SunDance: Black-box Behind-the-Meter Solar Disaggregation

Dong Chen and David Irwin
University of Massachusetts Amherst
{dongchen,deirwin}@umass.edu

ABSTRACT

The increasing penetration of grid-tied solar is complicating utilities' ability to balance electricity's real-time supply and demand. While advancements in solar forecast modeling are enabling utilities to better predict and react to future variations in solar power, these models require historical solar generation data for training. Unfortunately, pure solar generation data is often not available, as the vast majority of grid-tied solar deployments are "behind the meter," such that utilities only have access to net meter data that represents the sum of each building's solar generation and its energy consumption. To address the problem, we design SunDance, a "black box" technique for disaggregating solar generation from net meter data that requires only a building's location and a minimal amount of historical net meter data, e.g., as few as two datapoints.

SunDance leverages multiple insights into well-known fundamental relationships between location, weather, solar irradiance, and physical deployment characteristics to accurately disaggregate solar generation from net meter data *without access to a building's pure solar generation data for training*. We also identify and leverage a new fundamental relationship, which we call the Universal Weather-Solar Effect, that, to the best of our knowledge, has not been articulated in the past and is broadly applicable to other solar energy analytics. We evaluate SunDance using net meter data from 100 buildings and show that its black-box approach achieves similar accuracy without access to any solar training data as a fully supervised approach with complete access to such training data.

CCS CONCEPTS

• Computing methodologies → Model development and analysis; **Model verification and validation**;

KEYWORDS

Solar Disaggregation, Solar Modeling and Analysis

ACM Reference format:
Dong Chen and David Irwin. 2017. SunDance: Black-box Behind-the-Meter Solar Disaggregation. In *Proceedings of e-Energy '17, Shatin, Hong Kong, May 16-19, 2017,* 11 pages.
DOI: http://dx.doi.org/10.1145/3077839.3077848

978-1-4503-5036-5/17/05…$15.00
DOI: http://dx.doi.org/10.1145/3077839.3077848

1 INTRODUCTION

The aggregate solar capacity in the U.S. is growing rapidly, as the most recent estimates predict the U.S. solar market grew by 119% in 2016 alone [3]. Importantly, nearly all solar deployments are "grid-tied," such that they feed any solar power generated into the electric grid. Grid-tied deployments impose operational challenges on utilities in balancing electricity's real-time supply and demand. In particular, utilities plan generator "dispatch" schedules based on predictions of future load. Unfortunately, the increasing penetration of grid-tied solar is decreasing the accuracy of net load predictions. Solar power, even when aggregated, is more stochastic and less predictable than aggregate consumption largely because it depends on multiple factors that are specific to each site and highly localized.

Due to its increasing importance in grid operations, numerous prior and ongoing work focuses on accurately forecasting the grid's solar generation [4, 11–13, 16, 18, 24, 25]. While many of these forecast models offer coarse grid-level predictions of net load, recent work increasingly focuses on automatically generating customized forecast models via machine learning for each solar deployment based on its unique characteristics [13, 24]. These models can then be combined to generate a more accurate fine-grained grid-level forecast of solar generation and net load. Importantly, these custom solar forecasting techniques leverage supervised machine learning: they use a site's historical solar generation as training data to automatically learn a model that maps weather metrics to solar output at each time interval. The models then use standard forecasts of these weather metrics as input, e.g., from the National Weather Service (NWS), to predict future solar output.

Thus, the key to constructing sophisticated forecast models is access to historical solar generation data for training. Utilities are rapidly installing advanced or "smart" meters, which record energy flow at fine-grained intervals ranging from five minutes to every hour, that can provide such historical data. Smart meter installations are estimated to hit 70M by the start of 2017 and 90M by 2020 [14]. Thus, ample training data from smart meters is typically available for large residential solar deployments (>10kW) and solar farms, as these deployments are often required to be independently metered. As a result, these deployments' meter data represents pure solar data. However, nearly all small-scale residential solar deployments (<10kW) are "behind the meter" (BTM), such that the smart meter data exposed to utilities represents only the net of a building's solar generation and its energy consumption. Thus, constructing the forecast models above for BTM solar is not possible, as there is no pure solar data available for model training.

To address the problem, we present a new system, called SunDance, that accurately separates (or "disaggregates") a building's net meter data into its solar generation and energy consumption.[1]

[1]Note that solar disaggregation differs from energy disaggregation [5], as it only separates out solar generation from energy data and not appliance-specific consumption.

Importantly, SunDance employs a "black box" technique that requires *no training data* from the building itself, i.e., historical data separated into solar generation and energy usage, and instead only requires a minimal amount of net meter data and a location, both of which are available to utilities. In lieu of training data, SunDance leverages multiple insights into fundamental relationships between location, weather, physical characteristics, and solar irradiance. In particular, SunDance combines two key insights.

- **Clear Sky Generation Model**. Our first insight is that it is possible to build an accurate customized model of each solar deployment's maximum "clear sky" generation potential based on fundamental relationships between the Sun, the Earth, and a deployment's location and custom physical characteristics, *even when using noisy net meter data that combines solar generation with significant energy consumption.*
- **Universal Weather-Solar Effect**. Our second related insight is that the same weather conditions reduce the maximum clear sky solar irradiance potential by the same percentage regardless of the magnitude of this solar irradiance, which is a well-known weather-independent function of time at each location. This property, which we call the *Universal Weather-Solar Effect*, enables SunDance to i) build a general model using supervised machine learning that maps weather metrics to the expected fraction of the maximum solar irradiance potential for locations where solar training data is available, and then ii) apply that model to accurately infer solar generation at other locations, *where solar training data is not available.*

SunDance combines the insights above to develop an accurate customized model of a solar deployment's maximum solar generation potential using only its noisy net meter data and location, and then determines the fraction of this maximum generation the deployment actually produces by using a general model of weather's fundamental impact on the maximum solar irradiance potential. In developing SunDance, we make the following contributions.

Solar Background. We discuss in detail the fundamental physical relationships that govern solar generation over time based on location, position of the Sun, physical characteristics, weather, temperature, etc., and provide empirical evidence for SunDance's key insights above. These relationships dictate each location's unique solar signature, and the impact of weather on solar generation.

SunDance Design. We present SunDance's solar disaggregation technique summarized above. To construct a customized model of a solar deployment's maximum clear sky generation, SunDance searches for a valid solar signature that represents the tightest strict upper bound on the noisy net meter data. SunDance then learns a general model that captures the Universal Weather-Solar Effect at all location(s) where solar training is available. SunDance then combines these models to disaggregate a location's net meter data.

Implementation and Evaluation. We implement SunDance and evaluate it on net meter data from 100 buildings. We show that SunDance's accuracy, in terms of its Mean Absolute Percentage Error (MAPE), when inferring solar generation *without access to any solar training data from the buildings under test* is comparable to the accuracy of a customized machine learning model built with complete access to a building's historical solar data for training.

2 SOLAR BACKGROUND

SunDance assumes access to average power data $P_{net}(t)$ from a building smart meter, which represents the sum of solar power generation $P_s(t)$ and energy consumption $P_c(t)$, as shown below, where $P_s(t) \geq 0$ and $P_c(t) \leq 0$.

$$P_{net}(t) = P_s(t) + P_c(t), \forall t > 0 \quad (1)$$

Given only $P_{net}(t)$ and the meter's location, SunDance's task is to infer $P_s(t)$ and $P_c(t)$ at each time t.[2] Below, we provide a brief background on the fundamental relationships that determine i) the maximum amount of solar irradiance that reaches the Earth's surface at any time at any location, ii) the physical characteristics of solar cells that dictate how much of this irradiance is converted to electrical power under ideal weather conditions, and iii) the impact of non-ideal weather conditions. We identify insights based on these fundamental relationships in designing SunDance.

