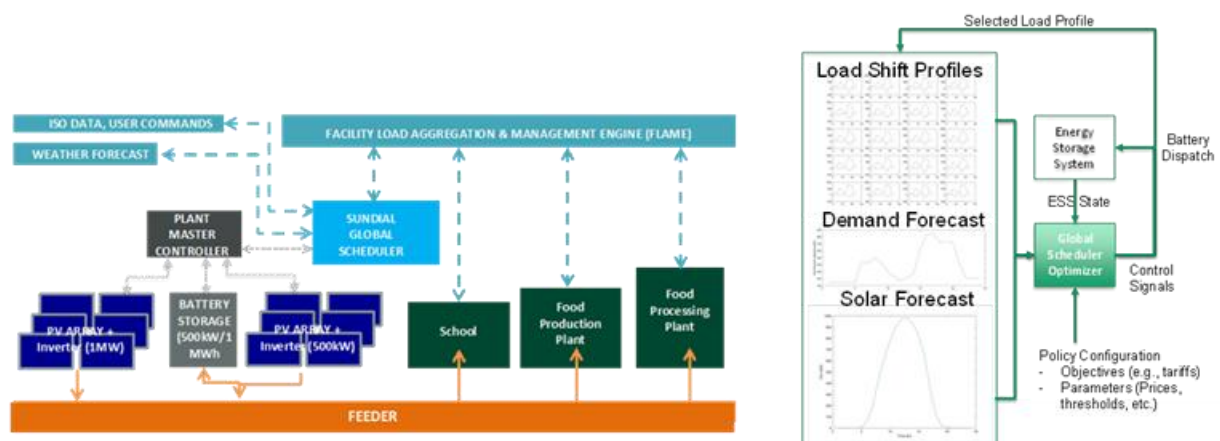


Figure 1: Block diagram of the SunDial system (left) and the SunDial Global Scheduler (GS; right).

The SunDial pilot described in this paper consisted of a total of 1.5 MW of solar at two different plants, 500 kW/1MWh Energy Storage, and a portfolio of three C&I facilities with some degree of load flexibility (approx. 200kW in total). In addition to these DERs, the key components of the SunDial System (Figure 1, left) consist of (1) the Global Scheduler (GS), which acquires and evaluates performance data from participating PV systems(s); (2) the Facility Load Aggregation and Management Engine (FLAME), which forecasts, aggregates, and manages building-level loads and their load flexibility; and (3) external data services, such as weather predictions and utility / independent system operator (ISO) electric market data.

¹ Due to large differences in the peak loads among the facilities and the total load (3.8MW) relative to the PV 0.5MW) and ESS resources (0.5MW/1.0MWh) available, we applied different scale factors to individual loads to approximately equalize peak demand across the three sites and normalize total demand to approximately equal total solar plus ESS capacity (food production facility = 0.1; food processing: 0.2; school = 1.0).

Global Scheduler

Overall control of the SunDial System was orchestrated by the Global Scheduler (GS), a vendor-agnostic optimal dispatch controller developed by Fraunhofer as part of this project. We implemented the GS as an extension of the VOLTTRON platform, an open-source DER operating system developed by PNNL, and deployed the GS on a Linux server at the Solar + ESS plant (see Kromer et al. 2020). Monitoring and control of endpoint devices at the PV and Storage sites was implemented via a local Modbus connection to a site plant master controller. These devices included the PV inverter; an ESS master controller (which aggregates data and controller for the ESS); meters located at the output of the ESS inverter, PV inverter, and the point of common coupling; a weather station; and the site's switchgear. Monitoring and control of demand side resources was implemented via connection to a cloud-based server (the FLAME) developed and operated by IPKeys, a curtailment services provider (CSP).

IPkeys instrumented participating C&I facilities with telemetry to provide 1-minute power data to the GS in near-real time (typical latency of 5 to 15 minutes), and to enable actuation of flexible loads based on GS commands. Data from individual customer facilities were collected by the FLAME and communicated to the GS. The GS also queried external data sources, including Clean Power Research's satellite-based solar prediction service, ISO New England's web API, a power meter installed on the local distribution system (to monitor feeder power flow), and an SSH connection accessible by the project team for remote monitoring and control. We logged system data to a local database at a time resolution down to 1 second (e.g., PV production and ESS data), and was replicated on a remote Fraunhofer server nightly.

Based on the current and forecasted state of the system and pre-configured policy objectives, the GS determines a least-cost deployment strategy for its portfolio of DERs and issues control signals to actuate this least-cost pathway. This cost minimization function primarily focuses on controlling power flow within a feeder to, e.g., minimize reverse power flow on the transmission system, minimize demand peaks within a nodal area, or follow a load shape, while simultaneously minimizing cost. The specific optimization objectives may be reconfigured to reflect the tariff structure relevant to each deployment.

The GS's optimizer is implemented using a simulated annealing (SA) optimization methodology. We selected SA due to its ability to generate an approximately optimal solution for large combinatorial problems, to address non-convex and discontinuous search spaces, and its applicability to a broad range of optimization problems (see the literature review in Kromer et al. 2020 for additional discussion, as well as Zeifman, Kromer, and Roth (2017).

Facility Load Aggregation & Management Engine (FLAME)

The FLAME has the following basic functions:

1. Generate baseline aggregate load predictions for the next 48 hours for the participating facilities (discussed in detail in Kromer and Roth 2020).
2. Generate potential aggregate managed load profiles for the next 48 hours that the participating facilities could provide based on their collective load flexibility.
3. Identify the n number of managed load profiles yielding the lowest operating cost for the participating facilities in aggregate, based on 48-hour GS electricity cost projections.
4. Communicate real-time facility load, baseline aggregate load predictions to the GS.
5. Negotiate managed load prediction strategy with the GS.

