

Mechanisms and Policies for Controlling Distributed Solar Capacity

NOMAN BASHIR, DAVID IRWIN, PRASHANT SHENOY, and JAY TANEJA,
University of Massachusetts Amherst

The rapid expansion of intermittent grid-tied solar capacity is making the job of balancing electricity's real-time supply and demand increasingly challenging. Recent work proposes mechanisms for actively controlling solar power in the grid at individual sites by enabling software to cap it as a fraction of its time-varying maximum output. However, while enforcing an equal fraction of each solar site's time-varying maximum output results in "fair" short-term contributions of solar power across all sites, it does not result in "fair" long-term contributions of solar energy. Enforcing fair long-term energy access is important when controlling distributed solar capacity, since limits on solar output impact the compensation users receive for net metering and the battery capacity required to store excess solar energy. This discrepancy arises from fundamental differences in enforcing "fair" access to the grid to contribute solar energy, compared to analogous fair sharing in networks and processors. To address the problem, we first present both a centralized and distributed algorithm to enable control of distributed solar capacity that enforces fair grid energy access. We then present multiple policies that show how utilities can leverage this new distributed rate-limiting mechanism to reduce variations in grid demand from intermittent solar generation.

CCS Concepts: • **Hardware** → **Power and energy**; **Energy distribution**; **Smart grid**;

Additional Key Words and Phrases: Solar, fairness, net meter

ACM Reference format:

Noman Bashir, David Irwin, Prashant Shenoy, and Jay Taneja. 2018. Mechanisms and Policies for Controlling Distributed Solar Capacity. *ACM Trans. Sen. Netw.* 14, 3-4, Article 25 (December 2018), 28 pages.
<https://doi.org/10.1145/3219811>

1 INTRODUCTION

The amount of grid-tied solar power continues to grow at an exponential rate, with capacity increasing by an average of 33% each year over the past 6 years (Hill 2017).¹ This growth is driven by consistent drops in solar module prices, which have fallen 10% per year on average over the past three decades, due to both advances in solar module design and increasing economies of scale in manufacturing. For example, Swanson's Law—the solar equivalent of Moore's Law—observes

¹This article is an extended version of a paper previously published at ACM BuildSys (Bashir et al. 2017).

This research is supported by NSF grants CNS-1645952, IIP-1534080, CNS-1405826, CNS-1253063, and CNS-1505422 and the Massachusetts Department of Energy Resources.

Authors' addresses: N. Bashir, D. Irwin, P. Shenoy, and J. Taneja, University of Massachusetts Amherst; emails: nbashir@umass.edu, irwin@ecs.umass.edu, shenoy@cs.umass.edu, jtaneja@umass.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

© 2018 Association for Computing Machinery.

1550-4859/2018/12-ART25 \$15.00

<https://doi.org/10.1145/3219811>

that the price of solar modules tends to drop 20% for every doubling in the cumulative shipped volume (Swanson 2006). In many locations, the average cost of solar energy is now less than from fossil fuels. Some estimates project that solar could contribute as much as 20% of global electricity consumption as early as 2030 (Farmer and Lafond 2016).

Unfortunately, the increasing penetration of solar energy in the grid complicates utility operations. In particular, utilities are responsible for balancing electricity's real-time supply and demand by regulating the power generation of a portfolio of "dispatchable" generators. Historically, grid demand, when aggregated over a large number of customers, has been smooth and highly predictable based on the expected temperature and the day, e.g., weekend, weekday, or holiday. As a result, utilities have been able to effectively plan when and what generators to dispatch to satisfy expected demand. However, increasing solar penetration now requires utilities to also compensate for variations in solar output over multiple timescales. Solar output, even when aggregated, is much less predictable than grid demand, since it varies primarily based on cloud cover, which is more localized and stochastic than temperature variations. At short timescales, compensating for large solar variations due to clouds using mechanical generators is challenging, since generator ramp rates are less than solar ramp rates. Further, at longer timescales, utilities lose revenue from users generating their own solar power during the day but must still maintain the generating capacity to provide these users electricity when the sun is not shining, e.g., during cloudy weather, at night, and over the winter. This has serious implications to utilities' business model.

As a result, government regulations generally place *limits* on the amount of grid-tied solar capacity that can be installed and feed energy into the grid. These limits are currently set based on a complex political process that includes multiple stakeholders with competing interests, including politicians, utilities, environmental groups, and solar installers. In the United States, these limits vary widely by state and often restrict both the percentage of users with grid-tied solar and their aggregate solar power capacity. The rapid growth in solar power is now causing states to frequently hit these limits, triggering protracted negotiations (often taking many months) among the stakeholders to raise them. Since the limits, which are a form of admission control, are hard, once they are hit, additional users cannot install grid-tied solar until they are raised. For example, due to such limits, users in Hawaii were recently barred from installing grid-tied solar for 2 years (Cardwell 2015; Mulkern 2013). Similarly, Massachusetts reached its cap in summer 2015: it then took 9 months for the state legislature to negotiate and pass an increase in the cap and for the governor to sign it Massachusetts (2016). The solar cap in Massachusetts was then reached again in October 2017, halting \$78M in solar projects (Analysis 2017), spurring another round of protracted negotiations.

Importantly, the aggregate power limits or caps above are static and based on the *rated installed capacity* of each solar installation, not the amount of power they actually generate in real time. Standard Test Conditions (STCs) for rating solar module capacity specifies an irradiance of 1kW/m^2 with an air mass of 1.5, no wind speed, and a cell temperature of 25C. These conditions approximate the generation of a south-facing solar module (tilted at the same angle as the sun) at solar noon near the equinox on a clear sunny day in the United States with an ambient air temperature of °C.² Of course, weather conditions are rarely this "ideal": the ambient air temperature at STC is unrealistic, roof lines frequently dictate nonideal orientations and tilts, and solar irradiance is usually much less than 1kW/m^2 , e.g., during the morning, evening, over much of winter, and under cloudy skies. Given that STC conditions are unrealistic and essentially never naturally occur in the United States, alternative test conditions, such as PVUSA Test Conditions (PTCs), have been proposed (Durrenberger 2015).

²STC actually specifies that the ambient and solar cell temperature are 25C, which is physically impossible, as the solar irradiance increases the cell temperature much higher (+25C) than the ambient temperature.

Thus, the *actual* aggregate solar power generated is rarely, if ever, at (or even near) the rated capacity, and varies widely each day, over the year, and as the weather changes. For example, on cloudy days, the aggregate contribution of solar power across many distributed sites is much less than on sunny days, but also more variable. As a result, on a cloudy day, the grid could potentially accept solar power from many sites that are currently forced off-grid without exceeding its capacity limit, although the high variance in output may pose operational challenges in balancing supply and demand. To address the problem, recent work proposes mechanisms (Singh et al. 2017) and policies (Lee et al. 2017; Rongali et al. 2016) for actively controlling solar power output to the grid. This work enables software to cap the solar power injected to the grid as a configurable fraction of its time-varying maximum output (Singh et al. 2017) and then, inspired by similar rate control problems in networking, designs rate allocation policies to limit the aggregate contribution of distributed solar subject to the solar capacity the grid is willing to accept (Lee et al. 2017; Rongali et al. 2016).

An important metric when determining how to dynamically limit each solar site's power output is preserving *fairness* between sites. For example, one site should not have their entire solar output curtailed, while another site has none of its solar output curtailed. Prior work co-opts the traditional notion of "fairness" from the networking literature, which computes it with respect to the instantaneous sending rates of flows, and not the cumulative amount of network traffic flows sent over time. This makes sense in networking, as senders can potentially generate an arbitrary amount of network traffic at any time. Thus, if one idle sender does not generate traffic for a long period, then (1) other senders should be able to increase their sending rate to consume any resulting excess network bandwidth during this time, and (2) the idle sender should not be able to accumulate unlimited credit for their idleness, enabling them to monopolize the link once they resume sending. The former property ensures bandwidth allocations are work conserving, while the latter property prevents starvation of senders. Analogously, prior work attempts to maintain "fair" solar rate allocations, such that each solar site contributes near the same fraction of their time-varying maximum instantaneous solar power output. For example, the work always ensures that all sites contribute $X\%$ of their current maximum solar output, for some value of X .

The problem is that this traditional notion of fairness in networks does not map well to solar energy and the grid. Instead, we argue that the grid should express fairness in terms of the total fraction of *energy* users contribute over time (with respect to each other) rather than in terms of their instantaneous rates of *power*. Ultimately, users care about the amount of total solar energy they can feed into the grid (over some time window), as a fraction of the total solar energy they could possibly feed in, since this impacts both the cost of their system and the revenue it generates. In particular, users directly receive compensation for the energy they feed in, which decreases with the fraction of energy they can contribute. This compensation is used to offset the initial capital costs of installing solar and affects the time it takes to pay back the installation of the system, which is a key metric in making the decision as to whether to install solar. The expected fraction of energy users *cannot* feed into the grid may also necessitate additional system costs to store excess energy.

As we show, enforcing fair instantaneous rates, as in networking, may result in *unfair* contributions of total energy over time. Unlike in networking, solar sites can only generate "traffic," i.e., solar energy, at certain times based on the sun's irradiance, which is a function of location, time, local weather, and physical installation characteristics. Importantly, solar sites *cannot control their location, the sun, the weather, and often their physical characteristics*, and thus have little to no control over when and how much solar power they can generate. In contrast, network clients that are not generating traffic are doing so voluntarily, and *could generate traffic if desired*. However, network clients do not directly receive compensation for sending data. Clearly, if network

clients directly received compensation for the total amount of data they sent, they most certainly would generate traffic all the time, and the total amount of data they sent over time would be critically important. Traditionally, network providers place data caps on users, such that if they send data beyond a threshold within a billing period, e.g., a month, they must pay additional fees or receive degraded (or no) service. Again, this method does not apply to regulating solar energy, as all sites are generating energy, i.e., “traffic,” at similar times based on the same source—the sun. The generating capacity of solar sites also varies widely and is largely outside of users’ control.

This article identifies this fundamental difference between fair rate allocation in networks and fair grid energy access for solar, and discusses how and why it arises. We then design a suite of rate allocation algorithms to enforce weighted fair grid energy access and evaluate its tradeoffs. Collectively, these rate allocation algorithms represent a new *mechanism* for controlling the output of distributed solar capacity. Finally, we present multiple *policies* that show how utilities can leverage this new distributed rate-limiting mechanism to reduce both the magnitude and variations in grid demand from intermittent solar generation. In doing so, we make the following contributions.

Solar Fairness Definition. While preserving fairness is a first-class concern when sharing processors and networks in computer systems, it has generally not been a metric of interest in sharing the electric grid. We introduce and define the notion of distributed solar fairness (DSF) and discuss how it differs from similar notions of fairness in computer systems and networking. We also discuss how unfairness arises among distributed solar sites with limits on their aggregate solar output.

Mechanism for Controlling Distributed Solar. We propose a simple rate allocation algorithm to enforce fair grid energy access among distributed solar sites, which defines a new mechanism for fairly controlling distributed solar capacity. While this algorithm allocates rates to different solar “flows” over time, as with traditional fair-sharing algorithms in computer systems and networks, it varies these rates to ensure users contribute the same fraction of their actual solar energy capacity. The algorithm exposes tradeoffs in its convergence speed, fidelity to the aggregate limit it enforces, and robustness, i.e., the interval over which it must exchange data.