2.1 Computing Clear Sky Irradiance

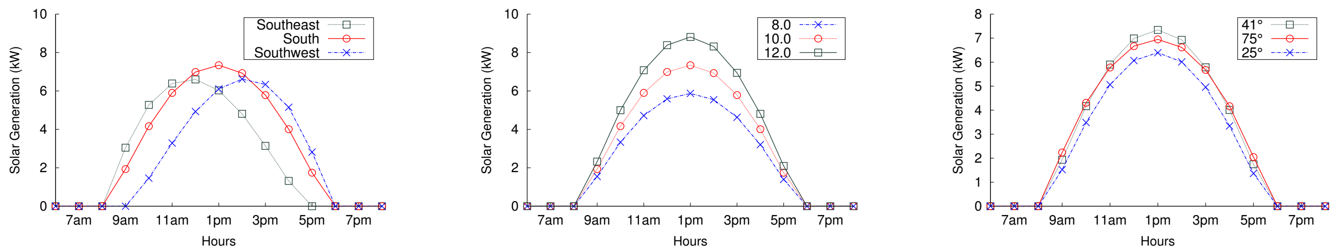
Solar irradiance is the power transmitted to the Earth by the Sun, and is measured in units of kilowatts per meter squared (kW/m^2). While the Total Solar Irradiance (TSI) that strikes perpendicular to the Earth's atmosphere is relatively constant and estimated at ~1.361 kW/m^2, the irradiance that reaches the ground is much less due to atmospheric losses (even under clear skies). The magnitude of these losses is largely a function of the Air Mass (AM) that light must travel through to reach the Earth, such that that the larger the AM the lower the fraction of TSI that reaches the ground. The AM is, in turn, a function of the Sun's position in the sky. For example, the fraction of TSI that reaches the ground is less closer to sunrise or sunset, as the Sun's light must pass through much more of the Earth's atmosphere at those times.

Since the Sun's position in the sky is a well-known function of location and time, it is possible to use the AM along with measurements of other atmospheric parameters to estimate the *clear sky irradiance* under a cloudless sky at any point on Earth at any time. There are many clear sky irradiance models that range from simple geometric formulas involving only the AM, the Sun's position, and experimentally-derived constants to highly complex models that require detailed data on the specific location's atmospheric conditions [19]. Note that evaluating the accuracy of these models is outside the scope of this paper, and has been the focus of significant prior work [19]. While SunDance is compatible with any of these models, our implementation in this paper uses a simple model that requires only a location's latitude and longitude and the Sun's position, which is a function of location and time. These models enable SunDance to directly estimate the clear sky solar irradiance I_{total}(t) at any location at any time t. Importantly, the actual power a solar module generates at any time *must always be strictly less* than a location's clear sky solar irradiance, as a module can never generate more power than the Sun provides.

2.2 Effect of Physical Characteristics

Solar cells harness the photovoltaic effect to translate the Sun's irradiance into electrical energy. However, the efficiency of solar cells depends on a variety of physical characteristics specific to

[2] Note that this time t also includes the day and month of the year.



(a) Orientation (k=10, tilt=41°) (b) Size and Efficiency (orientation=180°, tilt=41°) (c) Tilt (orientation=180°, k=10)

Figure 1: Maximum clear sky solar generation potential near NYC on 1/1/2016 for different physical deployment characteristics, including different orientations α (a), sizes and efficiencies k (b), and tilts ϕ(c).

each solar deployment. For example, the efficiency of commercial solar modules varies widely due to different materials and manufacturing processes, e.g., mono- versus poly-crystalline modules. In addition, a number of other physical characteristics further reduce solar module efficiency. The most important physical characteristics that affect efficiency are a solar module's size, tilt, and orientation. For example, the clear sky irradiance models above assume a 100% efficient solar module lying flat on the ground, such that its directional orientation and vertical tilt are equal to 0°. However, if a solar module is tilted upward and facing away from the Sun, not all of the available solar irradiance will reach it. As before, the effect of solar module size, tilt, and orientation are well-known and can be expressed using the closed-form equation below that relate a module's solar power generation P_s to the solar irradiance incident on the module $I_{incident}$ and the physical characteristics above.

$$P_s = I_{incident} * k * [\cos(90 - \Theta) * \sin(\beta) * \cos(\phi - \alpha) + \sin(90 - \Theta) * \cos(\beta)] \quad (2)$$

Here, Θ is the Sun's zenith angle above (such that 90-Θ is the Sun's elevation angle), α is the Sun's azimuth (or orientation) angle, β is the solar module's tilt angle, and ϕ is the solar module's azimuth (or orientation) angle. The Sun's zenith angle ranges from 0° (when the Sun is directly overhead) to 90° (at sunrise or sunset). Similarly, a solar module's tilt angle ranges from 0° when lying flat on the ground to 90° when vertical. The orientation angles for both the Sun and the module range from 0° (directly north) to 180° (directly south). Finally, the k parameter represents a combination of a solar module's size and its efficiency, expressed as a percentage of the incident solar irradiance $I_{incident}$ it converts to electrical energy. For example, a solar module that is 2× larger but half as efficient as another solar module would have the same value of k.

Figure 1 illustrates the physical effects on clear sky generation potential at a location just north of New York City, at 41° latitude and -74° longitude, on January 1st, 2016 for different solar module orientations (a), sizes and efficiencies (b), and tilts (c). As the figure shows, orienting the solar modules west or east shifts the peak solar generation later or earlier, respectively. In addition, since the k parameter from Equation 2 is a constant scaler it simply scales the curve up and down. The tilt parameter (β) has a similar effect as k, in that it also tends to scale the curve up and down for practical values, but is not a scaler, and thus also affects the orientation shift.

Thus, given k, β, and ϕ, we can compute a solar module's maximum power generation potential P_{smax} in clear skies at any location at any time by setting $I_{incident} = I_{total}$ from §2.1, as the other parameters are a function of location and time. In §3, we show how SunDance infers k, β, and ϕ from net meter data.

Other Effects. While module size, efficiency, tilt, and orientation have the largest impact on solar module output, other physical effects also exist that are not precisely modeled by the closed-form equation above. For example, a module's operating voltage affects its efficiency based on a solar module's IV curve. In this paper, we assume solar modules always operate at their maximum power point using standard tracking algorithms. In addition, while both i) multiple solar modules with different placements that are wired together (either in series or parallel) and ii) modules that track the Sun by changing their tilt and orientation also permit similar closed-form models, they are more complex. We focus use simple models, which apply to the vast majority of solar deployments, and leave extending them to more complex deployments as future work.

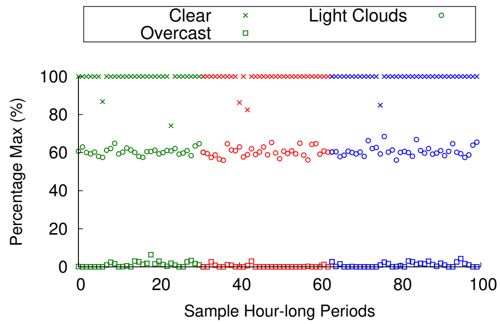
2.3 Weather Effects

The relationships above model the energy a solar module generates at any location at any time in ideal weather, e.g., under clear skies at an optimal temperature. Of course, non-ideal weather conditions can reduce both solar module efficiency and the amount of solar irradiance that reaches the ground. As we discuss below, the ambient temperature has a significant effect on solar module efficiency, while other weather conditions, such as clouds and humidity, affect the solar irradiance that reaches the ground.