6. Actuate the load control actions required to produce the aggregate managed load profile selected by the GS.

FLAME Architecture

In our model, the CSP (IPKeys) presents baseline load predictions and a portfolio of potential managed load profiles to the Global Scheduler through the FLAME interface. To integrate with the FLAME, we developed a new communications protocol, implemented as an extension of OpenADRB, supporting the deployment of load-shaping services as a means to facilitate high-penetration of solar. Current demand response protocols, such as that defined by OpenADR 2.0b, enable counter-parties to negotiate for blocks of DR services (e.g., “shed XX kW load for YY hrs at time ZZ”). However, these existing protocols do not offer counter-parties the ability to negotiate how to use demand-side assets for load-shaping, i.e., a combination of up- and down-regulation of loads at specific times of days. Load-shaping is a more sophisticated and valuable service than traditional DR for enabling high-penetration PV in that it supports local consumption of power during periods of excess generation to minimize power backfeed and reduction in consumption during system peaks, while providing facility operators with a tool to comprehensively assess the impacts of shifting load (Roth et al. 2020).

IPKeys led the FLAME development, leveraging IPKey’s Certified OpenADR 2.0b VTN and VEN technology. This system provides a secure and standardized control and monitoring system that allows real-time control of devices from any certified manufacturer. Consistent with the standard, we divided the duties of the FLAME into the Virtual Top Node (VTN) server, i.e., the FLAME, and Virtual End Node (VEN) client-side devices located at facilities. These interact as follows to generate aggregate baseline load predictions and load management potentials:

1. Each VEN transfers facility power data to the FLAME server via an OpenADR report.
2. The VTN receives these reports from all participating facilities and calculates the predicted baseline and load control offer curve for each facility.
3. The VTN combines them into a feeder level predicted baseline and a family of offer curves based on the projected needs of the system.
4. The VTN then uses hourly price signals from the GS to calculate the combination of load curves that yields the lowest cost over the next 48-hour time horizon.

Instead of a monolithic system where all decisions are made at the server, this system uses a Service Orientated Architecture that separates the information input and decisions to the area where they are best known. Defining the interface and the data exchange parameters creates a system that can incorporate any device that can publish its predicted consumption and control capabilities in an OpenADR 2.0b report. In this case, we expressed GS costs to the FLAME as a function of both time of day and demand level (potentially with multiple levels).

Ultimately, in response to a request from the GS, the FLAME sends an arbitrary number of low-cost curves (e.g., 20, 50, more) and their associated costs are sent to the GS. The GS selects the profile to implement, communicates that decision back to the FLAME VTN, and the FLAME VTN dispatches signals to the selected facilities to implement the control actions comprising the requested load shape. To simplify the search for the lowest cost managed load options to send to the GS, IPKeys applied a greedy version of the knapsack algorithm. Reviews of the options provided found that the facility-specific load shifts selected were consistent with the facility-specific constraints and aligned with the price maps provided by the GS.

The FLAME VTN is implemented as a cloud-based service to increase its resiliency. The FLAME VENs that actuate events at the facilities are local gateways (IPKeys EISSBoxes) that

control dry-contact switches at each facility. In addition, the FLAME sends emails to all facilities ahead of and upon initiation of events scheduled by the FLAME.

Facility Recruitment

We worked closely with National Grid’s energy efficiency (EE) program to recruit pilot participants. The team approached potential participants with the following value propositions:

1. Free audit of demand response (DR) potential.
2. Free installation of automated demand DR equipment (cost usually \$4,000 to \$8,000).
3. Free management of facility electrical demand with a total value of ~\$50-60,000/year for a 1MW customer, including:
 - Managing monthly demand costs (\$5.76 or \$8.50/kW, 15-minute peak power).
 - Managing annual ICAP tag: An annual charge by ISO New England (ISO NE) based on their demand during the annual peak hour for ISO NE. The ICAP charge grew from ~\$3/kW/month in 2015 to \$109/kW/year in 2018.
 - ISO NE DR markets: Customers with eligible DR capabilities participated in ISO NE’s forward-capacity markets, sharing the \$80/kW/year in 2018 with the CSP
 - National Grid DR Pilot: Incentives of ≤\$35/kW shed during peak periods
4. Free access to near real-time high-resolution (e.g., 1- to 5-min.) facility power data
5. Help the Shirley-Ayer area to become more sustainable by integrating more PV
6. Publication/press/media opportunities for participating customers.

Three C&I facilities decided to participate in the pilot (see Table 2), for reasons including a strong relationship with the National Grid EE program, desire to gain experience with forward-looking energy programs (for a multi-national), and town sustainability goals.

Table 2: Participating C&I facilities in the SunDial demonstration and their characteristics.

Facility	Peak Demand	Load Managed	Load Shift	Duration
Food Processing	1,300 kW	Conveyor drags	90kW	Runs 5 hours on weekdays between 7AM and 7PM, default estimated to be 9AM-2PM
		Secondary Process	70kW	Can be paused on occasion (e.g., once week maximum) for 1-2h
Food Production	2,450 kW	Electric Forklift Charging	Four-hour event: 24, 19, 12, 12kW	EVs typically used 2PM to midnight, default charging midnight to 4AM; no pausing during charging
School	290 kW	Packaged rooftop unit AC (zone temperature setpoints)	Varies, maximum ~35kW	Magnitudes and durations vary with T_{out} , during summer vs. school year (internal loads).

