DeepSnow: Modeling the Impact of Snow on Solar Generation

Noman Bashir
University of Massachusetts Amherst
nbashir@umass.edu

David Irwin
University of Massachusetts Amherst
deirwin@umass.edu

Prashant Shenoy
University of Massachusetts Amherst
shenoy@cs.umass.edu

ABSTRACT

The falling cost of solar energy deployments has resulted in ever-increasing growth in solar capacity worldwide. The primary challenge posed by increasing grid-tied solar capacity is handling its variability due to continuously changing conditions. Thus, prior work has developed highly sophisticated models to estimate and forecast solar power output based on many characteristics, including location, elevation, time, weather, shading, module type, wiring, etc. These models are highly accurate for estimating solar power, especially over long periods, for sites at low latitudes, i.e., closer to the equator. However, models for sites at higher latitudes are less accurate due to the effect of snow on solar output, since even a small amount of snow can cover panels and reduce power to zero. Improving the accuracy of these models for annual solar output by even 2-3% is significant, as power translates directly into revenue, which compounds over the system's lifetime. Thus, if a site produces 2-3% less power on average per year due to snow than a model predicts, it can mean the difference between a positive or negative return-on-investment.

To address the problem, we develop DeepSnow, a data-driven approach that models the effect of snow on solar power generation. DeepSnow integrates with existing solar modeling frameworks, and uses publicly available snow data to learn its effect on solar generation. We leverage deep learning to quantify the effect of different snow variables on solar power using 4 million hourly readings from 40 solar sites. We evaluate our approach on 10 solar sites, and show that it yields a higher accuracy than the current approach for modeling snow effects used by the U.S. Department of Energy's System Advisor Model (SAM), a popular solar modeling framework.

CCS CONCEPTS

• **Applied computing → Forecasting**; • **General and reference → Performance**.

KEYWORDS

solar performance modeling, deep learning

ACM Reference Format:
Noman Bashir, David Irwin, and Prashant Shenoy. 2020. DeepSnow: Modeling the Impact of Snow on Solar Generation. In *The 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '20), November 18–20, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3408308.3427620

BuildSys '20, November 18–20, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-8061-4/20/11…$15.00
https://doi.org/10.1145/3408308.3427620

1 INTRODUCTION

Solar photovoltaic (PV) power is on track to become the largest source of electricity generation by 2050 given current trends. Solar PV has numerous benefits that other forms of energy cannot compete with: it has an increasingly low hardware cost, provides a limit-less energy supply without fuel, is passive with no moving parts and thus has low maintenance costs, and can be used effectively anywhere. In addition, the cost of solar PV is continuing to decrease as manufacturing volume and efficiency increases. Similarly, researchers are continuing to improve solar efficiency in multiple ways, e.g., by using multiple p-n junctions to overcome the Shockley-Queisser limit [18], using bi-facial panels, integrating solar PV and solar thermal, and making panels smaller and more flexible (as part of roof tiles) to increase their coverage. By contrast, conventional thermal generators are already highly cost- and energy-optimized, and are unlikely to experience significant further improvement.

The primary challenge with increasing solar capacity is its high variability due to changing environmental conditions. There are two ways to mitigate this challenge. One way is to install large-scale energy storage to store excess solar energy, and supply energy when there is a solar deficit. While there is significant active research in energy storage, the current offerings remain too expensive for grid-scale deployment, and improvements in energy storage's cost and efficiency are lagging improvements in solar. Thus, solar capacity will likely increase at a much faster rate than energy storage capacity in the near-term future. Without sufficient energy storage, utilities will need to be able to accurately model and predict sites' solar output to enable better advance planning of the energy supply in the grid. Models of current and future solar output are also useful for solar operators in estimating their future year-over-year revenue, valuing solar sites, and potentially participating in electricity markets, e.g., by committing solar power in the day-ahead market.

Thus, solar performance modeling, which estimates a site's current or future solar output based on a site's characteristics and environment, is a highly active research area. There are multiple open-source and proprietary models available. For example, the U.S. Department of Energy (DOE) has at least two open-source solar modeling frameworks: pvlib [11] and the System Advisor Model (SAM) [9]. These frameworks enable users to virtually configure detailed solar systems, and then provides an estimate of solar output over time. Other approaches, such as Solar-TK [6], develop models directly from data, rather than requiring users to know the details of their system to virtually configure it. These modeling frameworks all use the same basic set of physical models that determine solar power output, including a site's location, elevation, time-of-day, day-of-year, panel efficiency, temperature coefficient, wiring topology, inverter efficiency, panel tilt and orientation, shading effects, cloud cover and global horizontal irradiance (GHI), etc.