Policies for Controlling Distributed Solar. The mechanism above simply enables utilities to set fixed caps on the amount of aggregate power solar sites contribute to the grid. However, since aggregate solar output changes over time, utilities must change these caps over time to effectively control solar output. Thus, we define two distinct policies that show how utilities can leverage this new distributed rate-limiting mechanism to reduce both the magnitude and variations from intermittent solar generation in the grid. These policies vary the solar generation based on its fraction of grid demand and its fraction of the average aggregate solar generation. We discuss the general fairness properties of regulating solar capacity based on these different policies.

Implementation and Evaluation. We implement our mechanisms and policies above and evaluate them on both synthetic data and real data from 50 solar sites. We show that traditional equal-rate allocation results in solar sites contributing up to 18.9% less energy over a single month than our mechanism that enforces fair grid energy access. Finally, we show that our policies that vary solar capacity over time reduce the variations in aggregate solar power compared to a fixed limit, resulting in a more stable grid demand, while also preserving fairness.

2 SOLAR FAIRNESS IN THE ELECTRIC GRID

In this article, we consider grid-tied solar arrays with “net metering” capabilities, which enable solar arrays to synchronize with and feed their energy into the electric grid. The current grid allows a net-metered grid-tied solar array to feed any amount of power into the grid, up to its maximum

installed capacity, with no restrictions. Thus, the “admission control” decision of whether to allow a solar array to net meter at all must be made at installation time. Once a solar array is installed and tied to the grid, there are no restrictions on the amount of power it can net meter. As discussed earlier, this severely limits the number of solar installations the grid can permit, since admission control policies must plan for the worst-case scenario, i.e., where all solar arrays concurrently feed in their maximum capacity, even though this scenario is highly unlikely (if not impossible), and by definition can only occur one time per year, i.e., at solar noon on the summer solstice under clear skies at an ambient temperature of °C. This type of static admission control is both imprecise and inefficient. By comparison, if admission control in the Internet worked this way, ISPs would decide whether users could obtain an Internet connection based on whether they could accommodate all users concurrently operating at maximum bandwidth, i.e., all simultaneously watching a high-definition movie. Such an admission control policy would result in only a small fraction of users being able to access the Internet. Of course, the Internet relies on statistical multiplexing, under the assumption that users are not frequently, if ever, using the network at the same time, and, if they do, it relies on network protocols, such as TCP, to fairly share the available bandwidth and prevent congestion collapse.

Thus, enforcing such limits at “runtime,” rather than at install time, has the potential to enable a much larger number of grid-tied solar arrays, while still limiting the total net-metered power to a prespecified capacity. In the future, we expect the grid to have the capability to rate control the amount of power that can be injected by a grid-tied solar array at any instant. These rate control capabilities are increasingly being included in so-called smart solar inverters. While smart inverters are currently being tested in small-scale demonstration projects, we expect them to gain broader adoption as solar penetration increases and the technology becomes more proven (Kroposki 2016). Since the allowed rate may vary over time, each solar array will need to control the setting of its smart inverter to enforce the assigned rate. The ability to rate control solar arrays at the timescale of minutes or hours has many benefits. For example, it can simplify the creation of generator dispatch schedules in the presence of high renewable penetration, since it places an upper bound on solar generation. It can also allow the installation of a much larger number of solar arrays while limiting their stochasticity. Finally, it can incentivize the use of local energy storage to store any surplus solar power that cannot be net metered into the grid due to capacity limitations.

Given such a scenario, we examine the problem of how the grid should assign rates to different solar arrays while maintaining both an aggregate limit on solar output and fairness across users. Prior work has used an analogy to the rate allocation problem in computer networks and applied the notion of fairness from networking to address this problem. Specifically, prior work uses analytical models of TCP’s rate control algorithm, which achieve network fairness, and weighted versions of this rate allocation problem to model the problem (Lee et al. 2017; Ardakanian et al. 2013; Rongali et al. 2016). However, with solar, owners directly receive compensation for the solar energy they contribute and thus are incentivized to always produce as much power as possible.

2.1 Defining Solar Fairness

As a result, rather than using a notion of fairness from networking, we instead propose a new fairness metric for rate-controlled solar arrays called *distributed solar fairness (DSF)* that is based on the cumulative net-metered energy. Let $E_i^{actual}(t_2 - t_1)$ denote the actual energy net metered by a solar array i over a duration $[t_1, t_2]$ in the presence of rate control, and $E_i^{max}(t_2 - t_1)$ denote the maximum amount of energy it *could have produced* in this time period with no rate control, e.g., using standard techniques such as maximum power point tracking (MPPT). Note that a site’s maximum generation potential varies over time based on a site’s unique location, weather, and physical

characteristics. Since rate control reduces the total energy that can be produced, the reduction in net-metered revenues over the interval $[t_1, t_2]$, which we term as $loss_i(t_2 - t_1)$, is $1 - \frac{E_i^{actual}(t_2 - t_1)}{E_i^{max}(t_2 - t_1)}$. This can be viewed as a direct monetary loss incurred by solar array i over the specified time interval due to rate control.

To be fair across users, we require that the percentage loss in energy, and thus compensation, is the same for all arrays over any time interval $[t_1, t_2]$. Thus, our notion of fairness requires that for any two arrays i and j ,

$$|loss_i(t_2 - t_1) - loss_j(t_2 - t_1)| < \epsilon. \quad (1)$$

While our ideal definition of fairness requires that this condition be true over any arbitrary time interval, in practice, achieving fairness over very short timescales may be infeasible. For example, if the sun has risen at the location of array i but it has yet to rise at the location of array j , it is not possible to guarantee fairness over a small timescale, since array j is unable to produce any power. In the next section, we describe a number of factors that complicate enforcing fairness at short timescales. However, it is both acceptable and feasible to enforce fairness over the much longer timescale of hours or days, or even at the timescale of a monthly billing cycle. In general, consumers' primary concern is whether their monetary percentage loss based on the energy they feed in is fairly distributed across all arrays over these longer timescales. Thus, in practice, the grid only needs to ensure fairness over these longer intervals $[t_1, t_2]$.

In the case of networks, fairness guarantees are provided only when the network flows are *backlogged*, which requires that the flows can continuously send data when network capacity is available. In our case, providing fairness over very short timescales also requires that the solar arrays be capable of producing enough power to use their allocated rates. However, over longer timescales, it is possible for an array to not use its instantaneous allocation, since it is unable to produce sufficient power, and yet "catch" up later by injecting power at higher rates than other arrays.

Even when enforcing fairness over longer timescales, the problem of allocating rates to each array is complicated by many factors. For instance, a simple approach that allocates identical rates to two arrays of identical size can yield unfair results. This occurs because arrays of identical size can still produce vastly different power output at any instant due to local differences in weather, as well as differences in physical installation factors, such as tilt, orientation, and location. Ignoring these differences can cause the fairness measure to diverge for various arrays. Thus, a fair rate allocation algorithm must consider several factors: assuming identical weather conditions, two arrays at two different locations will have slightly different sunrise, sunset, and solar noon times, yielding solar output curves that are time-shifted with respect to one another. In the networking case, this is analogous to enforcing fairness for *time-shifted* flows, where two identical flows are time-shifted and start transmitting data with different start times. Similarly, two solar arrays that are in proximity to one another may also produce different output due to microclimates, different shading effects, and so forth. Finally, different arrays may have vastly different capacities and thus rates must be computed to equalize the percentage loss for such heterogeneous size arrays.

2.2 Causes of Solar Unfairness: Solar Shape Diversity

Having discussed the notion of solar fairness, we now examine how unfairness arises from the differences in the shape of solar output across multiple sites. We specifically discuss how the different types of effects cause the "shape" of a solar curve to differ even across sites that are near to each other.

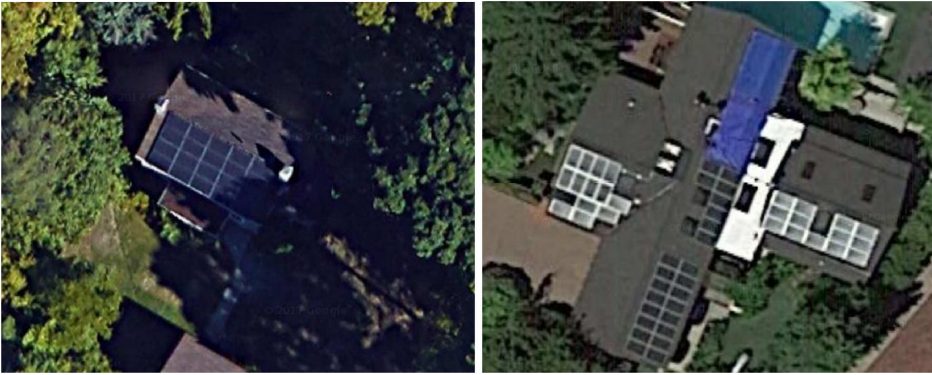


Fig. 1. Illustrative examples of nonideal solar sites.

Unfairness in solar energy access to the grid derives from the difference in output between solar sites, even when they are near each other. There are many reasons solar output between solar sites can differ. We outline some reasons below.

Solar Potential. The sun's position in the sky is unique at each location on earth at each instant of time. The sun's position in the sky, in turn, affects the air mass light must travel through to reach the earth, which reduces the amount of irradiance that reaches the ground. The solar potential is also a function of elevation, such that higher elevations have more potential than lower elevations at the same location. As a result, even with clear skies, the maximum solar-generating potential is different at every solar site at any moment. It is even possible for one site to generate solar power at the same time that another site is physically unable to generate any power.

Weather Effects. The weather also affects solar generation potential. In particular, solar power correlates with cloud cover, which is much more stochastic and localized than other weather metrics, such as temperature. For example, a cloud can cover one solar site, while a neighboring solar site is uncovered. As scattered clouds pass by, they can repeatedly cover and uncover solar sites at different times. In addition, solar cell temperature also affects generation and is a complex function of solar irradiance and ambient air temperature. Thus, microclimates where temperature and cloud cover vary significantly over small distances, such as those near large bodies water, can cause weather, and thus solar potential, to significantly differ at two nearby locations.

Solar Degradation. While solar modules are passive devices with no moving parts, they do degrade over time and experience faults, which affects the efficiency at which they convert solar irradiance into solar energy. For example, repeated exposure to extreme temperature changes can cause modules to discolor, increasing the cell temperatures, or even crack, causing small amounts of moisture to seep into the cells. The increased cell temperatures can degrade the materials that separate the P-N junction that electrons move through. Solar modules typically have a 25-year estimated lifetime with an expected degradation of 1% to 2% in efficiency per year. Thus, the output of older solar modules may differ significantly from the output of newer ones.