Flexible Facility Loads

The SunDial architecture delegates the complexity of facility energy management to the FLAME because facilities can have quite different load profiles, load management potentials, and tariff structures. That is, individual facilities (or their CSP) best understand their specific opportunities and constraints and how those affect the grid services they can provide and their cost. Thus, the FLAME calculates managed load profile options derived from models for the facility-specific load management options informed by current and forecast operating states.

We worked with operations staff at each facility to identify significant loads that could be shifted without compromising core operations and determined the acceptable operating envelope for each flexible load (see Table 2). Discussions follow for the individual facilities.

Food Processing: This facility was not willing to automate actuation of flexible loads due to concerns about affecting production. Instead, the FLAME communicated to the operations manager preferred times to run different processes via a daily email and indicators lights on a panel installed by IPKeys (lights actuated via an EISSBox gateway). Another challenge was that the start and run times for the conveyor drags varied day-to-day, e.g., run-time data from the control system showed that runtime over seven “events” varied between 2:37 and 12:53. Based on that and conversations with the operator, we estimated a typical start time of 8AM and a 5-hour runtime, and determined that a 5-hour event could occur anytime between 7AM and 7PM.

Since we lacked access to submetered power data for the drags and the other process we could pause, IPKeys used minute-resolution facility-level power data to calculate the power consumption differences between the 60 minutes prior to and after actuation of a specific event and the power consumption during an event. Due to the multi-hour duration of transfer drag runs and other processes potentially coming on line during that time, we found appreciable variability in the estimated power increase among “tests.” Averaging over 12 events yielded 90kW for the drags; we used a similar process to estimate the power for the other process we could control.

Food Production: We controlled the charging time for 12 electric pallet rider vehicles (EVs) at this facility. To automatically enable and disable EV charging, we installed an EISSBox that reports the state of and controls the 20A 480V three-phase contactors of the power circuits controlled, as well as total site power consumption.

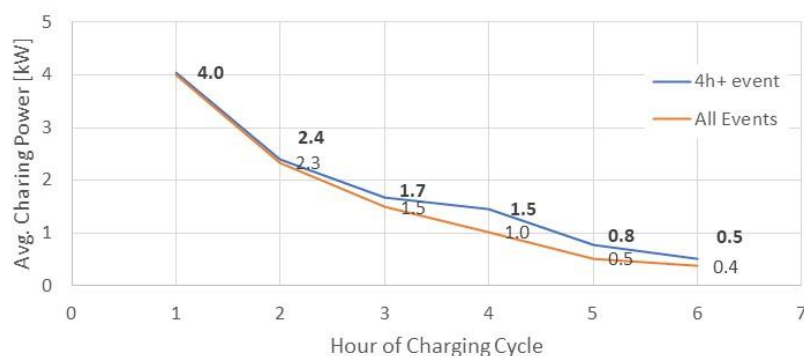


Figure 2: Average EV power for all charging events and charging events of 4+ hours, for two pallet rider EVs.

The EV charging power values were relatively small relative to total facility power draw, i.e., ~30-50kW versus ~250kW overnight and 600-1,200kW when operating. We estimated typical charging profiles for the pallet rider EVs using power data for two EV charging circuits (each serving two EVs), confirming with video data that the other EVs had similar usage schedules. The data for the two pallet riders show appreciable variability in the duration of charging event, with a four-hour charging event most common. The start times for charging vary appreciably as well, with about half of charging events starting between 9 and 11PM. Analysis of 90 charging events yields the average profiles shown in Figure 2; the fleet of 12 pallet-rider EVs produces hourly loads of 24, 14, 10, and 9kWh for events lasting at least four hours.

School: We shifted loads by adjusting the zone-temperature setpoints, $T_{z,set}$, that control the AC systems in the packaged rooftop units (RTUs) via the school’s building automation

system (BAS). Increasing $T_{z,set}$ turns off the AC in zones until the zone temperatures rise to the new $T_{z,set}$, decreasing school power draw (shed), while decreasing $T_{z,set}$ causes the AC in more zones to run and for a longer time. Working with the school, we decided that load shifting events would have a 3°F precool followed by a 3°F shed (see Figure 3). This minimizes overheating during the summer, maintaining occupant comfort, while providing approximately an energy-neutral load shifting (at least for an RTU; see Lin et al. 2017). In contrast, a straight load shed would reduce demand but compromise comfort, making it of limited value for applications with numerous events. Over most of the demonstration period, we could control 6 of the 7 RTUs; one RTU's control board had failed earlier, preventing its control via the BAS, and another board RTU failed after the other one was repaired. The FLAME actuated events via a second EISSBox gateway, switching dry-contact switches that provide control inputs to the BAS.

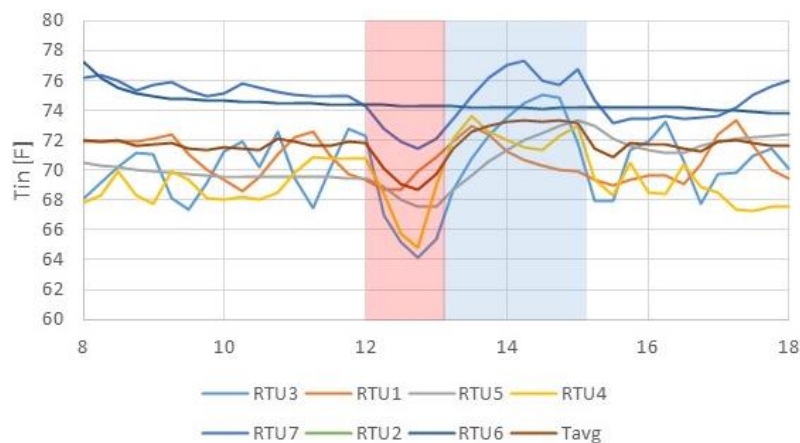


Figure 3: Zone T_{in} profiles for summer 2019 event with 3°F precool from 12-13h and 3°F shed from 13-15h.