The modeling frameworks above are generally highly accurate after configuring or calibrating them for a particular solar site. Unfortunately, this accuracy significantly degrades for sites at higher latitudes due to the effect of snow on solar output. While snow effects only occur during winter, they are large enough to degrade the accuracy of yearly estimates of solar output. Even a small amount of snow (less than 1 inch) can completely cover panels, and reduce power output to zero for extended periods of time. Solar output at higher latitudes in winter months tends to also be highly volatile, and almost completely based on snowfall. For example, at one site we monitor in the Northeast U.S., we recorded a monthly solar output in March, 2018 of 256kWh, and in March, 2019 of 774 kWh. The 302% increase in monthly solar output was due almost exclusively to snow cover, with 2018 having much more snow in March than 2019. This snowfall had a significant impact on the annual solar output as well. For the same site, the 2018 annual output was 8,141kWh, while the 2019 annual output was 9,091kWh, or a difference of 950kWh.

Importantly, at this site, the difference in solar output in the winter months (January, February, March) in 2018 and 2019 accounted for 617kWh (or 65%) of the difference in output over the year, even though the winter months only accounted for 8% (2018) and 14% (2019) of the yearly output. Thus, for high latitude solar sites, the variability in solar output is disproportionately affected by snowfall in the winter. As a result, accurately modeling the effect of snowfall is important both at short-term intervals, e.g., every hour, and over long-term intervals, e.g., over months and years. At short-term intervals as snow melts, solar output is determined by the rate of snow melt as well physical features, such as a panel's tilt angle. As our example shows, accurately modeling snow over long-term intervals is also important: improving accuracy by even a small amount (2-3%) can be significant, since power translates to monetary revenue, which compounds over time. Thus, a 2-3% difference when compounded over a 25-year lifetime is significant and can make the difference between a positive and negative return-on-investment (ROI).

To address the problem, we develop DeepSnow, a data-driven approach to modeling snow effects on solar power. Our hypothesis is that DeepSnow can significantly improve the accuracy of existing solar modeling frameworks for higher latitudes with snowfall. In evaluating our hypothesis, we make the following contributions.

Snow Data Analysis. We analyze public snow data to learn the snow variables that most affect solar generation. Snow is highly complex, and public snow data sites record over 40 different variables that describe snow. We analyze this dataset and identify the variables that most correlate with solar output. In total, our analysis covers 4 million hourly snow readings from 40 solar sites.

Deep Learning Approach. We combine the snow data with other data, such as the shading level, tilt angle, and temperature, to learn the effect of snow on solar output. Our approach builds on other solar modeling frameworks by augmenting them to account for the effect of snow. When snow is not present, our approach simply devolves to the current state-of-the-art. Thus, our model distills only the effects of snow on solar output and does not conflate it with the effect of other variables. We learn a common model for snow that can be applied to any new or existing solar site without re-training.

Implementation and Evaluation. We integrate our approach into an existing solar modeling framework [6], and evaluate its performance on 10 solar sites. We compare it to the same modeling framework without considering snow, and to DOE's SAM model, which includes a physical snow model. We show that our approach significantly outperforms the SAM model, consistently yielding better accuracy across all 10 sites by up to 7% over a year, and much more over the winter months and at shorter time intervals. We also show that DeepSnow is more accurate than SAM even over the summer.

2 BACKGROUND

We briefly summarize the impact of snow on solar power and discuss physical models that describe the relationship between different snow properties and solar power.

2.1 Solar Performance Models

There has been significant prior work on solar performance modeling and forecasting. Solar performance modeling infers a site's solar power output at the current time based on known conditions, while solar forecasting predicts a site's solar output at a future time based on forecasted conditions. While there has been hundreds of papers on this topic, there are a limited number of toolkits using these models that are open-source and available for public use. Below, we describe the three most prominent solar performance modeling tools. While we discuss these approaches, our main focus is on describing how these toolkits incorporate snow effects on solar panel output.

System Advisor Model (SAM). The System Advisor Model (SAM) is a financial and performance model that estimates the cost of energy for grid-connected power projects [7]. SAM contains performance models for a variety of energy resources and their relevant financial models, but we focus only the performance model for photovoltaic (PV) systems. SAM's performance model requires users to provide detailed information about the site, i.e., type of PV module used, wiring of panels, inverter model, site's tilt, and orientation. Researchers may not have access to such detailed information when performing analyses on a large number of solar sites. In addition, SAM takes the weather data in the Typical Meteorological Year (TMY) format and provides support to download TMY weather files for locations across the United States. However, if users must evaluate the performance of a site for a given year, they must go through a non-trivial process of constructing TMY files for the desired location and year. Therefore, SAM is more suited for a one-time feasibility analysis of a PV site. Despite SAM's configuration challenges, it provides a good baseline for our analysis, as we expect DeepSnow to be used for energy estimates of solar sites similar to SAM.