Physical Characteristics. Finally, the physical characteristics of a solar site also affect its solar output. These include the solar module's tilt and orientation, as well as any occlusions from surrounding buildings, trees, or mountains that may shade them. For example, an east-facing solar module will both start and stop generating power well before a west-facing one in the morning and evening, respectively. In general, rooftop solar deployments are complex and not ideal. Figure 1 illustrates typical rooftop solar deployments with multiple modules at different nonideal

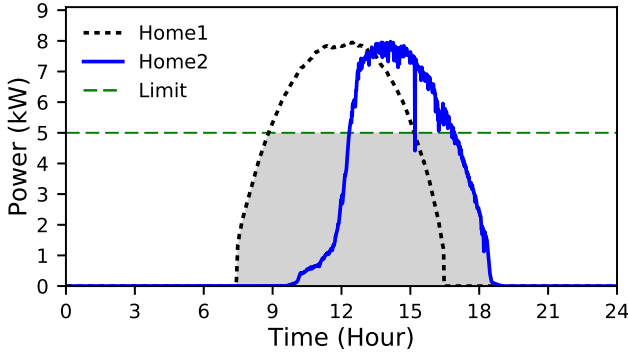


Fig. 2. Profile of solar output for two homes 80km apart.

tilts and orientations with significant shading from trees and other surroundings. In addition, soiling from debris, e.g., dust, mud, snow, leaves, and so forth, can also cause solar generation to differ between two nearby sites with identical solar modules. The differences above manifest themselves as differences in the shape of solar output at each site. We characterize these differences below, which are the root of unfairness in solar allocation.

Shifts. Shifts occur when a solar curve is shifted with respect to another solar curve, such that the first curve starts before or ends after another curve. Shifts occur due to either differences in the orientation of modules or differences in location. For example, east- and west-facing modules at the same location will be shifted with respect to each other. Similarly, a difference in longitude between two locations also results in a shift, since the sun rises and sets at different times (for the same day length).

Squeeze. Squeezes occur when a solar curve is narrower with respect to another solar curve, such that the first curve starts before and ends after another curve. Squeezes occur either due to differences in the tilt of modules or differences in location. For example, a south-facing vertically tilted module will be squeezed with respect to a horizontally flat tilted one. Similarly, a difference in latitude between two locations also results in a squeeze, since the length of a day changes with latitude.

Dips and Cuts. Dips occur when the solar output drops below the power level seen when the sky is clear. Dips may be caused by clouds, shade from trees, or nearby buildings and reduce the amount of sunlight seen by an array. The amount of the power dip depends on the magnitude of the reduction in the sunlight incident on the array. Similarly, cuts occur when a solar curve's power is cut off (or blocked) with respect to another solar curve, such that the first curve generates power normally while the second curve generates nothing. Cuts typically occur in the morning and evening, since these blockages are more prevalent when the sun is low in the sky. A cut is a special case of a dip where the output drops to zero.

Each solar site can exhibit an arbitrary combination of the three characteristics above. These characteristics are also static, since they are purely a function of a site's location, physical characteristics, and surroundings. As a result, if a solar site experiences a shift, squeeze, or dip relative to another solar site one day, it will often experience it every day, although the extent of it may change over the year, e.g., due to the sun's position or changes in foliage. In addition, different weather conditions between sites also create differences in the solar curves. Figure 2 illustrates how two nearby homes can exhibit different solar output over a day. In this case, Home 2 is more east-facing, as in Figure 1 (right), than Home 1, and thus its power generation is shifted with respect to Home 1 on this day. However, Home 2 has a cut near the end of the day, indicating a

blockage in solar output that causes its output to drop to zero, as in Figure 1 (left), which has trees on its west side that block sunlight near the end of the day. In this case, imposing a limit on the aggregate power from the two homes, and then satisfying this limit by allocating equal rates of solar power output between the two homes, results in an unequal solar energy contribution at the end of the day.

This occurs because at the beginning of the day, Home 2 is generating no power, and thus Home 1 is able to contribute a high fraction of its generation up to the limit. Due to the cut in power, once Home 2 starts generating power, it must share the grid with Home 1 by contributing an equal fraction of its time-varying maximum power potential up to the limit, even though Home 1 has already contributed a significant amount of energy to the grid. Thus, even though Home 2 contributes the same fraction of power as Home 1 at all times, its *fraction of energy* always remains less than Home 1, since it is never able to catch up from its lack of output at the beginning of the day.

Next, we present our mechanism for allocating solar rates that limits their aggregate output in real time and achieves solar fairness, while accounting for these factors.

3 MECHANISM: FAIR SOLAR ENERGY ALLOCATION ALGORITHMS

We assume a mechanism exists to remotely control the time-varying fraction of maximum power an individual solar site contributes to the grid, as described in recent work (Singh et al. 2017). We expect such a mechanism to be included in future smart inverters, which are already remotely accessible via the Internet (Kroposki 2016). We also assume that a grid-balancing authority exists and sets limits on the aggregate solar energy output across all solar sites by controlling this mechanism at each individual site. This control is similar, in effect, to current demand response programs that enable utilities to remotely control HVAC systems and thermostats to regulate grid demand. The primary difference is that solar modules provide more precise, fine-grained control that is transparent to users. We assume that the grid's transmission and distribution infrastructure, e.g., its transformers and feeders, are well provisioned to handle the maximum solar generation, such that the transformers never exceed their capacity and feeders do not reverse their power flow. These assumptions are likely true for the foreseeable future, as transformers and feeders are generally overprovisioned for energy consumption, and grid-tied solar power actually reduces the energy consumption. As a result, we need not consider the impact of the grid's network topology or the capacity of its distribution infrastructure in determining solar rate allocations.

Instead, the grid-balancing authority sets aggregate limits on the distributed solar output based solely on net metering regulations and their operational constraints, e.g., based on the characteristics of their generators. However, we assume the net metering regulations and operational constraints are dynamic and based on actual solar generation rather than static and based on the rated capacity of solar sites, as is the case today. That is, the limits are defined based on the actual aggregate solar power that may feed into the grid at any time, rather than on the number of installations that may connect to it. The balancing authority may also alter the limit to improve operations, such as increasing it during times of peak demand to allow more solar energy to flow into the grid. In this case, the curtailed solar power operates like high-quality reserve capacity or a demand response resource that a utility can use to balance the grid.

Our problem is to allocate the fraction of maximum power contributed by each site such that all sites contribute the same fraction of energy over each time window T . In general, we assume T is a long period, such as a week or a month, since, as discussed earlier, it may be difficult or infeasible to ensure fairness over shorter time periods. The analogous rate allocation problem in networking, if we assume the grid's transformers and feeders are well provisioned, is to simply enable all sites to contribute the same fraction (or rate) of their time-varying maximum power at

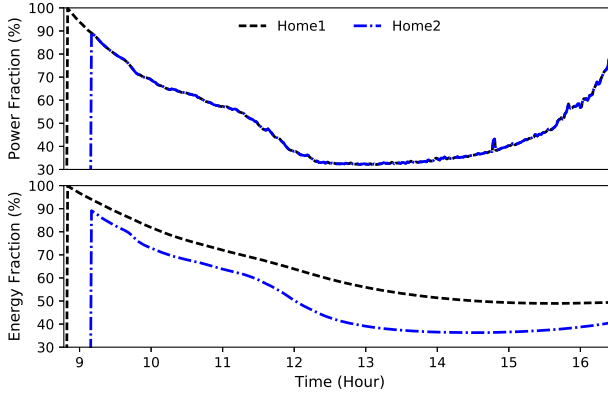


Fig. 3. Divergence in the fraction of energy contributed by Homes 1 and 2 from Figure 2, even when the fraction of power they contribute is equal, assuming a 5kW limit.

all times. Thus, to enforce an aggregate limit, the balancing authority might enforce that all sites contribute only 50% of their maximum power. Note that we assume the grid-balancing authority specifies the aggregate limit in terms of absolute power (as in current net metering policies), and thus it will have to adjust the equal fraction of power contributed by each site over time as it varies to maintain the limit. In this case, we can compute this *equal rate* across all sites as simply the aggregate limit (L) divided by the sum of the current power output (P) of each of the n sites at any time t . We can augment this approach to include a weight, as in weighted fairness (Demers et al. 1989), such that the allocated rates are in proportion to each site’s weights, rather than being equal:

$$Rate(t) = \min \left(\frac{L(t)}{\sum_{i=1}^n P_i(t)}, 1 \right). \quad (2)$$

However, as discussed above, this does not result in an equal (or weighted) contribution of energy over time. Figure 3 illustrates this behavior for Homes 1 and 2 in Figure 2. When the rate, expressed as a fraction of each site’s maximum generation potential, is always equal (top), the fraction of energy each contributes diverges (bottom). In this case, the aggregate limit is static and set to 5kW throughout the day. Since Home 2 does not generate any power early in the day, Home 1 is able to feed a disproportionate amount of energy into the grid. Then, once Home 1 starts generating power, Home 1 and Home 2 each feed power in with equal rates. However, as the bottom graph indicates, the initial generation early in the day enabled Home 1 to feed in more energy (as a fraction of its total energy generation potential) relative to Home 2. In this case, Home 1 fed in 10% more energy than Home 2 in only a single day. Since this behavior is the result of a fixed object, e.g., trees, shading Home 2 early in the day, the unfairness will manifest itself every day of the year.

To address this problem, we design a rate allocation algorithm that enforces fair energy access to the grid. We first discuss a centralized version of this algorithm, assuming a tightly coupled system, and then present a distributed version. In both cases, the algorithms first start by computing the equal rates above and then determine which and how much sites can deviate from this equal rate based on their current cumulative fraction of energy. We use the equal rate allocation as a starting point, since we require some basis for assigning initial rates to users. Equal rate allocation represents a good starting point, since under ideal conditions, i.e., where sites have exactly the same solar profile at all times, setting equal rates above *will* result in equal long-term energy

Table 1. Variable Definitions for Algorithms 1 to 3

Variable	Description
n	Number of solar sites
i	Index of sorted homes
T	Duration over which the fairness is enforced
$P_i(t)$	Maximum power that a site i can generate at time t
$D_i(t)$	Power demand of a site i at time t
$p_i^{assigned}(t)$	Fraction of maximum power assigned to site i at time t
Energy Fraction (EF)	Fraction of solar energy fed into the grid over interval T for a given site i
Fair Energy Fraction (FEF)	Fair fraction of solar energy over interval T
$F(t)$	Fractional limit on solar capacity with respect to either grid demand or solar power at time t
$L(t)$	Aggregate limit on solar capacity at time t
$p^{avail}(t)$	Difference between aggregate limit and assigned power at sites at time t
$p_{est}^{agg}(t)$	Estimated aggregate power
K	Correction gain

contributions. Only when the solar profiles diverge does the equal rate allocation also diverge from a fair long-term solar energy allocation.

3.1 Centralized Algorithms

Algorithm 1 shows the pseudocode for our centralized algorithm, which we label as fast centralized allocation. Table 1 defines the algorithm's variables. In the centralized case, we assume that each solar site knows the fraction of solar energy each other site has fed into the grid over the current time window T , e.g., a month, which we call the energy fraction (EF). The algorithm then simply sorts each solar site by its EF and assigns rates based on a solar site's position in the list. In particular, lower-ranked solar sites get allocated higher rates than higher-ranked solar sites to allow them to "catch up." The algorithm enables sites to catch up fast, since it allocates rates to 100% of solar power in sorted order, starting with the lowest-ranked site, until it reaches the aggregate power limit or it reaches a site that has an energy fraction equal to the mean across all sites, which we call the fair energy fraction (FEF). At this point, the algorithm sets the rates of sites with energy fractions above the FEF based on the fair rate allocation algorithm above, but where the limit $L(t)$ is the remaining power after setting rates for the low-ranked sites. Thus, the algorithm is work conserving in that it does not penalize sites that have contributed more than their fair energy by not allowing them to feed solar into the grid. As above, we can also apply a weight to each site, such that the fraction of energy they feed in should be in proportion to their weight. Since this algorithm enables sites that are behind to catch up fast, we refer to it as the "fast" centralized allocation algorithm.