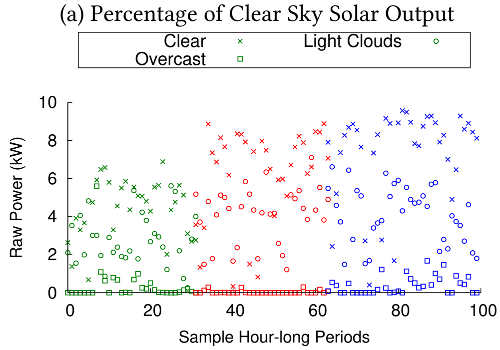
Efficiency Effects. While multiple weather metrics may affect solar cell efficiency, the most significant metric is the ambient temperature. The closed-form equation below estimates the cell temperature based on the temperature of the ambient air [23].

$$T_{cell} = T_{air} + S * \frac{NOCT - 20}{800} \quad (3)$$

Here, T_{cell} is the cell temperature in Celsius, T_{air} is the ambient air temperature in Celsius, S is the solar irradiance that is striking the panel (in W/m^2), and $NOCT$ is the Nominal Operating Cell Temperature. The NOCT varies between solar modules, but generally ranges from 33°C to 58°C with 48°C as a typical value. Importantly, for every degree increase (or decrease) in T_{cell}, the efficiency drops (or rises) by a constant percentage. While the precise temperature-based efficiency loss varies between modules, it is typically ∼0.5% per degree Celsius. Given these relationships, we discuss in the next section how SunDance adjusts for these temperature effects with only knowledge of the outdoor air temperature at a location, and without requiring a solar module's NOCT value or even knowing the precise effect of temperature on efficiency.



(a) Percentage of Clear Sky Solar Output



(b) Raw Solar Output

Figure 2: Percentage of maximum clear sky solar generation (a) and raw solar output (b) for multiple locations (indicated by 3 colors) and time periods with the same weather.

We do not model the effect of other weather metrics on efficiency, as these effects are typically not significant [20].

Irradiance Effects. A variety of other weather conditions serve to block some solar irradiance from reaching the module. Most importantly, the level of cloud cover, which can be quantified using satellite imagery, blocks solar irradiance. In addition, humidity, fog, and various forms of precipitation increase the particulates in the atmosphere (and on the module) that also block solar irradiance.

A key insight of our work, which we will leverage in the next section, is that the *exact same weather should have the same proportionate affect on the maximum solar irradiance potential* I_{total} *that reaches the ground, regardless of its magnitude,* which varies widely over time time at different locations. That is, if two different locations A and B experience the *exact same weather conditions* at two different times then the solar irradiance that reaches the ground $I_{incident}$ will be $c * I^A_{total}$ at location A and $c * I^B_{total}$ at location B, where c is a constant based on the weather and I_{total} is the maximum clear sky solar irradiance at those locations at those times. We call this the *Universal Weather-Solar Effect*, and, as we show, it is key to SunDance's approach.

While the Universal Weather-Solar Effect is intuitive, we provide some initial empirical evidence for it using solar and weather data from three solar deployments at different locations and times. A full evaluation of this effect is outside the scope of this paper. Our insight above implies that the same weather should yield the same percentage reduction in clear sky solar generation across all locations, regardless of time. Figure 2 shows results for three different weather conditions—clear, light clouds, and overcast—for three solar deployments (indicated by color) that >1000km apart. Here, "clear" corresponds to a dew point less than 6.1C, humidity less than 15%, no precipitation, and a "clear" sky condition; "light clouds" corresponds to a dew point of 7.3C-8.7C, a humidity between 40% and 55%, precipitation between 0 and 0.3cm, and a "mostly cloudy" sky condition; and "overcast" corresponds to a dew point greater than 16.7C, a humidity greater than 85%, precipitation greater than 0.3cm, and a "snow/rain/thunderstorm" sky condition.

At each location, we search for hour-long periods over multiple years that have precise weather conditions in the range above. We then plot the percentage of maximum solar generation at each hour under these weather conditions. Note that the sample time periods on the x-axis are random, non-contiguous, and drawn equally from the different locations, which each have widely different physical characteristics. Since we select sample periods based on similarity in their weather, they naturally cover different days of the year and times of day, which vary widely in their clear sky solar irradiance.

Despite these differences, Figure 2(a) demonstrates that the percentage of clear sky solar generation is remarkably constant for each weather condition across all times and all locations. In particular, our "clear" condition maps to near 100% clear sky generation, our "light clouds" condition maps to near 60% generation, and our "overcast" condition maps to near 0% generation. The slight variance in (a) occurs because our weather metrics are neither complete, i.e., we only use four weather metrics, nor precise, i.e., the sky condition reported by the NWS is a qualitative text string. In contrast, the raw solar output from these deployments across the sample periods, shown in Figure 2(b), is not constant, and varies widely based on weather, location, and time.

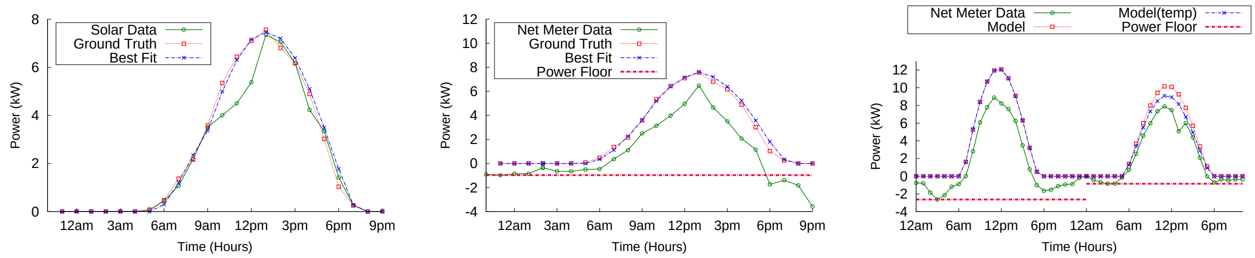
3 SUNDANCE DESIGN

Given only a solar-powered building's net energy meter data and its location, SunDance disaggregates the data into the two separate components in Equation 1: the building's solar generation $P_s(t)$ and its energy consumption $P_c(t)$. SunDance's design includes three key steps, which we summarize below, before detailing each.

1. Build a Custom Model of Maximum Solar Generation. SunDance uses historical net energy meter data to build a *custom model* of a solar deployment's maximum clear sky solar generation potential at any given time based on its location. This model incorporates each deployment's unique physical characteristics from the previous section, and its temperature effects, but focuses narrowly on modeling maximum generation potential and thus *does not* model any other weather-related effects, e.g., due to clouds, humidity, precipitation, etc. SunDance builds this model by finding the valid solar curve dictated by the fundamental relationships in the previous section that best "fits" the data.

2. Build a General Model of Weather's Effect on Irradiance. Separately, SunDance builds a general model that maps multiple weather metrics to the expected percentage reduction in clear sky solar irradiance potential. Due to the Universal Weather-Solar Effect, this model is general and can be built using solar training data from any (or multiple) locations, but then applied to accurately quantify the effect of weather on the clear sky solar irradiance potential at other locations, where such training data is not available.