An attractive feature of SAM's performance model is that it takes into account the energy loss due to snow. The snow model used by SAM is based on prior work that calculates the percentage of a PV array that will be covered by snow given the daily snow depth measurement, plane of array irradiance, temperature, and the tilt angle for the site [13]. This model considers the snow sliding off the surface of solar panels as the dominant snow removal process. However, the model does not take into account snow melting due to temperature, or the wind blowing light snow from the panel surface. SAM's snow model uses daily snow depth values. At the beginning of the day, the model will check if there is a snow event, and, if the snow depth is greater than 0, the model assumes that all solar panels on the site are covered by the snow. At each hour, the model will then check if the following condition is met.

$$T_a > \frac{I_{poa}}{m}$$

Here, T_a is the ambient temperature, I_{poa} is the plane of array irradiance, and m is an empirically determined constant with a value of -80$\frac{W}{m^{2\circ}C}$. This inequality checks if the ambient temperature and plane of array irradiance are high enough to allow some of the snow to slide off the array surface. If the given inequality is satisfied for a particular hour, some portion of the PV array will be exposed to direct sunlight. The amount of exposed surface is intuitively dependent upon the tilt angle for the site and is computed as follows.

$$Snow\ Slide\ Amount = 1.97 \times sin(tilt)$$

Here, 1.97 is the sliding coefficient determined empirically by the authors of the snow model [13]. The Snow Slide Amount computed from the previous equation is given as multiple of 1/10th of the array's height. For example, a value of 2 for the Snow Slide Amount means that 2/10th of the array is exposed to sunlight. Note that the model assumes that the snow completely slides off the panels and does not accumulate near the bottom of the panel. If the snow slides off in a given hour, the model determines the fraction of solar panels exposed at the end of the hour. Solar output for a corresponding number of panels is set to normal, while the covered panels' output is set to 0. The model makes a few simplistic assumptions, such as snow removal occurring only because of sliding, and snow completely sliding off the panels and not accumulating near the bottom, and the power output of the covered portion of the panels being 0. As we show, these assumptions significantly reduce the accuracy of this snow model compared to DeepSnow, which implicitly models these phenomenon by learning them from data.

PVlib. PVlib is a python library for modeling solar energy systems that, compared to SAM, focuses more on programmability and less on financial estimates [11]. PVlib's model library implements multiple physical models for various factors related to solar performance, i.e., pvlib has multiple models for solar irradiance to power conversion. The underlying models used by SAM are also available in PVlib. Like SAM, PVlib also requires the user to specify detailed information about a solar site to model it. PVlib uses the same snow model as SAM. Due to these overlapping characteristics, we do not consider PVlib for further analysis in this paper.

Solar-TK. Solar-TK is a black-box solar performance model that is open-source and publicly available [6]. This model calibrates the model parameters entirely from a small amount of historical generation data from a solar site. Once calibrated, the model only requires as input a site's location, time, and weather (cloud cover and temperature) over some time resolution, and returns as output an estimate of solar energy over that time resolution. Solar-TK also has an optional shade-adjustment module that uses machine learning (ML) to learn the effect of shading from nearby buildings and trees on solar output. The shade-adjustment module provides its output as a fraction between 0 and 1: a value of 0 indicates that power has dropped to zero because of shade and a value of 1 indicates that shading has no effect on the site. Such shading has an effect on snow melting, as shaded snow melts slower than the non-shaded snow. Therefore, we use the shade factor provided by Solar-TK as one of the features of DeepSnow's model. However, Solar-TK does not model the effect of snow on solar power output. Since Solar-TK is modular and extensible, we integrate a DeepSnow module that augments Solar-TK's existing framework to model snow.

2.2 Solar Performance Modeling under Snow

The key effect of snow on the solar sites is in the form of blocking the solar irradiation from reaching the panel surface. The importance of this can be assessed and even quantified by utilizing the physical models for the optical properties of snow. There are a large number of factors that affect these optical properties, such as the average snow flake/grain size, free water content, or the formation and density of the layers within the snowpack [10].

The most important property of the snow is its reflectance, which is also often called its *albedo*, that describes the ratio of reflected radiation from a surface to radiation incident on it. The *albedo* depends upon the snow depth for the thin snow layers, and prior work has shown that it will reach its maximum value when the snow depth is around 4cm [14]. The reflectance is not the only phenomenon that reduces the solar irradiation reaching the panel surface, as snow absorbs some of the radiation as well. However, the absorption effect is less significant when compared with the reflectance. As the snow melts, the transmission of solar irradiation to the surface may not increase as the increased snow water content increases the absorption of radiation in the snow water [17]. While there is a consensus that the combined effect of reflectance and absorption means that even a small amount of snow will greatly reduce the solar generation, the depth of snow needed to block 99% of light varies from 2-74cm in prior work [12, 17].

Researchers have established that the snow depth along with snow water is not enough to describe the effect of snow on solar irradiation. The top layer of snow has different optical properties than the snow in the deeper layers of the snowpack, which are compressed by the weight of upper snow layers. At snow depths lower than the thickness of the top layer, the optical properties of the snow pack will be influenced by its underlying surface. This top layer has a higher extinction coefficient than the deeper snow, resulting in a more rapid extinction of solar radiation passing through it [12, 14]. Further, the extinction coefficient of a typical snowpack, within the visible spectrum, varies across different wavelengths of radiation. At longer wavelengths, the solar radiation penetrating the snow layer is of a comparatively low wavelengths with the highest transmittance occurring at wavelengths between 450 and 550nm. The extinction coefficient of snow appears to be the lowest between 400 and 700nm, increasing sharply for wavelengths above 700nm. The extinction coefficient is somewhat lower for older snow than for new fluffy snow, and significantly higher for very wet, melting snow [14, 17].