One problem with the algorithm above is that it has the potential to starve out solar sites if other sites are not able to feed in solar for a long period. For example, after a snowstorm, the snow may melt off solar modules at different rates, enabling large differences in their maximum power. As a result, some solar site may not be able to feed power into the grid, and will thus "get behind" in terms of its energy contribution. Once the snow melts from this solar site, the algorithm above would set its rate to 100% until it catches up, which would reduce the rates of the other solar sites.

ALGORITHM 1: Centralized Energy Allocation (Fast)**Require:** $P_i(t)$ and $P_i^{assigned}(t)$ for all homes over time T , $L(t)$

- 1: Compute $EF_i = \frac{\sum_{t=0}^T P_i^{assigned}(t)}{\sum_{t=0}^T P_i(t)}$, $\forall i$
- 2: Compute $FEF = \frac{\sum_{i=0}^n \sum_{t=0}^T P_i^{assigned}(t)}{\sum_{i=1}^n \sum_{t=0}^T P_i(t)}$
- 3: Sort & index homes in ascending order of EF ($\uparrow_{i=1}^n$)
- 4: $p^{avail}(t) = L(t) - \sum_{i=1}^n P_i^{assigned}(t)$
- 5: **while** ($p^{avail}(t) > 0$) **do**
- 6: **if** ($EF_i < FEF$) **then**
- 7: $P_i^{assigned}(t) = P_i(t)$
- 8: Update $p^{avail}(t)$, $i++$
- 9: **else**
- 10: **break**
- 11: $Rate(t) = \frac{p^{avail}(t)}{\sum_{i=1}^n P_i(t)}$ for homes above FEF
- 12: Update EF for all homes

ALGORITHM 2: Centralized Energy Allocation (Slow)**Require:** $P_i(t)$ and $P_i^{assigned}(t)$ for all homes over time T , $L(t)$

- 1: Compute EF_i as in Algorithm 1
- 2: Compute FEF as in Algorithm 1
- 3: Sort & index homes in ascending order of EF ($\uparrow_{i=1}^n$)
- 4: Compute fair rate $Rate(t) = \frac{L(t)}{\sum_{i=1}^n P_i(t)}$
- 5: $p^{avail}(t) = L(t) - \sum_{i=1}^n P_i^{assigned}(t)$
- 6: **while** ($p^{avail}(t) > 0$) **do**
- 7: **if** ($EF_i < FEF$) **then**
- 8: $P_i^{assigned}(t) = (1 + (FEF - EF_i)) \times Rate(t)$
- 9: Update $p^{avail}(t)$, $i++$
- 10: **else**
- 11: **break**
- 12: $Rate(t) = \frac{p^{avail}(t)}{\sum_{i=1}^n P_i(t)}$ for homes above FEF
- 13: Update EF for all homes

To mitigate the starvation problem, we can limit the catch-up rates for sites that are behind. In this case, rather than set these sites to 100% of their maximum power, we can set a limit between the equal rates computed in Equation (2) and 100%. In our algorithm, we apply proportional control to set these rates, such that the more behind a solar site is, the faster it catches up. Specifically, we increase Equation (2)'s rate by the same proportion the site is behind in energy. Thus, if a solar site has 20% less than its "fair" fraction of aggregate energy, we allow it to increase its rate in Equation (2) by 20%. Algorithm 2 shows the pseudocode for this algorithm, which we label as the "slow" centralized allocation, where line 7 applies the proportional adjustment to the rate.

ALGORITHM 3: Distributed Energy Allocation**Require:** $L(t)$, $P_{agg}^{est}(t)$, and $Rate(t)$ over time T

- 1: Estimate aggregate power $P_{agg}^{est}(t)$ using gossip protocol
- 2: Compute EF_i as in Algorithm 1
- 3: Estimate $FEF = \frac{\sum_{t=0}^T (P_{agg}^{est}(t) \times Rate(t))}{\sum_{t=0}^T P_{agg}^{est}(t)}$
- 4: Compute fair rate $Rate(t) = \frac{L(t)}{P_{est}^{agg}(t)}$
- 5: $P_i^{assigned}(t) = (1 + K(FEF - EF_i)) \times Rate(t)$

3.2 Distributed Algorithm

The centralized algorithms above assume accurate generation information is available from all solar sites in real time, and that they are able to instantaneously set the rates of all solar sites without any delay. This implies that solar sites form a tightly coupled system with utilities, where they stream generation data to utilities in real time and utilities are able to instantaneously control their rates. Implementing such a tightly coupled system is not realistic today. Most smart meters communicate wirelessly over cellular networks and thus have limited bandwidth and connectivity issues. A centralized approach also represents a single point of failure and is not robust to network failures. Thus, we present a distributed approach that uses incomplete information propagated at lower rates, e.g., minutes to hours.

Our distributed approach assumes that individual sites do not have reliable network connectivity, and thus must set their own allocations based only on an estimate of the globally fair energy fraction. Individual sites do not know the specific power and energy generation of other sites, and thus cannot compute precise rates that satisfy the aggregate limit and correctly apportion fair rates across sites. Instead, individual sites can only increase or decrease their rate relative to the equal rates in Equation (2) and based on the difference between the globally fair energy fraction and their local fraction of energy. In this distributed approach, sites that are both above and below the globally fair energy fraction decrease and increase, respectively, the rate in Equation (2) by the same proportion that the site is ahead or behind in energy.

Algorithm 3 shows the pseudocode for this algorithm, which we label as distributed energy allocation. Each solar site independently runs the distributed algorithm at a specified interval to determine their solar rate. The length of this interval represents the expected time between disseminating new generation information to other solar sites. While each solar site can broadcast to all other solar sites, full mesh communication has the same issues as the tightly coupled centralized approach. Instead, inspired by distributed rate limiting in networks (Raghavan et al. 2007), our approach uses a more robust push-sum gossip protocol that periodically disseminates recent generation information to a random set of N other sites each interval (Kempe et al. 2003). This push-sum gossip protocol takes a few intervals to converge, such that each site has an accurate estimate of the “fair” fraction of global energy and the global equal rate from Equation (2). We also add a multiplicative gain factor, K , as a configurable parameter to adjust how fast sites catch up in the distributed algorithm, similar to Algorithm 2.

3.3 Fidelity of Control

Both the centralized and distributed algorithms must make decisions based on stale information, as solar power changes continuously in real time. In the centralized case, even though this time period may be small, e.g., 1 minute, solar output can fluctuate significantly even over these short time periods. Since large fluctuations can have a negative impact on electronics, the fidelity of

the control, i.e., how close the algorithm is able to maintain the aggregate limit that is set, is an important performance metric. In addition, large fluctuations in the rates from the algorithm can also have a negative impact on the electronics that control solar output, and thus are also undesirable. As we show, the centralized algorithm with a fast catch-up suffers from increased fluctuations as it periodically focuses solar allocation on a few sites by increasing their rates to 100% of maximum output, and thus causes large changes in allocated rates. Of course, the distributed algorithm may also take more time to propagate information, causing it to diverge more from the aggregate limit. We evaluate the fidelity of control and fairness of this algorithm under different conditions in Section 6.

4 POLICIES: CONTROLLING DISTRIBUTED SOLAR CAPACITY

The previous section's algorithms define a new mechanism that enables utilities to set a configurable absolute cap on distributed solar capacity, such that the aggregate generation never exceeds the cap *and* it maintains distributed solar fairness across sites; i.e., the energy contributed by each solar site converges to the same percentage of each site's maximum possible energy generation. This mechanism provides utilities a new tool in managing and integrating increasing penetrations of solar energy with their conventional generation resources. To effectively leverage this new tool, utilities must define *policies* that configure the mechanism and alter the caps over time to effectively control both the magnitude and variation of distributed solar output. There are a wide range of policies utilities could use in controlling the cap. We define two different types of policies below and discuss their tradeoffs in regulating solar power.

4.1 Demand-Based Policy

Our first class of policy sets the cap on solar as some fraction of the grid's demand. Thus, every interval T , the utility will set the limit on aggregate solar power equal to a fraction of the average grid demand expected over T . Utilities already have models that accurately predict grid demand based on weather conditions and the day, which they can leverage to estimate near-term grid demand. This policy is inspired by the current regulatory structure in most states, which seeks to indirectly limit the magnitude of solar penetration in the grid as a function of the grid's demand. The policy enables solar sites to contribute more solar power as grid demand increases within the day, and over time. The idea behind this policy, and current regulations, is that limiting the magnitude solar power fed into the grid (1) ensures that utilities are kept solvent by requiring their customer base to purchase a specified fraction of energy from them and (2) indirectly limits the relative magnitude of the variations in net demand that utilities must balance with conventional generators. In effect, this policy "shapes" the solar curve to match the shape of the demand curve, but with a smaller magnitude. As a result, the policy limits the additional variance caused by high penetrations of solar. Algorithm 4 shows the policy's pseudocode, which we label as the demand-based policy.

As with the current static first-come-first-serve policies, this policy benefits utilities, enabling them to maintain a fixed percentage of solar penetration. However, since the policy changes the limit based on real-time demand, it enables more sites to feed solar into the grid at nonideal times, e.g., cloudy days, relative to current regulations, which prevent many sites from ever connecting to the grid. In addition, the policy enables more solar to feed into the grid in the middle of the day, when demand and solar are both high. Of course, grid demand is the lowest overnight, when solar is not producing any power. Regulators can also leverage this policy to gradually increase solar penetration by increasing the fraction of grid demand that sets the limit, rather than preventing or halting connecting solar installations to the grid.

ALGORITHM 4: Demand-Based Policy**Require:** $D_i(t)$ for all homes over time $T, F(t)$

- 1: $L(t) = F(t) \times \sum_{i=1}^n D_i(t)$
- 2: Call Algorithm 1

ALGORITHM 5: Solar-Based Policy**Require:** $P_i(t)$ for all homes over time $T, F(t)$

- 1: $L(t) = F(t) \times \sum_{i=1}^n P_i(t)$
- 2: Call Algorithm 1

4.2 Solar-Based Policy

The policy above ties the aggregate limit on solar generation to grid demand and only indirectly reduces the magnitude of variations for high penetrations of solar power. Instead, we can define another policy that directly regulates solar output, independent of demand, to directly reduce the magnitude of variations in solar power. In this policy, we set the cap on solar as some fraction of the grid's current aggregate solar generation. Thus, every interval T , the utility will set the limit on aggregate solar power equal to a fraction of the average aggregate solar generation expected in the grid over interval T . This policy enables utilities to improve operations by smoothing the aggregate solar profile over the entire day to enable a more predictable net demand. Since the policy adjusts to set the limit below the maximum possible output, it is able to precisely control the amount of solar fed into the grid. In contrast, a fixed cap (as in the previous section) or a demand-based cap (as discussed above) may periodically impose a cap that is much greater than the current maximum generation. As a result, there will be no limit on the solar generation, enabling it to vary uncontrollably, as in today's grid. Algorithm 5 shows the pseudocode for this policy, which we label as solar-based policy.