3. Apply the Two Models Above to Disaggregate Solar Power. Given the two models above, disaggregating net energy meter data



(a) Modeling Pure Solar Data (b) Modeling Net Meter Data (c) Modeling using Minimal Data

Figure 3: SunDance's maximum clear sky generation model when built on pure solar data (a), net meter data (b), and on net meter data using historical data from only two days (c). In contrast, both (a) and (b) represent the best fit over one year of data.

is trivial. SunDance first uses weather data for the location as input to its general weather model to infer the percentage reduction in maximum clear sky solar irradiance potential. SunDance then applies this percentage reduction to the deployment's maximum solar generation, which is computed using the custom model in step one, to infer the absolute amount of solar generation $P_s(t)$ at each time t. Finally, to complete the disaggregation, SunDance subtracts this solar generation $P_s(t)$ from the net meter data $P_{net}(t)$ to yield the energy consumption $P_c(t)$ at the same time t.

3.1 Building a Maximum Generation Model

Inferring Physical Characteristics. The relationships in §2.1 and §2.2 enable us to define a range of valid solar curves at any location, which dictate the shape of maximum clear sky solar generation potential over each day of the year based on a deployment's physical characteristics, e.g., ϕ, k, and β from Equation 2.

SunDance builds a maximum clear sky generation model for a solar deployment by finding the ϕ, k, and β that defines the valid solar curve that best "fits" the location's energy data. We first discuss building this model for pure solar generation data and then describe how to translate it to net meter data that combines solar generation and energy consumption. Even pure solar generation data is stochastic, exhibiting many rapid variations in power due to changing weather conditions that diverge from its maximum power. For example, Figure 3(a) depicts solar generation on a partially cloudy day for a 10kW residential solar deployment, where output dips in the morning. Since generation deviates from its maximum in the morning, finding the valid solar curve that simply minimizes the Root Mean Squared Error (RMSE) with the data is not appropriate: the non-ideal weather will *always* result in fitting a solar curve that is lower than the maximum clear sky solar generation.

As a result, SunDance instead finds the best fit valid solar curve that represents the tightest upper bound on the data, since we know that the observed solar generation should never exceed the maximum clear sky generation. That is, among the valid solar curves that are equal to or greater than all datapoints, we find the one that minimizes the RMSE with the data. As a result, the curve SunDance finds will be dictated entirely by *the single datapoint that experiences the highest percentage of its maximum generation potential.* Thus, even if a day is cloudy, if there is even one datapoint that is near the maximum generation, this datapoint will dictate the best fit for the entire day (since the best fit must be a strict upper bound on the data). For example, even on the cloudy day in Figure 3(a), the best fit curve closely matches the ground truth maximum solar generation (which we approximate using the next day's solar generation under a clear sky), since it is dictated by the points in the day that are sunny. SunDance can apply this approach to multi-day time periods where the likelihood of a deployment experiencing its maximum generation at some point is high.

SunDance must search for the ϕ, k, and β from Equation 2 to find the best fit. This search is challenging since the parameters are dependent, e.g., modifying the tilt changes the effect of orientation, and conducting a brute force search across the entire parameter space is too computationally expensive, especially for fine-grained data. However, searching the entire parameter space appears necessary, as the tilt β and size and efficiency k have a similar effect on generation, which can lead to finding local maxima in isolated areas of the parameter space. For example, in one part of the parameter space, we may find a best fit curve that has a high tilt and low k, whereas the actual deployment has a high k and low tilt.

To address this problem, we observe that, while the physical characteristics of solar deployments are not always ideal, installers generally attempt to make them as ideal as possible. As a result, SunDance is able to accurately estimate a starting condition for its search based on the ideal physical characteristics to ensure it starts in the "right" region of the parameter space. In particular, the ideal orientation angle is south-facing in the northern hemisphere (and north-facing in the southern hemisphere) with a tilt angle equal to the latitude. Given these starting conditions, SunDance conducts an iterative search that first finds the value of k that best fits the data using a binary search, while keeping the other values constant. Given the new value of k, the search process then proceeds iteratively by next searching for the orientation that best fits the data using a binary search. After finding this orientation, we adjust the tilt in the same way. The search continues iteratively by adjusting each parameter in turn until they do not change significantly.

In practice, this approach efficiently finds tilt and orientation angles that are close to the ground truth tilt and orientation angles, since most solar deployments have physical characteristics near the ideal. For example, the ground truth tilt and orientation in Figure 3 are 35° and 190°, respectively, while the tilt and orientation angles SunDance finds in (a) when using pure solar data are 36° and 189°.

Modeling Temperature Effects. While the approach above finds a model with tilt and orientation angles that are close to the ground truth tilt and orientation angles, it is not accurate when applied over the entire year due to the effect of temperature on solar cell efficiency, which is captured by the k parameter. Since the best fit curve must be an upper bound on the data, this curve is dictated

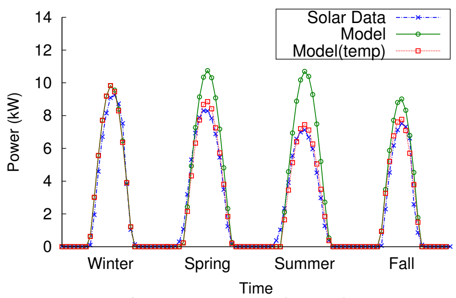


Figure 4: SunDance's maximum clear sky generation model both before and after adjusting for temperature effects.

by the point that achieves the highest percentage of its maximum clear sky generation potential at the lowest temperature, which is the most efficient operating point for the solar cell. Thus, to adjust for these temperature effects, SunDance finds the datapoint that is closest to the initial upper bound solar curve found above and then finds the location's ambient temperature at that time to use as a baseline $T_{baseline}$. This datapoint represents the coldest time period that maximizes solar cell efficiency under clear skies. SunDance then applies a temperature adjustment to k in the model.

$$k'(t) = k * (1 + c * (T_{baseline} - T_{air}(t))) \tag{4}$$

The temperature adjustment function reflects the constant factor c increase (or decrease) in efficiency when the ambient temperature is below or above the baseline temperature. Here, T_{air} is the location's ambient temperature at time t. While a typical value of c is 0.5% for solar modules [26], SunDance searches for the precise value of c for each deployment that represents the tightest upper bound on the data. While efficiency is a linear function of cell temperature, and not ambient temperature, since the temperature adjustment function subtracts the current temperature from the baseline it cancels out the constant values in Equation 3.

Figure 4 shows a maximum generation model for a sunny day in each season both before and after our temperature adjustment. Before the temperature adjustment, the model is highly accurate in January, since these cold weather days represent the most efficient operating points that dictate the upper bound, but is highly inaccurate in July when the temperature is 40C greater than on the coldest days. This is expected, since with a typical solar module, a 40C increase in temperature decreases the efficiency (and maximum generation potential) by $40 * 0.5\%$=20%. After the temperature adjustment, the maximum generation model closely matches the generation on these sunny days. In this case, the factor c we found was 0.57%, which is near the typical value of 0.5%.

Modeling using Net Meter Data. The discussion above builds a model of maximum clear sky solar generation using pure solar generation data. Modeling the maximum solar generation using net meter data differs in two key respects, which require simple extensions to our methodology above.