The primary point of our discussion above is that the impact of snow on solar power is highly complex, and not a straightforward phenomenon. The snow depth alone is not enough to quantify the effect of snow and its additional variables, such as snow water amount, snow layers, density of different layers, snow layer temperature, humidity, windspeed, and types of incident irradiation. Understanding these additional variables is necessary to fully understand the impact of snow. While the effect of these different parameters is well-understood individually, there is no well-known physical model that incorporates the effect of snow on the solar irradiation passing

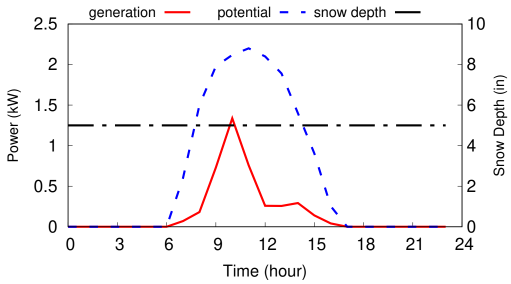


Figure 1: Solar power output and snow depth for an example home on a sunny day following a snow event.

through it. These effects are likely too numerous and complex to physically model. However, the complex relationships between different snow optical and other properties makes this an ideal problem for applying deep learning.

2.3 Snow Measurement Data

The National Operational Hydrologic Remote Sensing Center (NOHRSC) provides observed and modeled snow data for locations across United States [3]. NOHSRC's observed data comes from ground-based, satellite, and airborne observations from all the snow stations across U.S. In addition to the observed values, NOHSRC also runs a snow model that uses the observed data along with estimates of snow pack characteristics to generate NOAA's National Snow Analyses (NSA) on daily basis. The NOHRSC snow model has $1km^2$ spatial resolution and hourly temporal resolution.

The NOHRSC provides information about the snow water equivalent, snow depth, snow pack temperatures, snow sublimation, snow evaporation, estimates of blowing snow, modeled and observed snow information, airborne snow data, satellite, snow cover, and others. The snow data is available through the SNOw Data Assimilation System (SNODAS), which is a modeling and data collection system developed by NOHSRC. SNODAS provides observed and modeled data for all the variables above for $1km^2$ spatial and daily temporal resolution. We used NOHSRC's interactive maps to retrieve the hourly resolution data for the sites we evaluate [4].

2.4 Problem Statement

Our goal is to develop a solar performance model that uses site-independent parameters, such as snow depth, precipitation, as well as site-specific parameters, such as tilt, orientation and shading, to capture the effect of snow to predict the solar output of a solar site during the winter months. We intend our model to integrate with an existing data-driven modeling framework, particularly Solar-TK [6], that is able to accurately estimate the site-specific parameters, e.g., tilt, orientation, shading, etc., and thus we focus solely on snow modeling in this paper. Our evaluation incorporates DeepSnow as an additional independent module into Solar-TK that adjusts its models for the presence of snow.

3 DEEPSNOW DESIGN

We present DeepSnow's design using a ML-based approach that integrates into Solar-TK to model snow's effect on solar output.

Figure 2: Flatter surface leads to more snow accumulation and longer snow melt time.

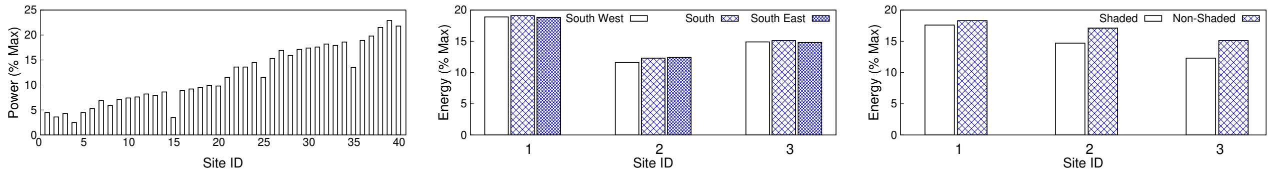
3.1 Solar Performance Under Snow

Solar power output under snow depends on snow's optical properties that determine how much sunlight reaches the solar panel surface. Snow's optical properties are dependent on factors such as snow flake size, free water content, formation of snow layers, and snow density. Physical models exist that model the impact of one or a few parameters on the optical properties of snow, and ultimately solar power. However, there is no single physical model that takes into account all of these snow properties and describes the relationship between snow properties and solar power. The problem is further complicated as publicly-available snow observation data is collected at weather stations, dozens of miles away from solar sites in most cases. The snow observations are typically made on a flat ground surface, a flat wooden board, or on the grass. These observations do not necessarily represent the snow condition at nearby solar sites. Solar panels are installed at a certain tilt angle and have different surface friction than ground, which results in snow sliding. This results in actual snow on the panel surface being different from the snow observation at the weather station.

Figure 1 illustrates how a solar site can produce power even when the observed snow depth at a nearby weather station is greater than 2 inches (5.08cm). The solar site produces >12% of its potential power at all times and 27% of its potential energy across the day. This behavior indicates that the snow on the solar panel surface has not completely melted to allow unrestricted generation typical for a sunny day. However, this also indicates that snow is not completely covering the panel surface and the site is able to produce some power from partially covered panels. There are many solar site properties that affect how much solar power is generated under snow. We outline the different properties below.