5 IMPLEMENTATION

We evaluate our centralized and distributed algorithms from Section 3, and our demand- and solar-based policies from Section 4, in simulation using both real and synthetic solar traces. We derive our synthetic solar traces from clear sky solar irradiance models implemented in the Pysolar Python library (pys 2017). The resolution of this synthetic solar data is 1 minute, and we convert the irradiance into power assuming a typical solar module efficiency of 18%. We then vary the maximum solar capacity of different sites from 1 to 20kW, and also vary the orientation and tilt angles of the simulated modules. For our real solar sites, we use data from 50 solar sites in the western part of the United States. We implement our simulator in Python and vary the simulated interval by which each site propagates its generation information.

6 EVALUATION

We first evaluate our mechanism for fairly allocated distributed solar capacity from Section 3 under different fixed caps, and then evaluate our policies from Section 4 for using this mechanism by changing the cap over time to control the net grid demand.

6.1 Mechanism Evaluation

We evaluate both the impact of diversity in solar output on fairness using the equal rate allocation algorithm, as well using the different variants of our fair solar energy allocation algorithm. In addition, we also evaluate the tradeoff between fairness and the fidelity of the algorithm to maintain

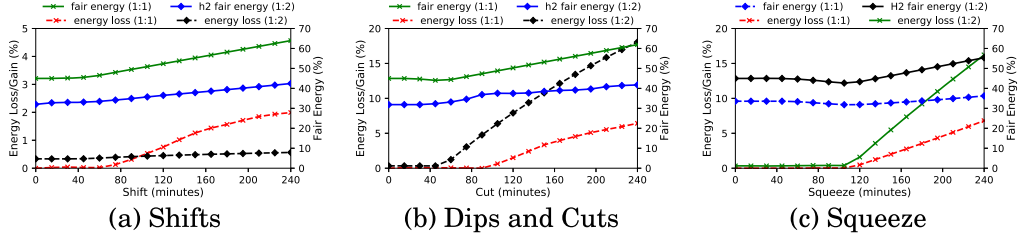


Fig. 4. Fairness as a function of the magnitude of shifts (a), cuts (b), and squeezes (c) for two solar sites.

an aggregate limit. We quantify the fidelity using the Mean Absolute Percentage Error (MAPE) between the limit and the actual aggregate generation, as shown below. For these experiments, we use a set of 18 homes over a month-long period:

$$MAPE = \frac{100}{T} \sum_{t=0}^T \left| \frac{L(t) - \sum_{i=0}^n P_i^{assigned}(t)}{L(t)} \right|. \quad (3)$$

Note that, in this section, we only compute the MAPE for all t where we enforce the aggregate limit, i.e., $\sum_{i=0}^n P_i(t) > L(t)$. The fidelity is not relevant when the mechanism is not used to enforce the limit, similar to how fair queuing results in networks are only relevant when all flows are backlogged. Quantifying fairness is more challenging than accuracy, since average fairness metrics, such as Jain’s fairness index, can obscure highly unfair behavior between any two sites by averaging over many sites. For example, if there are many flows, Jain’s fairness index can be close to 1 (indicating a fair allocation) even though some set of solar sites (or solar “flows”) may experience highly unfair allocations. Since energy fed into the grid directly correlates with money, unfairness even among a few users is problematic. Thus, we avoid aggregate measures of fairness across many sites, and instead quantify fairness by examining the distribution of energy allocations across sites.

6.1.1 Microbenchmarks: Shape Diversity. Figure 4 first looks at the impact on fairness between two solar sites for different magnitude shifts, dips and cuts, and squeezes. We use the equal rate allocation algorithm, which always satisfies the aggregate limit by setting rates equal to each other. For this experiment, we use synthetic data based on clear sky generation for two sites at the same location, and then alter one site’s generation to shift it, cut it, or squeeze it by a certain amount of time. Thus, these results do not include other effects that could impact energy fairness, such as weather, location, or tilts. The results are also a function of the aggregate limit, which we set to 14kW in this case, where the maximum power of the sites is 10kW (or 20kW total). These experiments quantify the effects over an entire year, and include two scenarios: one where the weights are equal (where each site should contribute the same fraction of their maximum solar energy potential) and one where the weights are in a 1:2 ratio.

Figure 4(a) shows the effect of a shift, where the x-axis indicates the duration of the shift, the right y-axis is the percentage of energy lost due to unfairness in the allocation, and the left y-axis is the mean fraction of energy the solar site should have fed into the grid. The figure shows that the energy loss is only modestly impacted by shifts (1% to 2%), in large part because they cancel each other out, such that a shift increases one site’s allocation at the beginning of each day but decreases it at the end of each day. As also illustrated in Figure 2, cuts (in Figure 4(b)) have a much larger impact on the energy loss, causing one site to lose nearly 10% of its energy relative to a fair allocation in the case of equal weights, and nearly 20% when weights are in a 1:2 ratio. The unequal weights increase the relative loss, since it exacerbates the amount of solar power one site

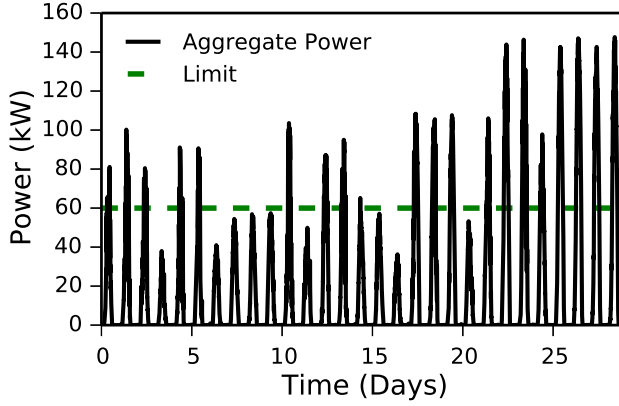
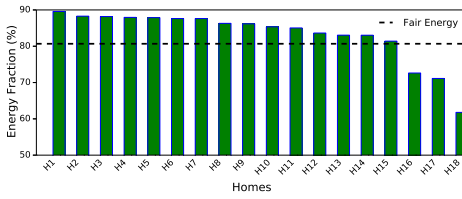
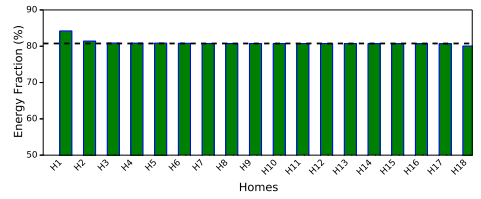


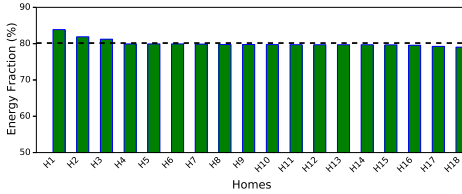
Fig. 5. Aggregate power of the 18 homes over 30 days.



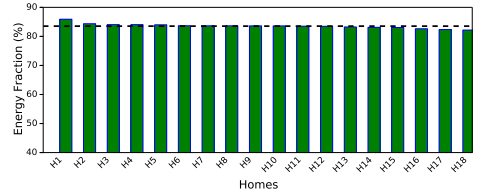
(a) Equal Rate Allocation



(b) Centralized Energy Allocation (Fast)



(c) Centralized Energy Allocation (Slow)



(d) Distributed Energy Allocation

Fig. 6. Distribution of solar energy allocation under a limit of 60kW for the equal rate (a), centralized fair energy (fast) (b), centralized fair energy (slow) (c), and distributed fair energy algorithms (d).

is able to feed into the grid when another site is unable to generate power. This effect is similar for squeeze with losses near 10% and 20%, respectively, with equal and weighted rates.

6.1.2 Fair Solar Energy Allocation. The previous subsection demonstrated the relative difference in fairness between two ideal synthetic homes with different shifts, cuts, and squeezes. We also experiment with controlling a small group of 18 homes in the western United States to get a sense of the differences in energy allocation across many homes with real solar power. In this case, we experiment with the equal rate allocation algorithm, as well as the three different variants of our fair energy allocation algorithm, including the centralized algorithm with fast catchup, the centralized algorithm with slow (proportional) catchup, and our distributed algorithm. For these experiments, we assume all the rates are equal and set the limit to 60kW. Figure 5 shows the aggregate power across all the homes over a month-long period and the 60kW limit. Here, we maintain a fixed limit: our policies in the next subsection present results with a variable limit.

Figure 6 then shows the distribution of the energy gain/loss relative to the fair energy in each case over a 1-month period, which corresponds to a typical billing cycle. Note that this

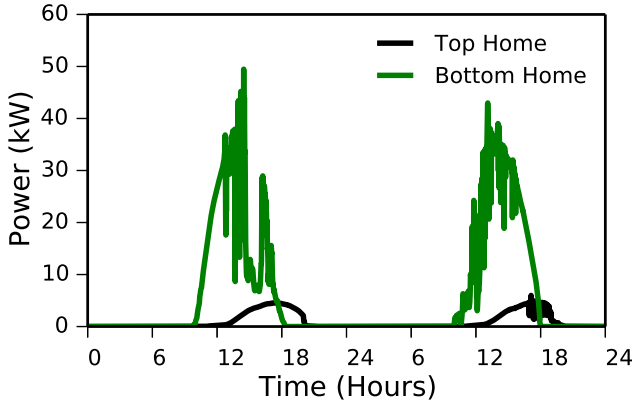


Fig. 7. Daily profiles of home H1 and H18 in Figure 6(a).

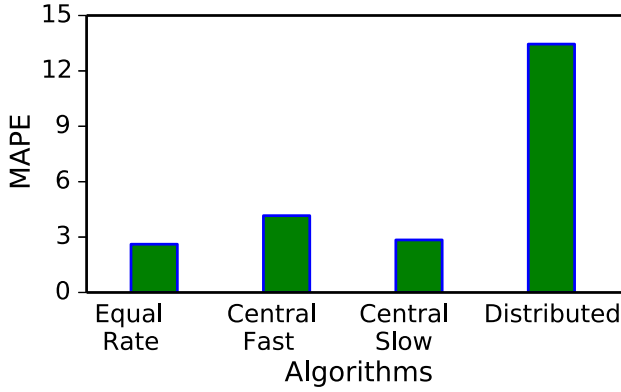


Fig. 8. Fidelity of each algorithm at enforcing a limit with a 1-minute communication interval.

percentage directly translates into the fraction of money gained and lost from net metering. In the equal rate allocation case (a), the largest difference is over 27%, such that one home gets 27.8% less than another home and 18.9% less than its fair energy allocation. For each of the other algorithms, the percentage drops to near 0%, since they explicitly attempt to maintain a fair energy allocation over time. Figure 7 then shows a sample sunny day for both the most advantaged and disadvantaged solar sites with equal rate allocation; we can see from this graph the impact of shifts, cuts, and squeezes on fairness, as these two homes have significantly different solar curves. As the figure shows, these sites have significantly different capacities, with one site having a capacity near 50kW and the other having a capacity of only 7kW. Note that a goal of our fair solar energy access algorithm is to enable both of these sites to contribute the same fraction of their maximum generation potential, which is relative to their capacity. In contrast, despite these differences, in all variants of the fair energy allocation algorithms, we see this difference narrowing significantly, with all having a difference of less than 1% in terms of grid energy access over the month.