First, adding energy consumption introduces additional "consumption noise" to pure solar data that causes it to deviate from its maximum generation. Of course, this consumption-induced noise has the same effect as variable non-ideal weather conditions in decreasing the recorded generation. However, as discussed above, our modeling approach above is robust to non-ideal weather conditions, since the best fit must be a strict upper bound on the data, which

	Days	*Tilt*	*Orientation*	*k*	*Area (m^2)*	*c*
Ground Truth	NA	35°	190°	12.3	48.88	NA
Pure Solar	365	36°	189°	10.6	48.18	0.57
Net Meter	365	36°	188°	10.7	48.63	0.58
Net Meter (Temp)	365	34°	186°	10.9	49.55	0.72
Net Meter (Temp)	2	36°	185°	11.6	52.73	0.69

Table 1: The model parameters SunDance finds in the Figure 3 variants are all similar to the ground truth parameters.

is dictated the datapoint(s) that are closest to the maximum generation potential. This logic also applies to the non-ideal "weather" created by adding energy consumption: as long as datapoints exist where solar generation is near its maximum potential and energy consumption is low, e.g., when a home is unoccupied on a sunny day, our best fit upper bound model will be dictated by these few datapoints. This insight enables us to model a deployment's maximum solar generation even on noisy net meter data, where we cannot directly model the actual (disaggregated) solar generation.

Second, energy consumption in modern buildings generally never drops to zero. Thus, SunDance must estimate a building's minimum power consumption floor for the datapoint(s) above that dictate the model. To do so, SunDance simply uses the minimum power consumption at night, when solar generation is guaranteed to be zero. In many climates, the minimum power consumption occurs at night, while occupants are sleeping. However, in some cases, such as homes in cold climates that use electric heating, it is possible the minimum power consumption may not occur at night. In these cases, SunDance could use another approach to estimate the power consumption floor.

SunDance then subtracts the power consumption floor above from the data before constructing its model, where the minimum nightly consumption in the adjusted net meter data is zero. Figure 3(b) shows our model built using net meter data from the same deployment and time as in Figure 3(a). Note that, with (negative) consumption included, the net meter data is strictly less than the solar model; the figure also highlights the power floor SunDance uses to adjust the net meter data. Recall that the model in (a) finds a tilt of 36°, an orientation of 189°, a k of 10.6, and a c of 0.57%. Our model in (b) using the net meter data is similar, finding a tilt of 34°, an orientation of 186°, a k of 10.9, and a c of 0.72%.

Historical Data Requirements. SunDance requires remarkably little data to construct an accurate custom model of solar generation. In the limit, our approach needs only two datapoints during clear skies with low energy consumption, such that there is a significant temperature difference between the two points. Since the model finds the tightest upper bound on the available data, it is entirely dictated by the single point of maximum net generation (or, equivalently, the minimum net consumption). An additional point is needed at a different temperature to model the effect of temperature on efficiency. To illustrate, the models in Figures 3(a) and (b) were built by finding the tightest upper bound across an entire year of net meter data. In contrast, Figure 3(c) shows our model (with and without a temperature adjustment) using two sunny days in January on net meter data. Using only these two days, instead of an entire year, SunDance finds similar model parameters, with a tilt of 36°, an orientation of 185°, $k = 11.6$, and $c = 0.69\%$.

Table 1 summarizes the model parameters found on the different datasets and model variants in Figure 3. In all cases, the tilt, orientation, and c SunDance finds are close to the ground truth.

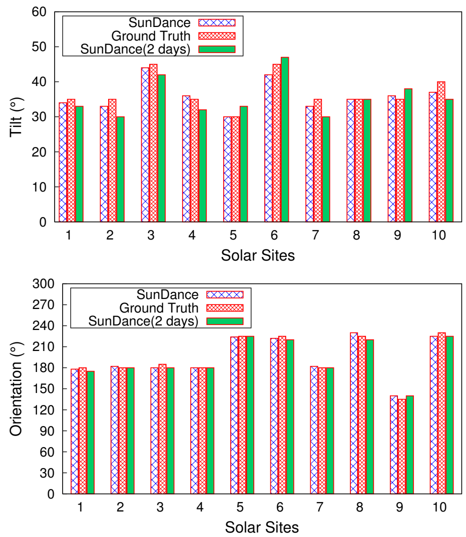


Figure 5: SunDance finds accurate tilt (top) and orientation (bottom) angles using as little as two days of net meter data.

In addition, the k value, which is the product of a module's size and efficiency at the baseline temperature, is also accurate. In this case, the module is 48.88m^2. While we cannot separate a module's size and efficiency, if we assume a typical commercial module with ~22% efficiency at 25C, SunDance can estimate the module size. As the table shows, SunDance finds sizes close to the ground truth for this module. While we evaluated SunDance across 100 buildings, we were only able to verify ground truth tilt and orientation angles of buildings that were clearly visible from Google street view data. Figure 5 shows that using noisy net meter data from 10 buildings where we could manually verify the ground truth tilt and orientation angles, SunDance consistently finds angles that are near the ground truth even when using only two days of data.

3.2 Building a General Weather Model

The models described above are highly customized to each deployment, incorporating its unique orientation, tilt, size, efficiency, and temperature effects. In contrast, our weather model, which leverages the Universal Weather-Solar Effect from §2.3, is general and thus applies to *any solar deployment*. As a result, our weather modeling is a one-time exercise that can use any solar irradiance (or solar power data) from any location. We construct our weather models similar to prior work on solar power forecasting models that use supervised machine learning [13, 24]. These approaches train a model based on labeled data that associates standard weather metrics, such as sky condition, temperature, humidity, dew point, precipitation, etc., with a deployment's solar generation.

Thus, the output of these existing models is a deployment's absolute solar generation, which is not general, but instead custom to each deployment's unique physical characteristics, particular its size. In contrast, SunDance generalizes these models by changing the output to be the fraction of maximum solar irradiance potential that reaches the ground. Based on the Universal Weather-Solar Effect, this approach can include solar irradiance data from *many* locations (and many times) to use for training a single general model. In addition, since pyranometer deployments, which record solar irradiance, are rare, SunDance can equivalently use any pure solar generation data (adjusted for temperature effects) that is available to build these models. Based on Equation 2, when dividing a deployment's solar output by its maximum solar generation potential, the factors based on the physical deployment characteristics cancel out, such that the resulting ratio is equivalent to the ratio of observed solar irradiance to maximum solar irradiance potential.

Importantly, our insight above means that *we can build a general weather model using pure solar power data from one (or many) deployments where it is available, and then use that model to accurately infer the reduction in solar power from its maximum potential at other solar deployments, where pure solar power data is not available.* This is a significant insight not only for our work on solar disaggregation, but also for work on solar forecasting based on pure solar data. For example, recent work highlights the importance of reducing the amount of training data necessary to build custom solar forecast models, especially for new solar deployments coming online [13]. However, based on the insight above, our general weather model *requires zero training data* from a new solar deployment under test. In addition, as discussed above, we can build an accurate maximum generation model using as few as two datapoints. In contrast, prior work requires from months [13] to years [24] of historical data to construct an accurate model.

Prior work has evaluated a wide range of supervised machine learning techniques for modeling the effect of weather on solar output, including least squares regression, Support Vector Machines (SVMs) using different kernel functions, and deep neural nets. While we evaluate different modeling techniques in §5, SunDance is orthogonal to the specific machine learning technique. SunDance's contribution instead lies in identifying the Universal Weather-Solar Effect and designing input and output features to leverage it to build a general weather model. We use the weather metrics from Weather Underground as input features to our model, including temperature, humidity, dew point, barometric pressure, precipitation, and sky condition. We map the qualitative descriptions for sky condition, e.g., scattered clouds, sunny, etc. to a numerical percentage of cloud cover using the mapping suggested by the NWS. More precise numerical percentages, which would improve model accuracy, can be derived from satellite imagery.