A. Tilt. Figure 2 illustrates that the amount of snow on solar panels, installed at a tilt angle, can be different from the snow observed at ground level. Snow slides down a panel installed at a higher tilt angle easily as compared to a flatter panel. However, note that the panels at the same tilt angle may not have the same amount of snow due to different surface frictions, snow accumulated at downstream panels, and wind direction. Figure 3a shows the effect of tilt on the amount of power produced by solar panels when the snow depth is >2 inches. Sites are arranged in the order of increasing tilt, where the left most site has a tilt of 15° and right most has a 60° tilt. A higher tilt generally yields more energy following a snow event.

B. Orientation. Solar panel orientation also affects power generation under snow. Figure 3b shows the amount of energy produced by solar sites with different orientations at three different locations. A south-facing panel produces more energy as it receives sunlight throughout the day, which melts the snow on the solar panel faster.



(a) Tilt Effect (b) Orientation Effect (c) Shading Effect

Figure 3: Effect of site properties on how snow on a solar panel melts: (a) tilt effect, (b) orientation effect, and (c) shading effect.

Variable Name	Data Type	Reporting Percentage (%)
Snow Depth	Numerical Model	94.85
Snow Depth (Top)	Numerical Model	93.10
Snow Depth (Middle)	Numerical Model	93.24
Snow Depth (Bottom)	Numerical Model	93.24
Snow Water Equivalent	Numerical Model	94.85
Snow Precipitation	Numerical Model	94.99
Non-Snow Precipitation	Numerical Model	94.99
Cumulative Snow Precipitation	Numerical Model	94.99
Cumulative Non-Snow Precipitation	Numerical Model	94.99
Cumulative Total Precipitation	Numerical Model	94.99
Air Temperature	Numerical Model	94.99
Relative Humidity	Numerical Model	94.99
Wind Speed	Numerical Model	94.97
Air Temperature	Observed	91.18
Dew Point Temperature	Observed	91.26
Wind Speed	Observed	91.42

Table 1: Reporting percentage and data source of variables in NOHSRC hourly data. Note that we discard variables that are reported less than 90% of the time.

In contrast, a south-west facing panel receives higher sunlight only in the evening period. These observations hold for all three locations.

C. Shading. Similar to orientation, shading impacts how much direct sunlight falls on the snow on a particular surface. Shaded areas are also cooler than the places that are directly under sunlight. Nearby trees and structures may partially shade the solar panels at least for a part of the day. This shade results in lower snow melt in shaded areas than the ones that are fully exposed. In the case of severe shading, the snow on the ground can fully melt, while the panels still are partially shaded with snow, leading to a discrepancy in snow observation data and ground truth. Figure 3c illustrates how the power generated by two sites, with the same tilt, orientation, and snow depth, can differ due to shading. We present this observation for three different locations. In this case, site 1 has trees on its west-side that block sunlight near the end of the day. As a result, site 2 has a higher energy output than site 1.

We have discussed the site-specific properties that affect how snow on the solar panel surface melts. These effects result in the snow on solar panels differing from the observations taken on a flat surface at a nearby weather station.

3.2 Snow Data Analysis

The National Operational Hydrologic Remote Sensing Center (NOHRSC) stores archives of historical snow data for every location in the United States for the past decade. The dataset provides over 40 variables related to snow. The data reports observed and modeled values for some variables, such as snow density and snow depth. In this case, the modeled variables are not directly measured, but derive from applying numerical weather models to observed data to infer their values. We use the observed values for the variable when available and replace the missing values with the modeled data. However, even after this preprocessing, the data contains numerous missing values and not all the variables are reported regularly. Table 1 shows that the reporting percentage varies significantly across variables. The variables with low reporting percentage (<90%) are discarded from further analysis.

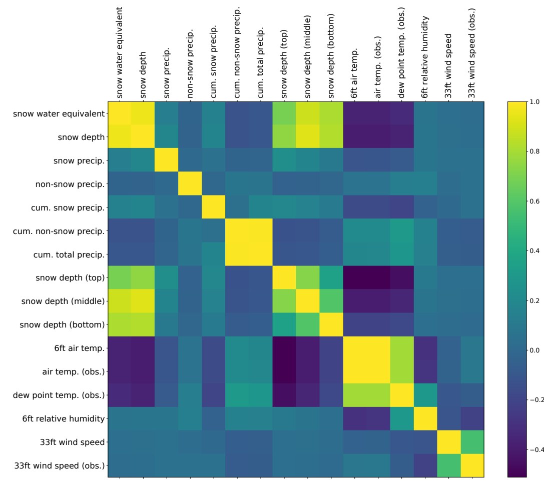


Figure 4: Coorelation Matrix for variables reporting at >90%.