Estimating baseline AC consumption and potential managed load profiles for AC systems is not straightforward. Space cooling loads vary as a function of T_{out} as well as outdoor (ventilation) air volumes, and internal heat gains, which vary over time and also by zone. In addition, many RTUs have multiple cooling stages, and AC capacity and COP vary with T_{out} , and can degrade due to faults. Shifting AC loads over longer time periods also requires accurate modeling of building thermal mass. IPKeys developed a model for school AC loads as a function of T_{out} and $T_{z,set}$ that they used to model baseline AC consumption as a function of time and T_{out} for baseline conditions and potential load-shifting events. Fraunhofer worked with IPKeys to refine the model in light of whole-building and RTU power, and zone temperature data from precooling (i.e., decrease $T_{z,set}$) and load shed events to better characterize the system's response.

IPKeys assumed that all of the AC load predicted by their model could be controlled; for multiple reasons, that was not the case. First, throughout most of the pilot one or two RTUs could not be controlled by the FLAME at different points in time due to control-board failure. Second, the RTUs account for only a portion of the school AC load, i.e., there are multiple additional split-system AC units not integrated with the BAS that we could not control (their nameplates were unreadable, precluding quantify their tonnages).

To validate the IPKeys model, Fraunhofer applied a cooling-degree-hour (CDH) approach to both whole-building and four RTUs where we had power telemetry (via SiteSage, with important limitations discussed later) to estimate cooling energy consumption as a function of T_{out} relative to the balance temperature, T_{bal} , the temperature above which the school

requires space cooling, i.e., $CDH = T_{out} - T_{bal}$. Our analysis focused on core school hours (7 AM to 3PM) in September 2018², with regressions for each hour since facility power draw varied with time of day (see Figure 4). We found that $T_{bal} = 60^\circ\text{F}$ yielded the best fit for whole-building and the four RTUs for which we had telemetry (see Kromer et al. 2020 for more detail).

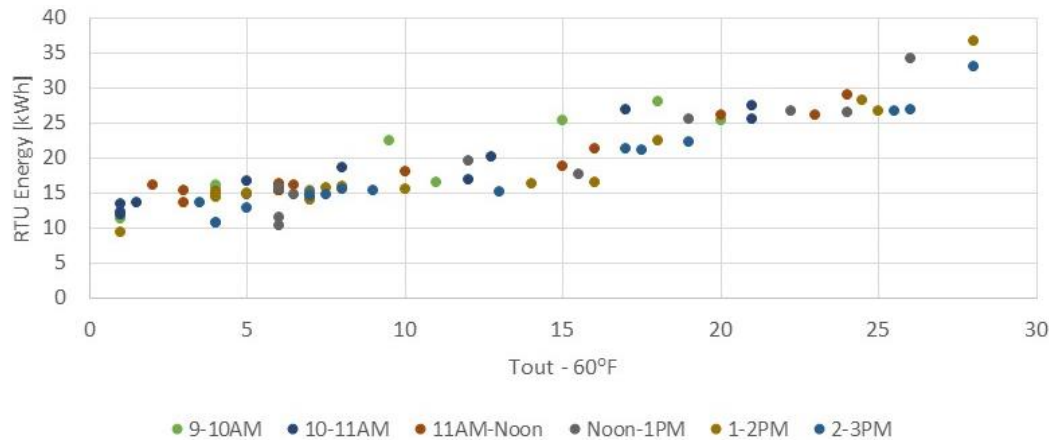


Figure 4: Hourly electricity consumption for four RTUs as a function of T_{out} and time of day (Sept. 2018).

The hourly regression coefficients ranged from 2.3 to 3.2 (highest in the morning, potentially due to nighttime recovery), with RMSE between 3 and 12%. On average, that whole-building hourly electricity consumption increases by about 2.5 kWh/ $^\circ\text{F}$ that T_{out} exceeds 60°F . The power draw of the four RTUs increased by about 1/3 of that value or $\sim 0.83\text{kWh}/^\circ\text{F}$; this reflects that the RTU power data exclude three RTUs and AC split systems that were not submetered (nor controlled). Based on these values, AC energy would reach approximately 75kWh/h at $T_{out} = 90^\circ\text{F}$, while the four RTUs would draw $\sim 25\text{kW}$ for AC ($\sim 34\text{kW}$ including ventilation fans). Indeed, the four RTUs peaked at 36.8kWh from 1-2PM on September 6 when $T_{out} = 88^\circ\text{F}$. In contrast, the whole-building power-CDH regressions for the fall of 2019, based on an appreciably smaller sample due to the large number of events called, had a much lower R^2 , an average hourly slope of ~ 1.0 , and achieved the best fit for $T_{bal} = 52^\circ\text{F}$. This suggests that whole-building AC power would peak around 38kWh at 90°F . In addition, we found that an increasing number of RTU AC systems were not operating after school resumed.

Field Testing Results

The SunDial System operated for 6,700 hours over 275 full days of testing; and operated in “application control” mode, which enables autonomous dispatch of DERs, for $\sim 5,900$ hours over 241 full days of testing. The system was offline for $\sim 4,275$ hours over 182 days, primarily due to: 1) failure of the PV site’s recloser; 2) site power outages (which often required manual reset of the site’s RTAC [real time automation controller] and/or switchgear); 3) planned site maintenance to support commissioning and customer acceptance that occurred in parallel with the SunDial demonstration; and 4) loss of communication between the GS and RTAC.

² We could not use data from 2019 due to numerous load-management events and the lack of RTU data for most of the cooling season. In addition, low school occupancy levels precluded comparing summer and September data.