In all of the algorithms above, we assume a 1-minute update interval, such that the solar rate is updated once every minute based on data from the previous minute. Figure 8 shows the fidelity of each algorithm in maintaining the limit with this update interval. We see that the equal rate allocation has the highest fidelity (corresponding to the lowest MAPE), since it adjusts rates instantaneously. The small divergence here is due to the minute-to-minute changes in solar power,

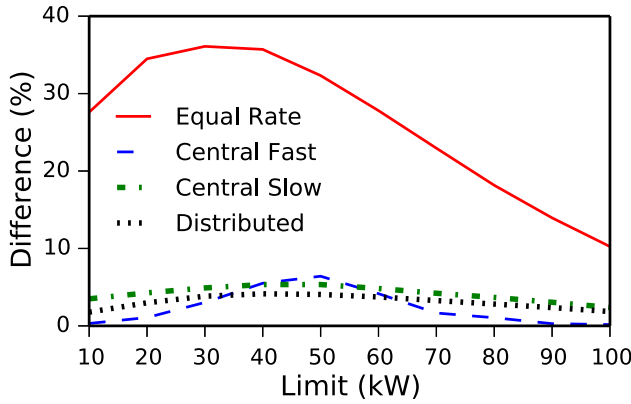


Fig. 9. Energy difference between H1 and H18 in Figure 6(a), as a function of the aggregate limit.

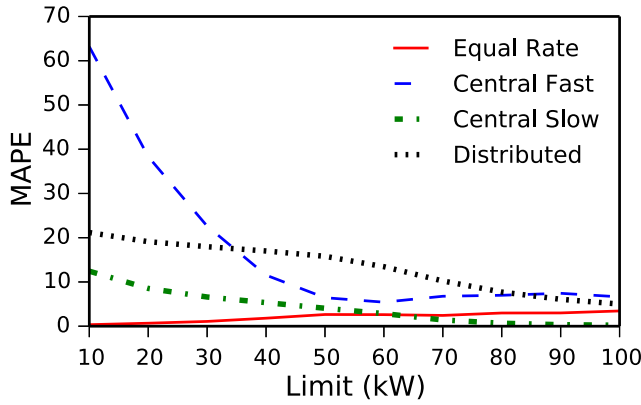


Fig. 10. Fidelity of maintaining the aggregate limit as a function of its magnitude for the different variants.

as the algorithm can only adjust rates after it senses that solar output has changed (which takes 1 minute in this experiment). The centralized algorithm with the fast catchup has a lower fidelity, which is also exacerbated due to the stochasticity in solar output at the minute level. This algorithm results in highly imbalanced rates during its catch-up phase, where some solar sites are contributing 100% of their energy generation. As a result, if these sites change their output significantly within a minute (before the rates are updated), the aggregate solar power will diverge from the limit. In the equal rate allocation case, the likelihood of such aggregate changes is low because it requires all sites to suddenly change their output in unison. However, when a small number of sites are catching up and have a disproportionate share of grid energy, it increases the likelihood that these sites will alter their generation within a minute. The centralized algorithm that uses a slower proportional catch-up mitigates this effect and has a MAPE near that of the equal rate algorithm. Finally, the distributed algorithm has significantly lower fidelity than the others due to its long propagation delays.

The deviation above changes with the limit as shown in Figure 9. For the equal rate algorithm, the unfairness decreases as the limit increases, since it mitigates the effect of differences in the solar curve between sites. However, the difference between the different variants of our fair energy algorithms remains largely constant and generally under 5%. However, Figure 10 shows that the equal rate algorithm has the highest fidelity across all aggregate limits. For the fair energy

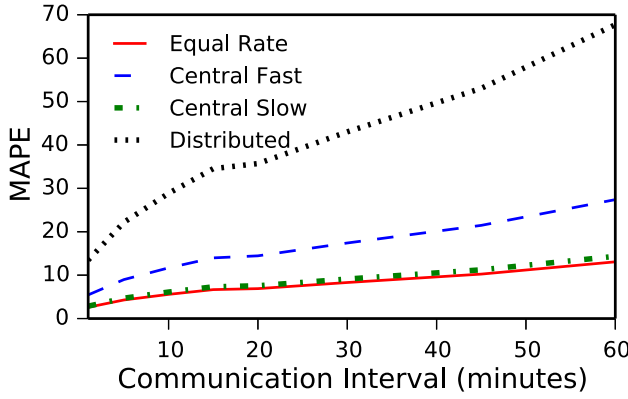


Fig. 11. As the propagation delay increases, the fidelity of control for the distributed algorithm decreases.

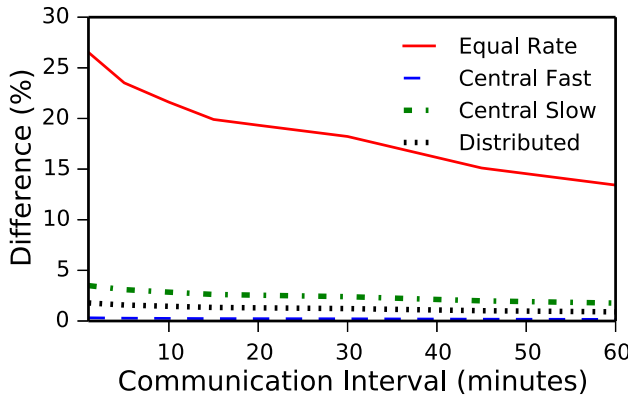


Fig. 12. As the propagation delay increases, the fairness for the distributed algorithm also increases.

algorithms, the lower the limit, the worse the fidelity at maintaining the aggregate limit. This impact of low limits is particularly severe for the centralized algorithm with fast catch-up, since at low limits it is subject to increasingly more extreme versions of the effects described above.

6.1.3 Distributed Algorithm. Finally, we explore the impact of information propagation delay in the distributed algorithm. Figure 11 shows this delay on the x-axis, while the y-axis shows the resulting MAPE relative to the limit. The graph demonstrates that, as expected, the fidelity of the control decreases (yielding a higher MAPE) as the propagation delay increases. This increase is faster for the distributed algorithm, since it takes some time for the rates to converge. However, in contrast, fairness actually *improves* as the delay increases. Figure 12 shows the percentage maximum difference in the percentage of energy gain/loss between any two homes (in this case, H1 and H18 from Figure 6(a)). The graph shows that as the propagation delay increases, this percentage trends toward 0%. Of course, the equal rate algorithm is unfair and thus takes longer to converge. With longer propagation delays, solar sites operate at the same fraction of power for longer windows of time. As a result, the amount of energy they contribute to the grid relative to each other converges. Thus, our fair energy access algorithms enable a tradeoff between propagation delay, fidelity, and fairness.

Figure 13 illustrates the fidelity of maintaining an aggregate 60kW limit for the distributed algorithm over a representative sunny day with a communication interval of 1 minute. The graph

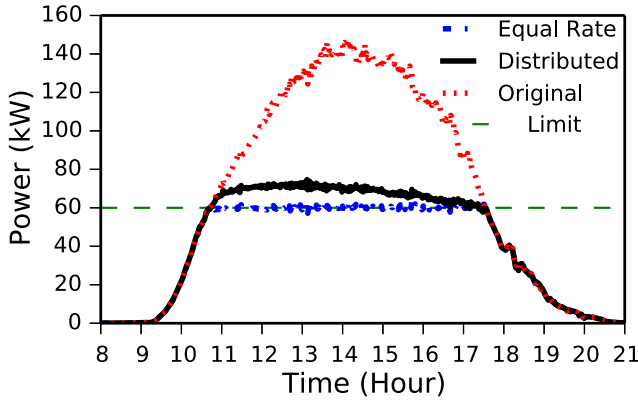


Fig. 13. Maintaining the limit at a 1-minute interval.

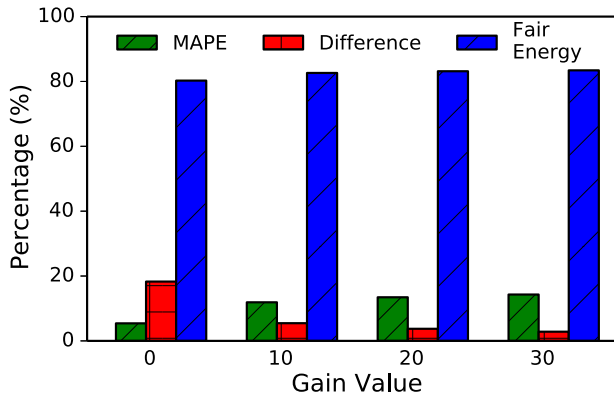


Fig. 14. Impact of accelerating the “catch up” of sites that are behind in their fair energy allocation by a multiplicative gain factor in the distributed algorithm.

shows that the centralized equal rate algorithm is able to maintain the 60kW limit precisely, while the distributed algorithm maintains a limit that is slightly above the 60kW threshold. Finally, Figure 14 shows how we mind the gap between fidelity and fairness by accelerating the catch-up amount in the distributed algorithm. In this case, we specify a gain value, which is a multiplicative factor applied to the typical rate computed by the distributed algorithm (which enables sites to increase their rate in proportion to the amount of energy they are behind). Here, a gain of 0 indicates no additional increase, while a gain of 10 increases the rate by a factor of 10. The graph illustrates the tradeoff between fairness and fidelity: as we increase the gain value (to accelerate catching-up sites that are behind in their energy allocation), the MAPE of the aggregate limit increases (reducing the fidelity of control), while the fairness increases (as specified by the decrease in the largest difference in energy allocation between two sites). For comparison, we also plot the fair energy fraction for the distributed algorithm, which increases slightly, as more power is fed into the grid (as a result of overshooting the limit as seen in Figure 13).

6.2 Policy Evaluation

The previous subsection evaluated the fidelity and fairness of our rate allocation mechanism using a fixed limit over time on aggregate solar power across many sites, as illustrated in Figure 13.

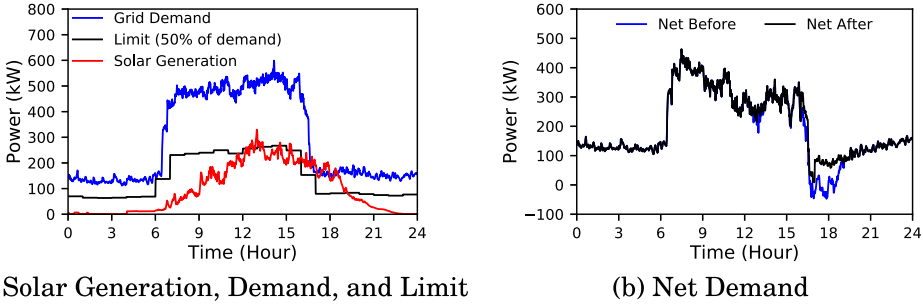
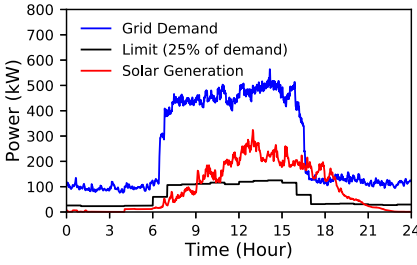


Fig. 15. Illustration of varying the solar limit based on grid demand (a). In this case, we set the limit to 50% of grid demand, resulting in only a short period where solar generation exceeds the limit. As a result, the net demand does not change significantly after setting the limit (b).