3.3 DeepSnow Modeling Approach

The previous sections outlined site-specific features as well as the snow related measurements from the publicly available datasets. In the next step, we identify additional sources of features that can potentially increase the accuracy of our modeling approach. After data cleaning, there are 16 variables with a reporting percentage of 90% or greater. Some of the variables quantify distinct properties of snow, i.e., snow depth and wind speed, and thus provide unique information. However, other variables are highly correlated because they provide the same or similar property using a different metric, i.e., snow depth and cumulative depth of snowfall. Figure 4 shows that there are some variables that are highly correlated. We keep only one of the variable pairs with correlation coefficient greater than 0.8. This feature reduction step removes 9 of the features.

A. Feature Engineering A distinctive feature of DeepSnow's modeling approach is that it incorporates data from multiple sources with each using a different modeling or data collection approach. In this section, we elaborate on the source for the site-specific and snow data features. We also describe additional features and their sources.

Feature	Data Type	Unit
Snow Depth (Hourly)	Numerical Model	inches
Cumulative Snow Precipitation	Numerical Model	%
Cumulative Total Precipitation	Numerical Model	%
Relative Humidity	Numerical Model	%
Snow Loss Factor (SAM)	Physical Model	%
Shading Factor (Solar-TK)	ML Model	%
Orientation (Solar-TK)	Calibrated Model	°
Tilt Angle (Solar-TK)	Calibrated Model	°
Air Temperature	Observed	C
Dew Point Temperature	Observed	C
Snow Depth (Daily)	Observed	inches
Wind Speed	Observed	mph

Table 2: Final set of features used for model training.

Tilt and Orientation. The site-specific parameters in our feature set are generated using a calibrated model from Solar-TK. Calibrated models fit the well-known physical models to the ground truth observations to determine the parameters or coefficients of the physical model. For example, Solar-TK's physical parameter module uses a physical model that describes the relationship between the solar panel output and its capacity, tilt, orientation, and temperature coefficients. Solar-TK calibrates the model to ground-truth data to find these physical parameters for each site. Since we do not have physical access to the solar sites, we use Solar-TK's calibration model to estimate the tilt and orientation of sites.

Snow Observations. The simplest source of data comes from the actual measurements of different snow variables at the snow station. In the final set of features, air temperature, dew point temperature, wind speed, and daily snow depth come from actual observations of these variables at the weather station. The variables, except for daily snow depth, have hourly resolution while the snow depth measurements are made every 24 hours. The observed values represent the most accurate form of data available to DeepSnow's model.

Snow-related Parameters. Most of the weather stations are not equipped with the most advanced equipment and therefore cannot measure the ground truth data for many of the snow-related variables. In order to provide full coverage, the NOHRSC dataset provides the values for the non-observed variables using Numerical Weather Prediction (NWP) models, i.e. hourly snow depth, snow precipitation, etc. NOHSRC runs these models for every $\sim$1km^2 area of the country, and releases them to the public. In our dataset, these measurements are available at an hourly granularity.

Snow Loss Factor. Physical models assume detailed knowledge of a site. These models translate this knowledge into the parameters that the model require. For example, the snow model used by SAM takes the tilt angle and the configuration of solar panels for a site as its input parameters. Given these parameters, it employs a physical model that estimates when the snow will slide off the solar panel surface and how much surface area of all the panels will be exposed to the direct sun after the slide. The proportion of surface area covered by snow determines how much the loss of power due to snow will be. While the snow model used by SAM is simple and makes many assumptions, it still is a good indicator of the snow that is present on a tilted surface. Our evaluation shows that using even a simplistic model can improve the accuracy of modeling considerably when compared with a model that does not use any snow modeling

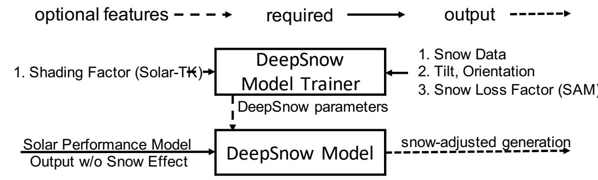


Figure 5: Flow diagram of DeepSnow, its inputs and outputs, and how it integrates into any solar performance model.

at all. Therefore, we use the Snow Loss Factor from SAM as one of the features for our model.

Shading Factor. Physical models or observed values do not exist for all factors that affect solar power output. For example, there is no physical model that describes how shading from nearby objects impact the solar output. Further, shading is unique to each site [6] and local effects needs to be modeled for each site. Machine learning is a useful tool for such situations and we use it to learn the non-linear effects of shading on individual sites. Specifically, we use the shade adjustment model, that employs an SVM kernel with RBF kernel, from Solar-TK to determine the shading factor.

Table 2 presents the final set of features used by the DeepSnow. We combine data from all these multiple sources to train our model rather than relying on any single data type.

B. DeepSnow Model. We have outlined a set of features that our model will use to learn how snow affects the power output of a solar site. Figure 5 illustrates the DeepSnow workflow, how it integrates into a solar performance model's existing workflow, and how a user can use it for a given solar site. DeepSnow's model takes as an input the output of a solar model that has not yet incorporated the effect of snow. For example, in case of Solar-TK the input to DeepSnow would be the shade-adjusted power output for a site. The assumption here is that the output of the solar performance model has already incorporated the effect of site-specific parameters, weather, and optionally shading. DeepSnow's model takes this modeled power as an input and multiplies it with a factor, whose values range from 0 to 1. A value of 1 for the snow adjustment factor means that there is no loss due to snow and the expected power of DeepSnow would be same as the input power. A value of 0 means that snow has completely covered all the solar panels on the site and the snow-adjusted power output for the site would be zero.