Excluding equipment failures and planned site maintenance, system uptime was ~98%. Table 3 summarizes the use cases tested over the course of the SunDial pilot.

Table3: Summary of use cases tested during the SunDial pilot

Use Case	Days	Description
Peak minimization + Energy Cost Optimization	104	Minimizes peak load and uses excess energy to shift energy from low to high cost periods
Backfeed Minimization + Energy Cost Optimization	15	Minimizes reverse power flow
Virtual Peaker Plant	53	Minimizes and firms net load from 3pm to 7PM
Load Shape	32	System minimizes error relative to a target power output (e.g., ramp rate)
Power Firming	12	System matches actual net load to predicted net load as predicted at a prior point in time
Cost Optimization	10	Shifts energy from low to high cost periods

Total ESS throughput during pilot was about 354MWh, equivalent to 177 full charge / discharge cycles of the 1 MWh ESS. Approximately 70 total DR events, 32 AC and 35 EV charging load shift events, were called. We tested the functionality of the GS in stages, with facility load management enabled primarily toward the end of the pilot. A discussion of the system-level and facility-level results follows.

System-level Results

We present results to illustrate the SunDial system's performance for the Virtual Peaker Plant and Backfeed Minimization use cases. Please see Kromer et al. (2020) for additional results for these and the other use cases.

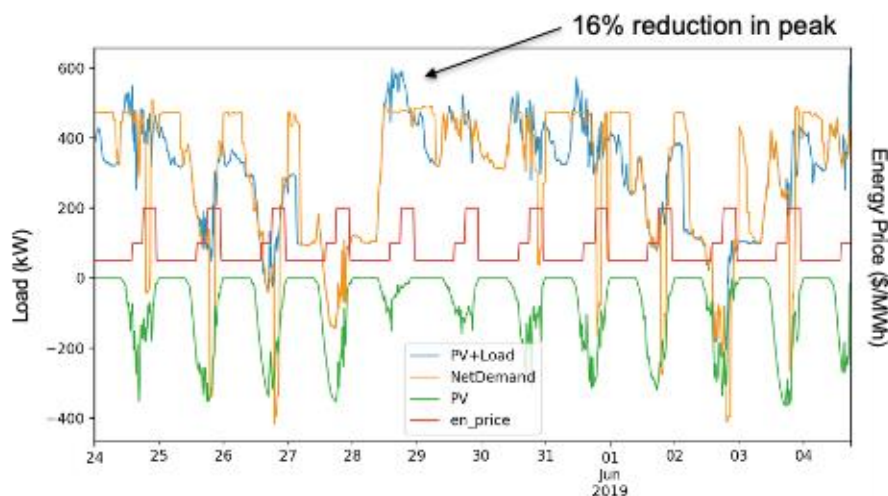


Figure 5: Peak shaving + energy cost optimization use case, May-June 2019 (13 days)

Peak minimization + Energy Cost Optimization: The primary objective in this use case is to minimize peak net demand over a defined time horizon, while also using excess storage capacity to optimize energy cost given a time-differentiated cost of energy. We varied the (1) peak demand minimization time horizon (daily vs monthly) and (2) the baseline threshold above which an additional kW of demand incurs a non-zero cost. We show results from a two-week

period in May-June 2019 in Figure 5; they are broadly illustrative of typical system behavior. Over this time period peak load is reduced by 16% relative to a non-storage baseline. On non-peak days, SunDial maximizes exports during a late afternoon increase in price.

For this use case, there is significant down-side risk in trying to regulate around an overly aggressive (low) peak, as smaller errors can result in premature use of the ESS and load shifting, resulting in a new peak. This occurred when the prediction algorithm *underestimated* net load, causing the system to regulate to a more aggressive (lower) peak load than feasible given the energy available for load shifting. The reverse scenario, i.e., the system over-estimates demand and deploys energy in an overly conservative way, also presents a challenge, as it increases the likelihood that generation will exceed demand. We mitigated these issues by improving the accuracy of the load and PV generation prediction algorithms (see Kromer and Roth 2020) and implementing additional control objectives to account for forecast uncertainty.

Results for this use case were evaluated for several alternate control scenarios:

- (1) “Heuristic”: Regulates net power output below a target threshold, but does not use best-available predictions to assess whether the system can likely meet this threshold.
- (2) “Perfect Information”: GS had perfect future (i.e., actual) load and PV predictions.
- (3) “Baseline”: GS uses prediction algorithms, prior to implementing prediction uncertainty.
- (4) “Uncertainty Mitigation”: GS uses prediction algorithms, with objectives tuned to mitigate prediction uncertainty.

As shown in Figure 6, modified “uncertainty mitigation” strategy outperforms the “Baseline” in every month, reducing monthly peak demand savings by an average of 92% and coming with 13% of the “perfect information” case. It also outperforms the “Heuristic” approach in 7 out of 9 months, achieving, on average, 41% greater average peak reduction.

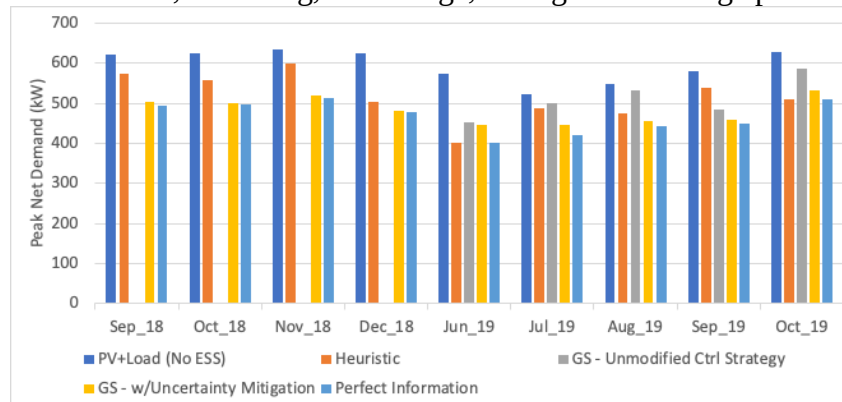


Figure 6: Monthly peak load as a function of ESS control strategy, relative to no-ESS case (“PV+Load”).