3.3 Disaggregating Net Meter Data

Given the two models above, solar disaggregation is trivial. For each datapoint in the net meter data, we use the weather metrics at that time as input to our general weather model above to infer the fraction of its maximum generation a solar deployment will output. To infer a building's actual solar generation $P_s(t)$, we then multiply this fraction by the maximum solar generation we infer based on our customized model in §3.1 at that time. We then simply subtract our inferred solar generation from the building's net meter data $P_{net}(t)$ to compute the corresponding energy consumption $P_c(t)$.

4 IMPLEMENTATION

We implement SunDance using a mixture of python and C++. We use simple well-known geometric formulas to compute a location's clear sky solar irradiance based on its latitude, longitude, elevation, time, and the Sun's position in the sky. To derive the Sun's position in the sky, we use the PSA algorithm, which takes as input the UTC time (to the second) and a location's latitude and longitude and outputs the Sun's precise azimuth and zenith angles [7]. High performance implementations of the PSA algorithm are publicly available that are accurate to within 0.0083° of the Sun's true position. We then compute the AM relative to an AM of 1 when the Sun is 90° overhead based on the well-known formula below.

$$AM = \frac{1}{\cos(\Theta) + 0.50572(96.07995 - \Theta)^{-1.6364}} \tag{5}$$

Given the AM above, we estimate the direct solar irradiance I_{direct} that reaches the ground using the Laue Model [17] as follows, where h is the location's elevation above sea level and 1.361 kW/m^2 represents the solar constant.

$$I_{direct} = 1.361 * [(1 - 0.14 * h)0.7^{AM^{0.678}} + 0.14 * h] \tag{6}$$

In addition, while variable, the amount of diffuse irradiance that is scattered by the atmosphere is generally estimated at ∼10% of the direct irradiance on a clear day. Thus, we compute the total solar irradiance I_{total} from §2.1 at any location as follows.

$$I_{total} = 1.1 * I_{direct} \tag{7}$$

Note that there are also packages available that implement other clear sky solar irradiance models, including PySolar [2] and NREL's library that implements the Bird model [1]. We leave evaluating SunDance's accuracy across these different models as future work.

Given a location's latitude and longitude, our implementation fetches historical weather data at one-hour granularity using Weather Underground's API. Since Weather Underground only has one-hour weather data, we can only disaggregate net meter data at the granularity of an hour. However, SunDance's approach is general and can be applied to weather and energy data at any granularity. We use the *scikit-learn* machine learning library in python to build our general weather model. The library supports multiple techniques including Support Vector Machines with different kernel functions and multiple linear regression models. We also use NumPy and Pandas for weather and energy data processing.

5 EVALUATION

We first evaluate SunDance's accuracy across 100 solar-powered buildings using one year of hour-level interval energy data. We then focus on a representative "net zero" building to understand the effect of energy consumption patterns, weather, and time on SunDance's accuracy. To quantify accuracy, we compute the Mean Absolute Percentage Error (MAPE), as follows, between the ground truth solar energy and the solar energy that SunDance infers over all time intervals t. A lower MAPE indicates higher accuracy with a 0% MAPE being perfectly accurate solar disaggregation.

$$MAPE = \frac{100}{n} \sum_{t=0}^{n} |\frac{S_t - P_t}{S_t}| \tag{8}$$

Here, S_t and P_t are the actual and inferred average solar power generation, respectively, over time t. We also compute the MAPE between the actual and inferred energy consumption using the same approach. We restrict all time periods to between sunrise and sunset, since SunDance is always perfectly accurate at night, as solar generation is always zero. Even so, MAPE is highly sensitive to periods of low absolute solar generation. For example, if sunrise falls near the end of an hour, the absolute generation of a 10kW solar deployment over the hour may only be 50W. If SunDance infers a generation of 100W, its MAPE for that period will be 100%. In contrast, the absolute generation during a cloudy mid-day period may be 5kW, such that if SunDance infers a generation of 6kW, its MAPE is only 20%. Thus, the absolute error of 50W contributes much more to the average MAPE than the absolute error of 1kW. To put our results in better context, we report overall MAPEs, as well as MAPEs for separate time periods and under different weather conditions. In particular, as in solar forecasting, we prioritize accuracy during cloudy periods in the middle parts of the day, where a significant amount of solar generation may fluctuate.

5.1 Comparing with a Supervised Approach

We compare SunDance's black-box approach to a fully supervised machine learning approach that has access to an entire year of historical solar generation and energy consumption data that has already been separated. In this case, the supervised approach works exactly like SunDance, except that, instead of our general weather model, we build a supervised model using the custom solar training data from each specific solar site. Thus, unlike other machine learning approaches [6, 13, 21, 24], our supervised approach incorporates the same physical solar models as SunDance. In recent work, we have show that this supervised approach is significantly more accurate than existing supervised approaches that do not incorporate physical solar models [9]. Our supervised approach represents a lower bound on the MAPE (and an upper bound on the accuracy) that SunDance can expect. As in prior work, we use a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel for our supervised approach [6, 21, 24]. SVM-RBF is common in solar modeling, since it attempts to fit a Gaussian curve to solar data and solar profiles are similar to Gaussian curves.

Figure 6 compares SunDance's accuracy with that of the supervised approach for each of the 100 buildings. In the graph, each stacked bar represents a building, such that the lower bar is the MAPE of a supervised approach, and the upper bar represents the increase in MAPE when using SunDance. The graph shows that across all buildings, the increase for SunDance in MAPE is generally small relative to the supervised approach. This result suggests the accuracy of Universal Weather-Solar Effect, as the only difference between the two approaches is data used for training.

In addition, the buildings are sorted by their ratio of solar energy generation to energy consumption, which is listed at the top of each bar. As the graph shows, the MAPE is partially a function of this ratio, such that a higher ratio generally yields more accurate results. This is intuitive, as increased energy consumption represents additional "noise" that SunDance must filter out. Note that a ratio of 100% represents a "net zero" building that has equal solar generation and energy consumption. Here, we see average MAPEs of ∼22% for SunDance on net zero buildings. Much of the imprecision derives from the absolute error in estimating each building's energy floor, which essentially requires an informed guess.

We also plot the same graph but only for the middle hours of the day (11am-3pm) to reduce the effect of small absolute errors that yield large percentage errors at the start and end of each day. This graph shows that MAPEs reduce by ∼22% during these important periods to an average of ∼17% for a net zero building. In addition, in

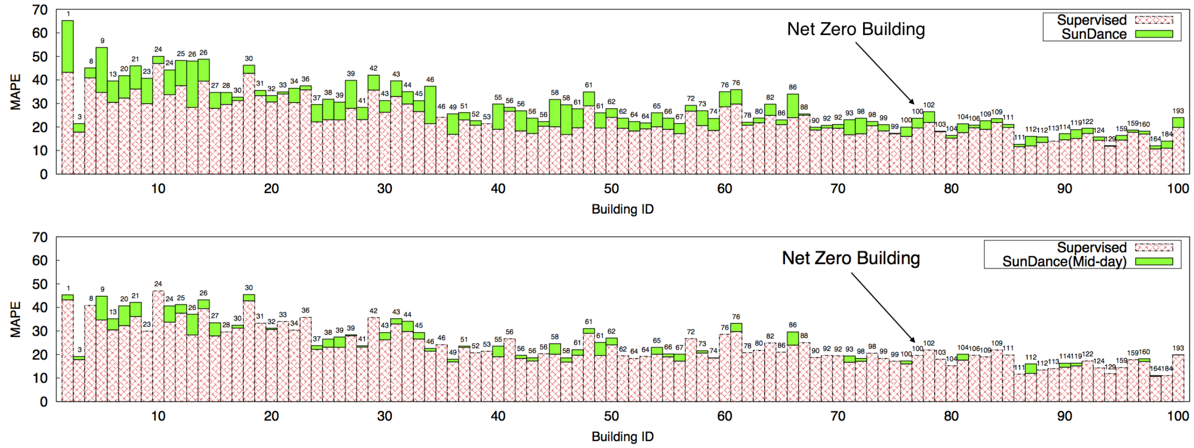


Figure 6: Daytime (top) and mid-day (bottom) MAPE for solar disaggregation using a supervised approach and SunDance for 100 buildings over a year. Buildings sorted by their ratio of solar generation to energy consumption (listed atop each bar).