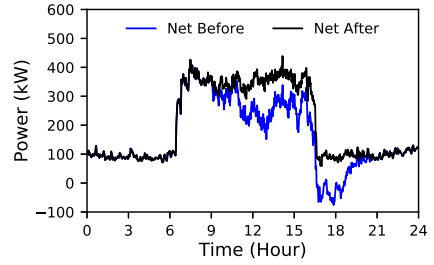
However, fixing the limit is not ideal as the aggregate solar generation across sites is highly variable throughout the day, e.g., rising in the morning and decreasing in the evening, while also changing due to weather. Thus, a fixed limit may be well below the aggregate solar generation at some times and well above it at other times. In the latter case, when the limit is well above the solar generation, our mechanism imposes no limitations on solar generation, leaving it uncontrollable with the potential for high variations that can complicate grid operations. In addition, a high limit does not restrict the magnitude of solar energy net metered in the grid. Likewise, in the former case, if the limit is well below the possible solar generation, then a utility may curtail more solar power than necessary.

Thus, our policies from Section 4 alter the limit over time, as the grid demand and solar capacity change. Note that, in the previous section, we compute the fidelity of our mechanism in satisfying the limit using the MAPE only over the time period when the limit is enforced. However, since our policies attempt to limit solar power at all times, by setting an appropriate limit, in this section, we compute fidelity, i.e., MAPE, and fairness over the entire time period. We run these experiments across a set of 50 homes in the western United States. We describe two classes of approaches: one that sets the limit as a fraction of grid demand and one that sets the limit as a fraction of solar generation. We evaluate each case below. We also compare with a policy that sets a fixed cap, as described in the previous section. In this case, we set the fixed limit for a given day to be $x\%$ of maximum predicted power generation for that day for some value of x . For example, if the maximum power generation is predicted to be 1,000kW, the 80% limit would mean that the fixed limit is set at 800kW. Note that, with the exception of the fixed cap, the limits for both the demand- and solar-based policies are dynamic, and set relative to different values, e.g., grid demand versus aggregate solar. Thus, setting the same limit for different policies may result in a different fraction of solar energy being net metered.

6.2.1 Demand-Based Policy. Figures 15 and 16 illustrate the impact of imposing a limit on solar generation across 50 homes based on the grid demand. In this case, our policy sets the limit at 50% and 25%, respectively, of average grid demand during the next control interval, which is 1 hour in these experiments. Figure 15(a) shows that solar generation only exceeds 50% of the grid demand for a brief window between 4 p.m. and 7 p.m. As a result, the difference between the unaltered net demand before imposing the limit is similar to the net demand after imposing the limit, as shown in Figure 15(b). However, Figure 16(a) shows that setting the limit at 25% of grid demand results in a more significant change in generation, since the limit is enforced for over half the day. Figure 16(b) shows a significant increase in the net demand after imposing the limit, demonstrating

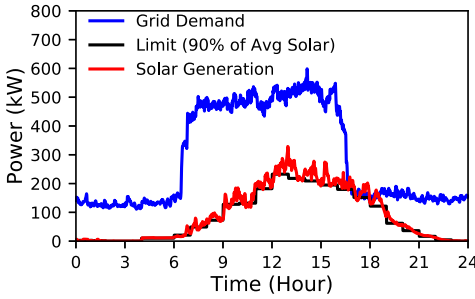


(a) Solar Generation, Grid Demand, and Limit

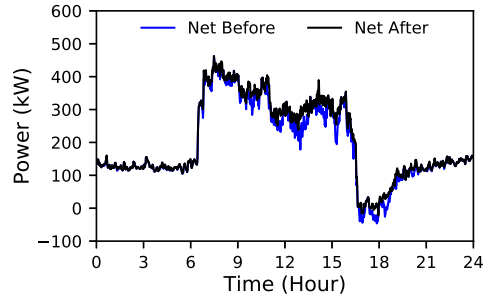


(b) Net Demand

Fig. 16. Illustration of varying the solar limit based on grid demand (a). In this case, we set the limit to 25% of grid demand, resulting in solar generation exceeding the limit for over half the day. As a result, the net demand increases significantly after setting the limit (b).

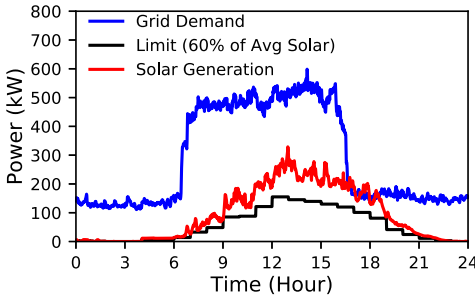


(a) Solar Generation, Demand, and Limit

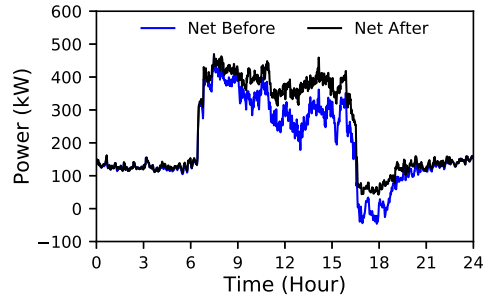


(b) Net Demand

Fig. 17. Illustration of varying the limit based on the solar generation (a). In this case, we set the limit to 90% of solar generation. As a result, the net demand does not change significantly after setting the limit (b).



(a) Solar Generation, Demand, and Limit



(b) Net Demand

Fig. 18. Illustration of varying the limit based on the solar generation (a). In this case, we set the limit to 60% of solar generation. As a result, the net demand changes significantly after setting the limit due to the reduced solar output (b).

that a demand-based policy is an effective tool for limiting the magnitude of solar in the grid. In this case, that increase in net demand would need to be satisfied by the utility using conventional mechanical generators. While not environmentally friendly, this scenario is preferable to utilities that must recoup their long-term investment in these generation resources to remain solvent.

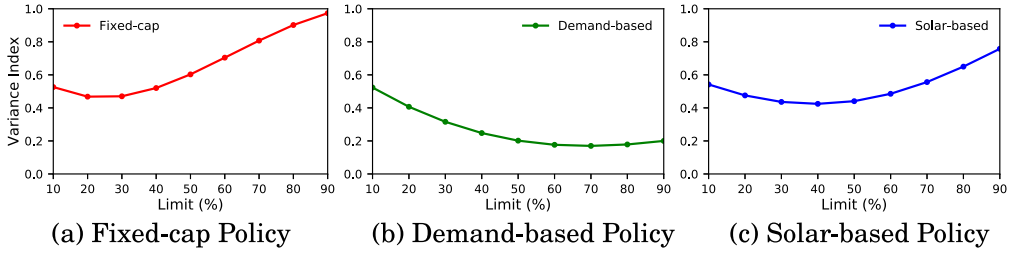


Fig. 19. The graph shows the variance index as a function of the limit for fixed, demand-based, and solar-based policies. The variance index is the ratio of variance in the net demand before the limit is imposed to the variance of the net demand after the limit.

6.2.2 Solar-Based Policy. Figures 17 and 18 illustrate the impact of imposing a limit on solar generation based on the average aggregate solar generation. In this case, we set the limit to 90% and 60% of the expected average solar generation during the next control interval, which is 1 hour for these experiments. Although we set the limit each hour, solar generation changes continuously over time. Figure 17(a) shows the grid demand, solar generation, and limit for 50 homes over a day. We can see that the limit serves to smooth the solar generation by reducing its variability. This occurs because the limit voluntarily restricts solar generation to 90% of the expected average generation over the hour. As a result, during each hour, the aggregate solar generation in the grid is near constant. Since 90% is a high cap, it only reduces the aggregate solar generation by a small amount, which keeps the net demand similar to the demand before imposing the limit, as shown in Figure 17(b). Figure 18 then shows the same scenario, but with a limit set to 60% of the solar generation. The lower limit smooths the solar generation even more (a), which in turn has a larger negative effect on the net grid demand (b). Compared to the demand-based policy, which is able to control the magnitude of solar in the grid as a fraction of demand, the solar-based policy is better able to control the variability in solar and the net demand profile.

6.2.3 Comparing the Grid- and Solar-Based Policies. The experiments above show the effect of our policies on solar generation and the net demand over 1 day. Below, we compare the variance, fidelity, and fairness of these policies over a month of data.

Variance. Figure 19 shows how the policies above are able to reduce the variability in the grid's net demand profile, which enables utilities to more effectively balance supply and demand. In the figure, the variance index on the y-axis is the ratio of variance in the net demand after the limit is imposed to the variance of the net demand before the limit. Thus, values less than 1 demonstrate that after imposing the limit, the variance decreases—so lower numbers are better. We see that imposing a limit at a very low value (10%) manages to reduce the variations by almost half (index ≈ 0.5). The fixed cap has the highest variance index, especially at high limits, since it rarely enforces a cap at these limits. As the limit increases, the variance index decreases to 70% to 80% for the demand-based policy and 40% to 50% for solar-based policy. The high value of the variance index at the lower limits is due to solar being capped well below its capacity, causing it to have little effect on the net demand. However, as the limit increases, the fraction of solar energy also increases, but the limit still caps the generation most of the time, which reduces the variance. As the limit increases further, the cap frequently exceeds solar generation, such that solar power is not capped. As a result, the limit is rarely imposed and the variance increases.

Fidelity. Figure 20 shows the fidelity of each of the policies in maintaining the limit with a 1-minute update interval. Since the limit for each policy is set in a different way, we list the fraction

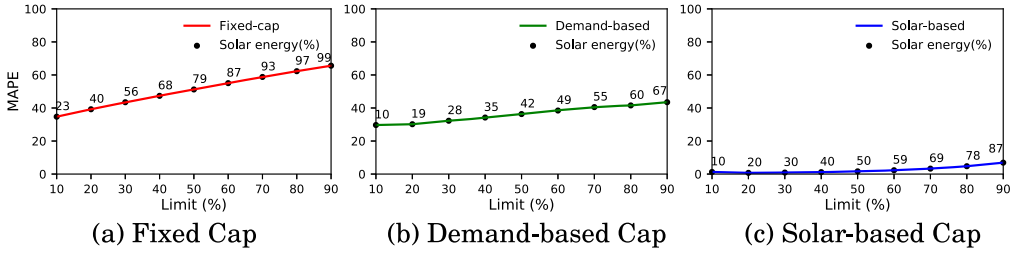


Fig. 20. The relationship between the fidelity of control for each of the different policies. The fixed cap has the worst fidelity since it does not alter the limit as solar generation changes over time.