Virtual Peaker Plant (VPP): The VPP provides firm capacity during periods of peak stress on the power distribution system, seeking to present a predictable, constant, and minimized demand profile from 3PM to 7PM eastern daylight time (6PM to 10PM UTC). This use case is representative of tariffs established by the Massachusetts Clean Peak Standard. Its secondary objectives are to minimize VPP overall peak and to minimize energy imports (i.e., charge using local solar). Figure 7 shows a 10-day VPP implementation, during which the VPP realized an average peak reduction of 245 kW over the targeted four-hour window relative to a “no storage” case. Relative to an identically dimensioned ESS using a constant discharge strategy, the VPP implementation reduces peak by an average of 35kW, or ~140kWh of additional ESS capacity (i.e., 35kW x 4 hours). In addition, daily peaks are reduced by ~23 kW.

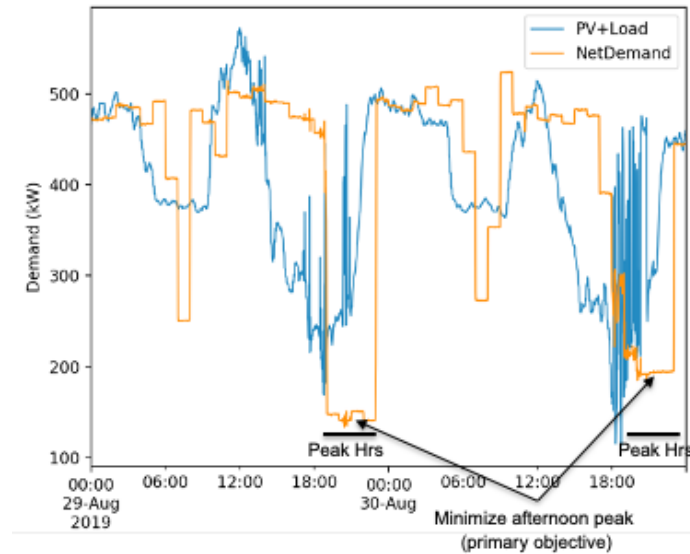


Figure 7: Virtual peaker plant example. Black lines indicate target periods for peak load reduction. All times UTC.

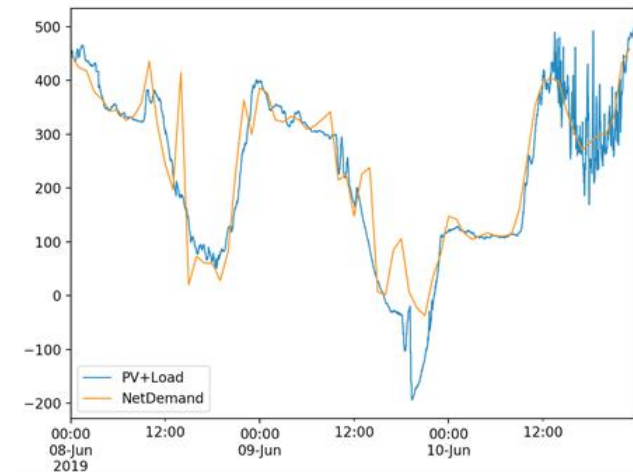


Figure 8: Power backfeed mitigation example

Table 4: Comparison of Virtual Peaker Plant afternoon peak load reduction scenarios

Scenario	Avg Afternoon Peak (kW)	Average Peak Load Reduction (kW)	
		Relative to No ESS	Relative to Constant Discharge
PV+Load Only (No ESS)	421	0	-207
Constant Discharge Strategy	214	207	0
GS, As-Implemented for field test	178	243	36
GS, Modified Control Objectives	166	255	48
Perfect Information	140	281	74

After field testing, we further refined the GS algorithms by implementing a non-linear cost to ESS discharge at low SOE to mitigate forecast uncertainty and improve solar and load prediction algorithms. Applied retrospectively (i.e., via modeling), these yielded, on average, a 48kW reduction in the afternoon peak, equivalent to 196kWh of additional ESS capacity relative to a

constant discharge case; and reducing overall peak by approximately 28kW. Results for four different control scenarios are summarized relative to a “No ESS” case in Table 4.

Backfeed Minimization: This use case simultaneously minimizes reverse power flow on potential solar surplus days while concurrently minimizing peak system load on high demand days (see Figure 8). PV generation exceeded demand for a portion of 24 days, with a total of 9,450kWh of energy backfed to the distribution system. With perfect information, the system would have been forced to shed 299 kWh and productively shift the balance of 9,155kWh that would otherwise be curtailed. The GS captures ~70% this total potential load shed.

Load-shifting Results

We found that the portfolio of flexible loads could repeatedly productively absorb ~50 kWh of solar generation during shoulder seasons and shed ~40kWh on a typical summer day to reduce peak load. These flexible load resources had significant constraints as to the time of day and duration of available load shifts. An additional large and highly flexible load (conveyor drag) would have increased total shiftable load to 96kWh during shoulder seasons and 85kWh during the summer season. However, it lacked automated control and we found no evidence that this resource performed as planned (the customer became non-responsive). This highlights the need to automate control of demand-side assets in use cases with frequent event calls; in addition, the lack of feedback about process status highlights the value of submetering key loads to accurately characterize performance of responsive loads.