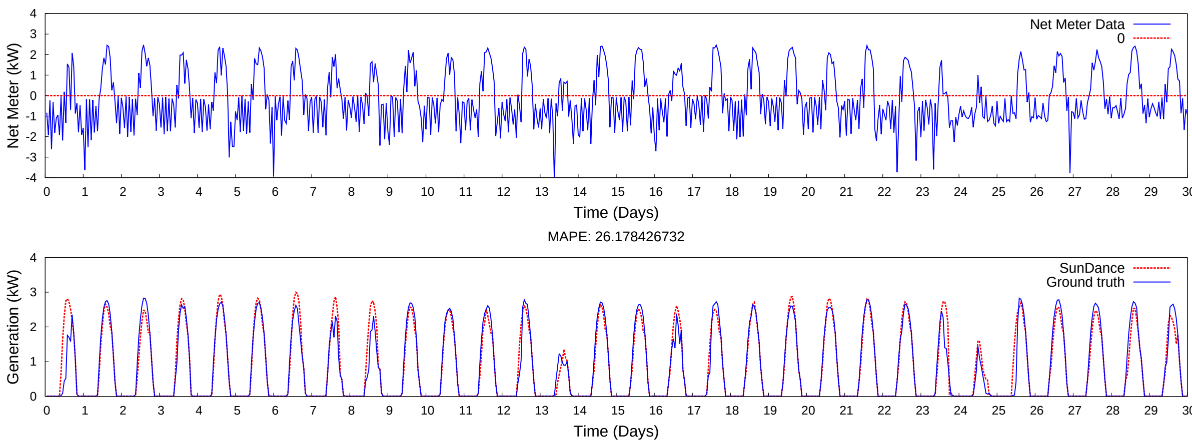


Figure 7: One month of net meter data (top) and ground truth and inferred solar generation (bottom) from a net zero building.

many cases for the mid-day results, SunDance performs as well as the supervised approach, in large part, because any small absolute error in the energy consumption floor has less effect on the MAPE as the absolute solar generation increases.

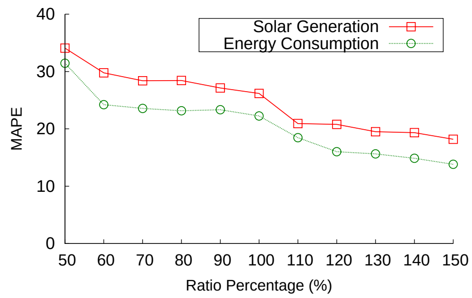
Result: *SunDance's black-box approach achieves similar accuracy without access to any solar training data from a site as a fully supervised approach with complete access to such training data.*

5.2 Quantifying SunDance's Accuracy

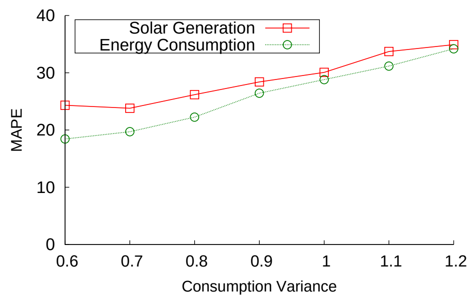
We next evaluate the different conditions that affect SunDance's accuracy on a representative net zero building (labeled in Figure 6). To provide a qualitative sense of SunDance's accuracy, Figure 7 shows the raw net meter data (top), as well as the ground truth and disaggregated solar generation (bottom). The figure shows that SunDance's inferred solar generation closely matches the ground truth solar generation, despite the stochasticity in the net meter data. In this case, the MAPE for solar generation is ~26%, while the MAPE for energy consumption is ~22%. The inferred energy consumption MAPE is typically lower because it is less affected by low absolute values, e.g., in the morning and evening.

We also examine the effect of changing both the ratio of solar generation to energy consumption and altering the variance of the energy consumption. In this case, to change the ratio, we alter the building's energy consumption at each time by a constant factor to increase and decrease the ratio. Similarly, we alter the variance by scaling the difference in energy consumption between two time periods by a constant factor, such that a value of 0 results in a completely flat consumption that never changes from the initial value. In both cases, the alterations produce a new set of net meter data, which we feed to SunDance for disaggregation. Figure 8 shows the results. As expected, as the ratio increases (a), and there is more solar generation to consumption, we see a linear decrease in MAPE (and corresponding increase in accuracy). Similarly, a low variance in consumption enables SunDance to more accurately model the solar generation and energy consumption floor (even if the ratio is large). Thus, in (b) we see a linear decrease in accuracy (and increase in MAPE) as the energy consumption variance increases.

We also break down our results based on weather conditions and time. Figure 9 breaks down accuracy based on weather conditions. In this case, we capture weather based on the percentage of the maximum generation a solar deployment is producing at any given time. Thus, if a solar deployment is only generating between 0% and 25% of its maximum clear sky potential, we assume that the weather is not good. Figure 9 shows that, as expected, our



(a) Generation:Consumption Ratio



(b) Energy Consumption Variance

Figure 8: Higher ratios of generation to consumption result in higher disaggregation accuracy (a). Lower variances in consumption result in higher disaggregation accuracy (b).

MAPE improves as the weather conditions improve. Importantly, for weather conditions that result in a ratio greater than 25%, SunDance yields near the same accuracy, indicating it performs well even under highly adverse weather conditions. While the MAPE is quite high during the worst weather conditions, this is largely due to small absolute errors from low generation that result in large percentage errors. To quantify this effect, we also plot the MAPE of energy consumption. Since this is a net zero building, we see that the small absolute errors in inferred solar generation during the worst weather conditions have little on effect on the energy consumption MAPE, which has similar accuracy across all weather.
Result: *SunDance accuracy is a linear function of the ratio of solar generation to energy usage and the variance in the energy usage. SunDance has the highest accuracy during the most critical period: adverse weather where solar generation is difficult to infer.*

6 RELATED WORK

The most similar work to SunDance is a recent approach for solar disaggregation (and product) from Bidgely, Inc. [21, 22]. Similar to SunDance, Bidgely trains a machine learning model (also using an SVM-RBF kernel) that maps weather metrics to normalized solar output on data from a set instrumented deployments, and then applies that model to estimate solar generation on a separate set of deployments. However, while SunDance normalizes solar output by constructing a maximum generation model based on the underlying physical characteristics of the deployment, Bidgely normalizes using a static value representing the maximum capacity of each deployment. As a result, Bidgely's model is significantly less accurate than SunDance's model, as we show in recent work [9]. Prior work also performs solar disaggregation by using data from microsynchrophasors at the feeder-level [15]. This approach differs in that it requires data from grid-level sensors.