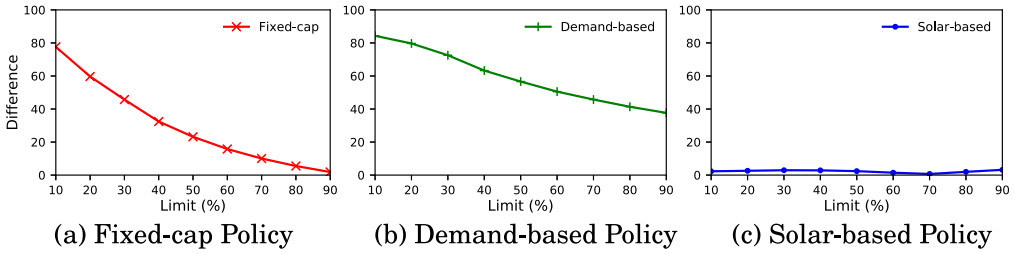


Fig. 21. For the fixed-limit policy and the demand-based capping, fairness decreases as the limit decreases. The solar-based capping policy exhibits the highest fairness as quantified by a low difference between the fraction of solar contributed by the highest site versus the lowest site.

of solar energy generated relative to the maximum solar energy atop each data point. As before, we quantify fidelity using the MAPE with respect to the ideal limit, but over the entire time period. In this case, we see that the fixed-limit case (a) has the lowest fidelity, which corresponds to the highest MAPE, since it is unable to alter the limit to match solar generation as it changes. As the limit increases on the x-axis, the MAPE increases, since less of the day is subject to the high fixed limit. The demand-based capping (b) does a better job of enforcing the limit simply because grid demand roughly approximates solar generation; i.e., it rises during the day and falls off at night. As a result, the limit is enforced for a longer period of time than in the fixed-limit case. Of course, capping based on solar generation has the highest fidelity (or the lowest MAPE) because the policy alters the limit to closely match the solar generation profile. As a result, the solar generation is nearly always less than the limit, such that the limit is always enforced.

Fairness. Figure 21 shows fairness, as a function of the difference in energy of the top and bottom homes in the distribution, among all 50 homes for each of the policies as the limit changes. We see that the demand-based capping policy (b) has the worst fairness properties. This occurs because altering the limit based on the grid's demand has a greater likelihood of setting a low limit when there is a difference in the generation of solar sites. Recall from Section 3 that unfairness arises when the solar generation profiles of sites differ significantly when the limit is enforced. The fixed limit case (a) also demonstrates poor fairness for the same reason. Thus, the fairness decreases as the limit decreases for both of these policies. In contrast, the solar-based capping policy (c) has the highest fairness, since it attempts to track the aggregate solar output more closely, i.e., by setting the cap as a fraction of the solar output. Thus, there are fewer time periods where a limit is enforced and sites' solar output differs significantly.

Summary. Our comparison above illustrates differences in the choice of policy for utilities. Ultimately, the choice of policy is subjective and dependent on a utility's goals. As we show, setting a

fixed cap is undesirable, since the magnitude of solar capacity varies throughout the day and with the weather. Thus, a fixed cap is unlikely to be satisfactory at all times, requiring utilities to vary the cap. Demand-based capping is effective at directly limiting the magnitude of solar generation to a fraction of demand, which is similar to what current admission control policies attempt to do. However, the policy has poor fidelity and fairness properties because it only indirectly regulates solar generation. In contrast, the solar-based capping has better fidelity and fairness properties because it directly limits solar generation.

7 RELATED WORK

There is a large body of work in the systems and networking literature on fair rate allocation and scheduling. This work differs from our work in that it focuses on maintaining instantaneous bandwidth fairness when flows are backlogged, and not fairness in the amount of data transmitted over long periods of time. Recently, there have been adaptations of this work to the electric grid to dynamically manage increasing penetrations of solar energy (Lee et al. 2017) and electric vehicles (Ardakanian et al. 2013). However, as we show, direct adaptations of instantaneous rate allocation from networks can result in *unfair* energy access. Similarly, iPlug (Rongali et al. 2016) proposes a policy for decentralized dispatch of solar power based on congestion-aware network protocols. iPlug differs from this work in that solar sites back off based on sensing grid congestion, e.g., due to a deviation in nominal values for voltage and frequency. One issue with this approach is that it requires degrading the power quality of the grid to send feedback signals. Balancing authorities are unlikely to allow such degradation in power quality. In addition, in modern countries, the grid's infrastructure is highly overprovisioned, requiring massive penetrations of distributed solar generation before iPlug would be able to sense any grid congestion, i.e., deviations in nominal voltage or frequency, that would act as a feedback signal. Thus, we adopt an approach that directly communicates generation via the network to maintain a fair energy allocation over time. Finally, iPlug's approach is not fair, since different users sense different voltages and frequencies depending on their position in the grid. For example, a user at the end of the distribution line will have lower voltages, and thus back off more than a user further up the line.

Enforcing fair energy access is important in the grid, since users directly receive compensation for the amount of energy that they net meter into the grid. Another key difference with prior work is that it generally assumes the key constraints are in the network: the capacity of the transformers and feeders that are analogous to network switches and routers. However, we assume the network is unconstrained, and that unfairness can arise simply from the differences in the generating potential (or "workload") between solar sites independent of network constraints. Importantly, sites are unable to control this generating potential in the same way that network clients can control when they send traffic. Prior work also does not explore the fidelity of control based on the time to propagate generation information in a distributed system. Prior work in the power systems community also explores different strategies for curtailing solar power. However, these approaches have largely focused on preserving the reliability of the grid and responding to overvoltage situations (Tonkoski and Lopes 2011; Lew et al. 2013; Tonkoski et al. 2009, 2010, 2011). Instead, our work focuses on enabling fair control of distributed solar capacity, which has not been a metric of interest in prior work.

Finally, much of the work above conflates mechanism and policy in controlling solar generation. For example, iPlug only requires that solar deployments back off when it senses grid congestion (Rongali et al. 2016). This is only one of many possible policies for controlling solar generation. In fact, most current policies for limiting solar generation are business oriented and not technically oriented in that their purpose is to encourage solar adoption while enabling utilities to effectively manage the grid. In contrast, our work clearly separates policy from mechanism. Our fair rate

allocation algorithms collectively define a mechanism that utilities can use to control the output of distributed solar capacity, and our proposed policies leverage mechanism in different ways to achieve different goals.

8 CONCLUSIONS

This article highlights an important difference between fair rate allocation in networking and enforcing “fairness” in the grid. In particular, enforcing fairness based on the relative amount of energy injected into the grid over time is more important than enforcing instantaneous rates. This discrepancy arises from fundamental differences in enforcing “fair” access to the grid to contribute solar energy, compared to analogous fair sharing in networks and processors. To address the problem, we present fair rate allocation algorithms to enable control of distributed solar capacity, while enforcing fair grid energy access. Collectively, these algorithms represent a new mechanism for controlling the output of distributed solar capacity. We then present multiple policies that show how utilities can leverage this new distributed rate-limiting mechanism to reduce variations in grid demand from intermittent solar generation. We implement our algorithm and evaluate it on both synthetic data and real data from up to 50 solar sites. We show that traditional rate allocation, which enforces equal rates, results in solar sites contributing up to 18.9% less energy than an algorithm that enforces fair grid energy access over a single month. Finally, we show that our policies that vary solar capacity over time reduce the variations in aggregate solar power compared to a fixed limit, resulting in a more stable grid demand, while also preserving fairness.

REFERENCES

2017. PySolar. Retrieved from <http://pysolar.org/>.
- Analysis. 2017. Analysis: Current Net Metering Cap Stalls \$78 Million of Solar Projects in Massachusetts. Retrieved from <https://www.seia.org/news/analysis-current-net-metering-cap-stalls-78-million-solar-projects-massachusetts>.
- O. Ardakanian, C. Rosenberg, and S. Keshav. 2013. Distributed control of electric vehicle charging. In *Proceedings of the Fourth International Conference on Future Energy Systems (e-Energy'13)*.
- N. Bashir, D. Irwin, P. Shenoy, and J. Taneja. 2017. Enforcing fair grid energy access for controllable distributed solar capacity. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys'17)*.
- D. Cardwell. 2015. Solar Power Battle Puts Hawaii at Forefront of Worldwide Changes. *New York Times*, April 18, 2015.
- A. Demers, S. Keshav, and S. Shenker. 1989. Analysis and simulation of a fair queuing algorithm. In *Symposium Proceedings on Communications Architectures & Protocols (SIGCOMM'89)*.
- M. Durrenberger. 2015. STC vs. PTC: Why Solar Panel Testing Matters. Retrieved from <https://newenglandcleanenergy.com/energymiser/2015/12/01/stc-vs-ptc-why-solar-panel-testing-matters/>.
- J. Farmer and F. Lafond. 2016. How predictable is technological progress? *Research Policy* 45, 3 (2016), 647–665.
- J. Hill. 2017. GTM forecasting more than 85 gigawatts of solar PV to be installed in 2017. In *CleanTechnica*.
- D. Kempe, A. Dobra, and J. Gehrke. 2003. Gossip-based computation of aggregate information. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS'03)*.
- B. Kroposki. 2016. Can Smarter Solar Inverters Save the Grid? *IEEE Spectrum*, October 20, 2016.
- S. Lee, S. Iyengar, D. Irwin, and P. Shenoy. 2017. Distributed rate control for smart solar-powered systems. In *Proceedings of the Eighth International Conference on Future Energy Systems (e-Energy'17)*.
- D. Lew, L. Bird, M. Milligan, B. Speer, X. Wang, E. Carlini, A. Estanqueiro, D. Flynn, E. Gomez-Lazaro, N. Menemenlis, A. Orths, I. Pineda, J. Smith, L. Soder, P. Sorensen, A. Altiparmakis, and Y. Yoh. 2013. *Wind and Solar Curtailment*. Technical Report. National Renewable Energy Laboratory.
- Massachusetts. 2016. UtilityDive, Massachusetts Gov. Baker Signs Bill Lifting Solar Cap, Lowering Net Metering Rates. Retrieved from <https://www.utilitydive.com/news/massachusetts-gov-baker-signs-bill-lifting-solar-cap-lowering-net-meterin-1/417201/>.
- A. Mulkern. 2013. A Solar Boom So Successful, It's Been Halted. *Scientific American*, December 20, 2013.
- B. Raghavan, K. Vishwanath, S. Ramabhadran, K. Yocum, and A. Snoeren. 2007. Cloud control with distributed rate limiting. In *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM'07)*.

- S. Rongali, T. Ganuy, M. Padmanabhan, V. Arya, S. Kalyanaraman, and M. Petra. 2016. iPlug: Decentralised dispatch of distributed generation. In *International Conference on Communication Systems and Networks (COMSNETS'16)*.
- A. Singh, S. Lee, D. Irwin, and P. Shenoy. 2017. SunShade: Enabling software-defined solar-powered systems. In *ACM/IEEE 8th International Conference on Cyber-Physical Systems (ICCPs'17)*.
- R. Swanson. 2006. A vision for crystalline silicon photovoltaics. *Progress in Photovoltaics: Research and Applications* 14, 5 (August 2006), 443–453.
- R. Tonkoski and L. Lopes. 2011. Impact of active power curtailment on overvoltage prevention and energy production of PV inverters connected to low voltage residential feeders. *Renewable Energy* 36 (June 2011), 3566–3574.
- R. Tonkoski, L. Lopes, and T. El-Fouly. 2010. Droop-based active power curtailment for overvoltage prevention in grid connected PV inverters. In *IEEE International Symposium on Industrial Electronics (ISIE'10)*.
- R. Tonkoski, L. Lopes, and T. El-Fouly. 2011. Coordinated active power curtailment of grid connected PV inverters for overvoltage prevention. *IEEE Transactions on Sustainable Energy* 2, 2 (April 2011), 139–147.
- R. Tonkoski, L. Lopes, and D. Turcotte. 2009. Active power curtailment of PV inverters in diesel hybrid mini-grids. In *IEEE Electrical Power and Energy Conference*.

Received January 2018; accepted May 2018