We discuss the findings for the AC temperature and EV charging applications further.

AC Set-point Adjustment: Evaluating the load-shifting performance of the school AC systems proved challenging for two reasons. First, the magnitude of the load shifts on many days similar to the variability in whole-building power draw (5-10%). This precludes using facility-level data to compare actual and predicted load shift and the net energy impact (i.e., net kWh relative to baseline power) for individual events. Second, the school’s SiteSage submetering system included only four of the seven RTUs, and we lost data access for those circuits because the school converted their wireless networks from 2 to 5GHz. Although we quickly procured the necessary hardware to submeter all seven RTUs, the school only brought the RTUs on line in September (three) and October (the other four) of 2019. Consequently, we acquired a very limited set of cooling-season data. Nonetheless, we were able to demonstrate that the FLAME called and actuated events over a significant portion of the summer and early fall of 2019.

During the summertime, school whole-building power levels are appreciably lower. We analyzed summertime events with a range of pre-cool and shed durations (from 1 to 4 hours), using a baseline incorporating a same-day adjustment based on the average power in the hour prior to an event to address variability in other end uses. Although the hourly baseline power regressions have low goodness of fit values, the events yielded reasonably consistent results among eight precool events and six shed events that followed a precool. Figure 9 shows the sum of load shed as a function of total CHD (left, including multi-hour shed events) and precool energy as function of CDH (right). As expected, both scale ~linearly with CDH, as longer events and higher T_{out} increase cooling loads. We also found that the incremental first hour kWh does not vary greatly among events; this is expected since the *incremental AC load* should be similar among events.³ Together,

³ To a first order, the incremental AC loads are due to the (up to 3°F) greater temperature difference between T_{out} and T_{in} (increases outdoor air, infiltration, and shell loads) and heat exchange with the building thermal mass as T_{in} decreases by 3°F. The AC COP also decreases as $T_{in}-T_{out}$ increases, increasing AC consumption during events.

this suggests that, with sufficient baseline performance data and data from functional tests (i.e., events), meaningful assessments of and models for the load management potentials of commercial building AC system under a wide range of conditions can be developed.

We did not receive any complaints from the school operations manager.

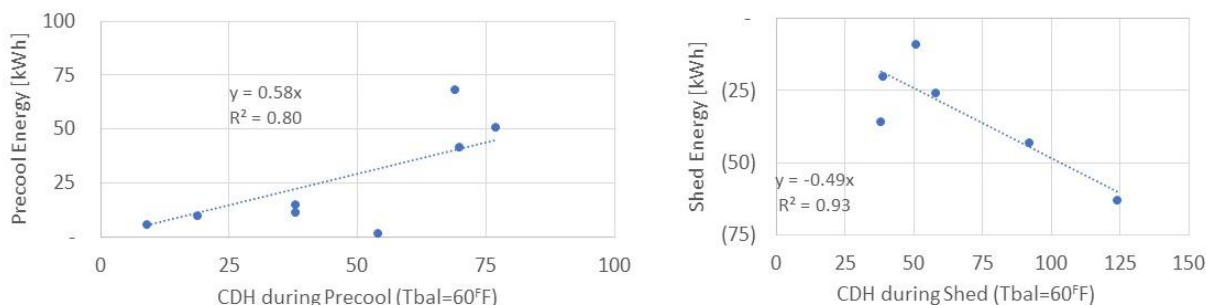


Figure 9: RTU total load shed (left) and precool energy (right) scale linearly with total event cooling-degree hours.

EV Charging: We called a total of 35 successful load-shifting events between August 6 and October 9, 2019 based on commands from the Global Scheduler. On average, metered data indicate that the average event shifted 54kWh, relative to the 67kWh predicted. Load shift predictions had a mean average error of 25kWh and a root-mean squared error of 30kWh.

The facility manager indicated that the EVs typically require 6-8h to charge, followed by a similar cool-down period prior to use to maximize battery lifetime. Since EV charging typically began in the evening, the EVs batteries would have plenty of time to complete charging prior to the resumption of their use the following afternoon. We called many events where we locked out charging until 10 or 11AM, i.e., they would finish charging close to when they would next be used. Nonetheless, we did not receive any complaints from the facility manager.

Lessons Learned and Recommendations

We share several lessons learned and recommendations from the pilot project that we hope can inform future RD&D work to help scale the integration of facility load management, energy storage, and solar PV. Notably, we hope that DOE-BTO's Connected Community projects to demonstrate viable solution packages and business models for grid-interactive efficient buildings (GEBs), can address many of these challenges.

Markets for providing services to the distribution grid remain nascent: While distribution-level VPPs can provide numerous functions to support operation of distribution networks, the markets for these services are not fully developed. Maturation of frameworks to enable local exchange of load-shaping services will be a critical enabler for operating renewables-dominant distribution grids.

C&I flexible loads can effectively supplement ESS for managing high penetration solar PV: For the cases studied, flexible loads comprising ~10% of the ESS capacity supported multi-hour load shaping objectives under a wide range of conditions without notably impacting customer operations. However, availability and predictability constraints made it challenging to treat them as firm capacity. These issues can be partially mitigated by load submetering, increasing the size of the load management portfolio, and improving HVAC models.

Energy storage can simultaneously manage short-term PV intermittency while supporting load shifting, but duty cycles need to account for real-world ESS performance:

The GS dispatched the ESS to simultaneously mitigate short-term intermittency changes and achieve daily performance objectives. Much of the ESS charge / discharge throughput used to smooth changes in PV production occurred at <10% of nameplate power, with measured real-world roundtrip efficiency of ~40 to 50% (vs. ~82-85% at higher charge / discharge levels).