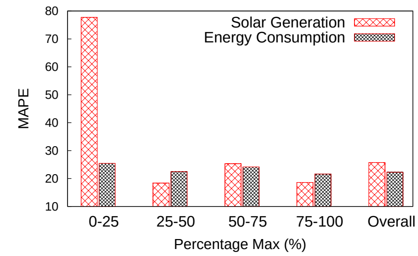


Figure 9: Solar generation and energy consumption MAPE during different weather conditions.

SunDance has many commonalities with existing solar forecast models based on machine learning [8, 13, 24]. However, these techniques use pure solar data to train their models, while SunDance focuses on disaggregating solar data from net meter data. SunDance's weather model uses a similar machine learning approach as prior forecasting techniques, except that its output is a fraction of the maximum clear sky generation, which varies over time at each location. This model is general due to the Universal Weather-Solar Effect. As a result, unlike prior forecasting approaches, SunDance requires *no training data* from the location under test to accurately model weather's effect on solar output.

Prior work on solar forecasting in SolarCast also performs feature engineering to reduce the amount training data necessary to build an accurate model that maps weather to solar output [13]. Similar to SunDance, SolarCast leverages the relationship between clear sky solar irradiance and solar output to build a single model of weather-to-solar output by normalizing its training data across time, e.g., by multiplying weather metrics by the clear sky irradiance. However, unlike SunDance, SolarCast's models are custom for each site, and not general, as their output is expressed in terms of raw solar power.

Finally, recent work shows how to extract the location where "anonymous" solar energy data was generated [10]. SunDance suggests the same approach can extract the location of anonymous net meter data that includes solar generation by first disaggregating the solar energy data. The potential to extract location from net meter data has serious privacy implications.

7 CONCLUSION

In this paper, we design SunDance, a new black-box technique for disaggregating BTM solar generation from net meter data. Importantly, SunDance requires only a deployment's location and a minimal amount of historical net meter data, e.g., as few as two datapoints. SunDance then leverages multiple insights into well-known fundamental relationships between location, weather, physical characteristics, and solar generation to build an accurate model of a deployment's solar generation. We implement SunDance and evaluate it on 100 buildings. Our evaluation shows that SunDance's black-box approach achieves similar accuracy without access to any solar training data from a deployment, as a fully supervised approach that has complete access to historical solar training data.
Acknowledgements. This research is supported by NSF grants IIP-1534080, CNS-1405826, CNS-1253063, CNS-1505422, and the Massachusetts Department of Energy Resources.

REFERENCES

[1] Bird Simple Spectral Model. http://rredc.nrel.gov/solar/models/spectral/. (????).
[2] PySolar. http://pysolar.org/. (????).
[3] 2016. Solar Energy Industries Association, Q4 2016 U.S. Solar Market Insight. http://www.seia.org/research-resources/us-solar-market-insight. (December 13th 2016).
[4] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez de Pison, and F. Antonanzas-Torres. 2016. Review of Photovoltaic Power Forecasting. *Elsevier Solar Energy* 136 (October 2016).
[5] K. Armel, A. Gupta, G. Shrimali, and A. Albert. 2013. Is Disaggregation the Holy Grail of Energy Efficiency? The Case of Electricity. *Energy Policy* 52, 1 (January 2013).
[6] M. Benghanem and A. Mellit. 2010. Radial Basis Function Network-based Prediction of Global Solar Radiation Data: Application for Sizing of a Stand-alone Photovoltaic System at Al-Madinah. *Energy* 35, 9 (2010).
[7] M. Blanco-Muriel, D.C. Alarcon-Padilla, T. Lopea-Moratalla, and M. Lara-Coira. 2001. Computing the Solar Vector. *Solar Energy* 70, 5 (August 2001).
[8] P. Chakraborty, M. Marwah, M. Arlitt, and N. Ramakrishnan. 2012. Fine-grained Photovoltaic Output Prediction using a Bayesian Ensemble. In *AAAI*.
[9] D. Chen and D. Irwin. 2017. Black-box Solar Performance Modeling: Comparing Physical, Machine Learning, and Hybrid Approaches. In *Greenmetrics*.
[10] D. Chen, S. Iyengar, D. Irwin, and P. Shenoy. 2016. SunSpot: Exposing the Location of Anonymous Solar-powered Homes. In *BuildSys*.
[11] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz. 2013. Review of Solar Irradiance Forecasting Methods and a Proposition for Small-scale Insular Grids. *Renewable and Sustainable Energy Reviews* 27 (November 2013).
[12] R. Inman, H. Pedro, and C. Coimbra. 2013. Solar Forecasting Methods for Renewable Energy Integration. *Elsevier Progress in Energy and Combustion Science* 39 (December 2013).
[13] S. Iyengar, N. Sharma, D. Irwin, P. Shenoy, and K. Ramamkritham. 2014. SolarCast - A Cloud-based Black Box Solar Predictor for Smart Homes. In *BuildSys*.
[14] J. John. 2016. US Smart Meter Deployments to Hit 70M in 2016, 90M in 2020. In *GreenTechMedia*.
[15] E. Kara, M. Tabone, C. Roberts, S. Kiliccote, and E. Stewart. 2016. Poster Abstract: Estimating Behind-the-meter Solar Generation with Existing Measurement Infrastructure. In *BuildSys*.
[16] A. Kaur, L. Nonnenmacher, and C. Coimbra. 2016. Net Load Forecasting for High Renewable Energy Penetration Grids. *Elsevier Energy* 114 (November 2016).
[17] E. G. Laue. 1970. The Measurement of Solar Spectral Irradiance at Different Terrestrial Elevations. *Solar Energy* 13 (1970).
[18] C. Martinez-Anido, B. Botor, A. Florita, C. Draxi, and S. Lu. 2016. The Value of Day-ahead Solar Power Forecasting Improvement. *Elsevier Solar Energy* 129 (May 2016).
[19] Joshua S. Stein Matthew J. Reno, Clifford W. Hansen. 2012. *Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis.* Technical Report. Sandia National Laboratories.
[20] S. Mekhilef, R. Saidur, and M. Kamalisarvestani. 2012. Effect of Dust, Humidity and Air Velocity on Efficiency of Photovoltaic Cells. *Renewable and Sustainable Energy Reviews* 16, 5 (2012).
[21] R. Mohan, T. Cheng, A. Gupta, V. Garud, and Y. He. 2014. Solar Energy Disaggregation using Whole-House Consumption Signals. In *NILM Workshop*.
[22] R. Mohan, C. Hsien-Teng, A. Gupta, Y. He, and V. Garud. 2015. *Bidgely, Inc., Solar Energy Disaggregation Techniques for Whole-house Energy Consumption Data.* Technical Report WO2015073996-A3. U.S. Patent Office.
[23] R. Ross. 1980. Flat-Plate Photovoltaic Array Design Optimization. In *IEEE Photovoltaic Specialists Conference*.
[24] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy. 2011. Predicting Solar Generation from Weather Forecasts Using Machine Learning. In *SmartGridComm*.
[25] C. Silva, L. Lim, D. Stevens, and D. Nakafuji. 2015. Probabilistic Models for One-Day Ahead Solar Irradiance Forecasting in Renewable Energy Applications. In *ICMLA*.
[26] E. Skoplaki and J. Palyvos. 2009. On the Temperature Dependence of Photovoltaic Module Electrical Performance: A Review of Efficiency/Power Correlations. *Solar Energy* 83, 5 (2009).