The value of the snow adjustment factor is dependent upon the set of features outlined in Table 2. The training of the DeepSnow model depends upon the availability of the features. Snow data is always required and is also publicly available for all sites in the United States. The tilt angle is another important feature as it is not only used as a direct feature to the model, but is also needed to determine the Snow Loss Factor from SAM. Solar-TK's physical parameter module can be used to determine the tilt angle for existing solar sites to which we do not have physical access but have some prior data, i.e. a few days. For the new sites yet to be deployed, the tilt angle can be set arbitrarily to any angle. The orientation feature offers similar challenges and options as the tilt angle. Finally, the shading factor is a feature that models the site-specific shading characteristics and needs prior solar data for the site. Prior work that has proposed or used the shading-adjustment has not fully evaluated its data requirements [6, 8]. However, intuitively, at least one year of data is necessary to account for the shading effect during all times of the year. However, one year of data may not be available for many sites. Thus, we use the shade adjustment as an optional feature.

We use the observed solar output divided by the output of the solar performance model without the snow effect, which is also used as the input to DeepSnow model, as the target output variable for our DeepSnow model training. This ratio gives us the effect of snow on the solar power output and also normalizes the target variable across all the sites. This normalization allows us to train DeepSnow's model using solar data from many sites with different characteristics.

The training of the DeepSnow model can be done for individual sites (local model) or a single model can be built for multiple sites (global model). In training a local model, we use the snow data, snow loss factor, and shade adjustment factor if available. Tilt and orientation angles for a given site would be fixed and therefore are not useful as features. While training a global model, we use the snow data, tilt and orientation angles, snow loss factor, and shade adjustment factor if available. Tilt and orientation angles remain constant for a given site, but the model learns to incorporate the effect of tilt as more and more sites are used for training. The benefit of a global model is it enables using DeepSnow for sites that do not have enough data to train an accurate local model.

In contrast, a local model has the potential to learn unique local factors that influence the effect of snow on a particular site's solar output. However, it requires a large amount of data to train an accurate model, which may not available for most sites. Our hypothesis is that a local model trained for a site with "enough" data can outperform a global model for that site. But, how much data is "enough" to build a superior local model? How much the performance improves by building a local model? What is the effect of optional shading adjustment factor? We answer these questions in our evaluation.

The choice of feature set, the target variable, and the tradeoff of local versus global model are irrespective of the machine learning or deep learning model used. Of course, some machine learning models would perform better with less data and will be suitable for a local model and vice versa. We evaluate these differences on 4 different ML models in our evaluation. The first model we chose is a simple linear regression model. We choose linear regression to evaluate the performance of the simplest model in modeling the effect of snow. SVM models have been used by prior work to model the effect of site characteristics on solar power output, i.e., modeling shading impact on solar sites [8]. Random Forest (RF) is one of the most popular ensemble learning methods as by using multiple samples of the original dataset while building trees, RF reduces the variance of the final model. A low variance means low overfitting and better generalizability. Finally, when building a global model, we will have access to millions of data points. Neural Networks (NN) tend to work better with large datasets. Therefore, we compare the accuracy of our DeepSnow model while using these machine learning models. Once the model is chosen and trained, our DeepSnow model takes the relevant set of features and computes the snow adjustment fraction. This fraction is then multiplied with the modeled input power to the DeepSnow to get the snow-adjusted power output for the site.

4 DEEPSNOW IMPLEMENTATION

We have implemented a prototype of DeepSnow in python. The implementation provides a DeepSnow python module that can be plugged into Solar-TK's modeling workflow. It also provides a DeepSnow module that is compatible with the PVlib python library. Our implementation extensively uses the pandas python data analysis library [2], and the NumPy scientific computing python library [1]. We use hourly snow data from NOHSRC Interactive Maps [4] and daily snow depth data from SNODAS [5]. We have used solar data from Solar-TK's solar data repository. We have released DeepSnow's code and model as part of Solar-TK's release, and made the data for our evaluation public at the UMass Trace Repository.[1]

We use the Scikit-learn [16], a Python library for machine learning, to implement the linear, SVM, and RF regression models. We use SVM with a Radial Basis Function (RBF) kernel. For RF, we use randomized search cross validation function from scikit-learn to determine the optimal hyper-parameters for the model. For the deep learning models, we use PyTorch [15], a deep learning library for Python. We use a 4 hidden layer feed-forward NN implemented using PyTorch. The NN has 128, 128, 128, and 64 neurons in the hidden layers. The hidden layers use the "relu" activation function, while the final output layer has a single output with linear activation. We use L2 regularization and dropout layers to reduce overfitting. A grid search was performed for the selection of hyper-parameters. For training, we use the Adam optimizer with mean absolute error as the loss function. The train and test split size was dependent on the choice of local and global models and is cited in the evaluation.