Accurate, reliable, and timely telemetry is critical to effectively leveraging distributed resources: This is critical for developing accurate predictions, particularly to rapidly detect major deviations from expected generation or loads and modifying operating decisions, and for detecting problems as they arise (ongoing commissioning). Submetering major loads could greatly improve the quantification and prediction of their flexibility.

Incorporating auto-calibrating, self-learning techniques into solar and load predictions is a critical enabler for scalability of integrated approaches: As discussed in Kromer and Roth (2020), the accuracy of solar generation and load predictions were significantly improved by feeding back near-real-time and historical data into machine-learning algorithms.

Cost-effectively integrating flexible loads to support high-penetration solar will require a highly streamlined process: The pilot took the CSP weeks of analysis to develop accurate prediction values. Importantly, CSPs also have a limited amount of customer attention for implementation, so excessive demands on customer time (e.g., more than a few hours total), will likely result in nonresponsive customers or customer attrition. Ideally, onboarding sites would take less than half a day, and site implementation less than one day.

Automated Load Management is Essential for Use Cases with Numerous Events: We called events on a majority of summer and fall weekdays and could detect a response at the two facilities with automated actuation, but not at the facility that declined automated load control. Most facilities will not manually implement events on many/most workdays.

Predicting AC Load Management Potentials for a Wide Range of Conditions and Durations is Challenging: AC loads and load shifting potential over different time horizons and conditions depend on several difficult and/or time-intensive factor to model. Developing these models is also currently not a core competency of most CSPs, who focus on predicting peak load sheds. We recommend, developing standard (ideally, automated) functional test and modeling and analysis procedures to acquire and analyze building data to generate accurate-enough models to predict AC load flexibility with minimal time investment. AC submetering also can provide valuable inputs. The facility-specific model should be automatically updated as additional events occur; this will require integration of the model platform with both power (whole building, ideally submetered) and connected AC (or potentially the BAS to feed model-predictive controls).

Controls Integration is Often Challenging and Not Reliable: We encountered multiple problems with BAS programming, including incorrect measurement point references and disabling of data acquisition functionality by the school's controls contractor.

OpenADR can be Expanded to Communicate Information Needed for a Wider Range of Use Cases: IPKeys developed and implemented a new communications protocol as an expansion of OpenADR, specifies communications to support the delivery load-shaping services.

High-Fidelity AMI Data are Essential to Model and Quantify Load Impacts: High-resolution (e.g., 1-5 minute) can help to quantify the load impact of process loads. Devices and meters that make high resolution data readily available on site to customers would help greatly.

Effective ESS use for strategic dispatch depends significantly on prediction accuracy and uncertainty mitigation: The biggest improvements in GS performance came from refining prediction methodologies and incorporating methods to account for prediction uncertainty by building in additional storage reserve capacity. As described in Kromer and Roth (2020) and

Kromer et al. (2020), the peak load management use cases, incorporating a predictive, adaptive controls approach had benefits equivalent to 20% to 40% of the total ESS capacity relative to a rules-based approach. We found that leaving ~30% of ESS capacity in reserve offered an effective hedge given the uncertainty; this approach could be refined by dynamically adjusting the reserve margin based on the current system state and the specific control objectives.

Acknowledgements

This work was funded in part by the U.S. Department of Energy Solar Energy Technologies Office (DOE-SETO) under Contract # DE-EE0007164 and the Massachusetts Clean Energy Center (MassCEC). The efforts of many people made this project possible:

DOE-SETO: Kemal Celik, Tassos Golnas, Guohui Yuan.

Facility Managers: John Blood, Kerry Henry, and Jon Snyder

Fraunhofer USA: Christian Hoepfner, Omri Schwarz.

IPKeys: Yuriy Kozhynov, John Kozyrakis, Robert Nawy, Victor Sadikov, and Bill Walker.

Massachusetts Clean Energy Center (MassCEC): Kavita Ravi.

National Grid: Justin Woodard, Aram Shishmanian, Samer Arafa, Jason Small, Fouad Dagher, Nick Corsetti, Peter Russo, and Paul Wassink.

References

- Kromer, M., K. Roth, M. Zeifman, T. Yip, J. Boch, S. Arafa, A. Shishmanian, and J. Woodard. 2020. “SunDial – An Integrated SHINES System to Enable High-penetration Feeder-level PV”. Final Report to U.S. Department of Energy Solar Energy Technology Office. January.
- Kromer, M. and K. Roth. 2020. “Seeing the Future: Evaluation of Time-Series Load-Prediction Methodologies for Optimal Energy Storage Dispatch in Solar-Dominant Distribution Grids.” *Proc. ACEEE Summer Study on Energy Efficiency in Buildings*. August.
- Kromer, M, K. Roth, and T. Yip. 2020. “Optimizing DER Dispatch in a Renewables Dominant Distribution Network Using a Virtual Power Plant”. *Proc. 2020 IEEE Power & Energy Society General Meeting (PESGM)*. Montreal, Canada. August.
- Merchant, E.F. 2020. “Solar Will Be Cheapest Form of Electricity Across US by 2030, Plus Other PV Trends.” Wood Mackenzie / Greentech Media. February 2.
- Roth, K., M. Kromer, and J. Boch. 2020. “Facilities Load Aggregation & Management Engine (FLAME)Development Report.” Report by Fraunhofer USA to DOE-SETO. January.
- Zeifman, M, M. Kromer, and K. Roth. 2017. “Integrated system to enable high-penetration feeder-level PV: Preliminary design and simulation results,” IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT). April.