5 EXPERIMENTAL EVALUATION

The objective of any solar performance model is to accurately model and infer the power generated by a solar site at any given time. As we established earlier, accurate energy estimates over some period, i.e. one year, are also important for the financial models of PV deployments. It is worth pointing out that, while low error in predicting instantaneous power values will generally result in lower error in energy estimations, it is necessarily not a linear relationship. The over and under power estimations may cancel out over longer periods, leading to low errors in energy estimations but high errors in power inference. Therefore, the performance analysis of a solar performance model should take error in both power and energy into account. Thus, we evaluate DeepSnow's accuracy in estimating both annual solar energy production and inferring instantaneous power.

We evaluate and compare the accuracy of DeepSnow with the solar performance models described in §2 on data from 10 solar sites at different locations with widely different characteristics in terms of snowfall. We picked sites across Colorado, Illinois, Massachusetts, Minnesota, and New York. All of the sites have 6-11 years of solar and snow data available. As a general rule, we keep the last 3 years data as the test dataset and train the different models on rest of the data. The rationale behind saving the last three years of data for each site is that we want to estimate annual energy for the sites and want to include all seasons of the year since the accuracy across seasons varies. As a result, the amount of training data for the different sites is not the same. Therefore, we clearly state the data characteristics in each subsequent experiment and discuss its effects on the results.

To quantify accuracy, we use Mean Absolute Percentage Error (MAPE), as follows, between the ground truth solar energy and the

[1] See https://github.com/sustainablecomputinglab/solar-tk and http://traces.cs.umass.edu/index.php/Smart/Smart

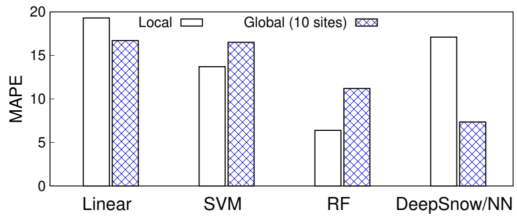


Figure 6: Global and local snow model accuracy in predicting solar output over one year using different ML models.

solar energy that DeepSnow estimates over the time interval, i.e. one year. A lower MAPE indicates higher accuracy with a 0% MAPE being perfectly accurate energy inference.

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

Here, S_t and P_t are the actual and inferred solar energy generation, respectively. We also compute the MAPE between the instantaneous power estimated by DeepSnow and the ground truth solar power.

5.1 Baseline

We compare the accuracy of DeepSnow with Solar-TK [6] and SAM [9]. We use our best effort to model sites equivalently in both tools, which we describe below.

Solar-TK. Solar-TK provides a physical parameter module that estimates the physical characteristics of a solar site and its temperature related parameters. These parameters are then used by the subsequent modules to estimate weather-adjusted power generation. We use one year of data for each site to estimate the physical parameters. We use all the training data for each site to train Solar-TK's shading module, which means that each site's shade training is done using at least 3 years of data. The output of the shading module is a shading index value at each timestamp that dictates how much shading reduces solar output. The shade-adjusted power values from Solar-TK are used as the first baseline. In addition to that, we use the shade index for our DeepSnow model as a feature. The shade index has a value between 0 and 1, a value of 0 means no power due to shading while 1 means shading has no effect for the site.

System Advisor Model (SAM). SAM's configuration for a site requires choosing the exact solar panels used, the wiring scheme for the modules, and the site's tilt, and panel orientation angles. Since we do not have physical access to any of the sites used for this analysis, we do our best effort to configure the sites in SAM. The first step in our approach is to estimate the overall capacity of the solar site, its tilt, and orientation angle. We do this by using Solar-TK's physical parameter module, which has been shown to have high accuracy in prior work [6]. Once we have an estimate of capacity, we manually examine satellite imagery of the site to infer the number of panels, their wiring scheme, and the size of each panel. Since we do not know the exact solar panel type, we tried multiple different modules with the same rating and selected the one that gave the best results. Furthermore, SAM requires input solar-irradiation and weather data in the form of Typical Meteorological Year (TMY) files, we prepare the data for all the sites and all years in the TMY format. We get solar irradiation data from NSRDB database and weather data from the Weather Underground; the same data source is used by DeepSnow and Solar-TK for non-snow weather data.

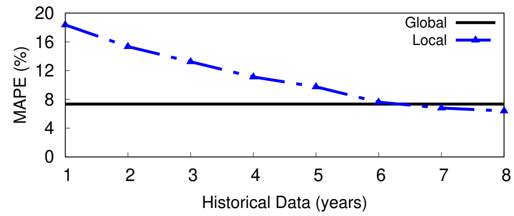


Figure 7: A local model requires more data than a global model.

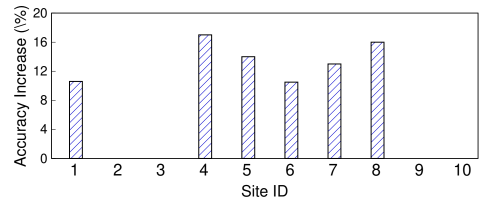


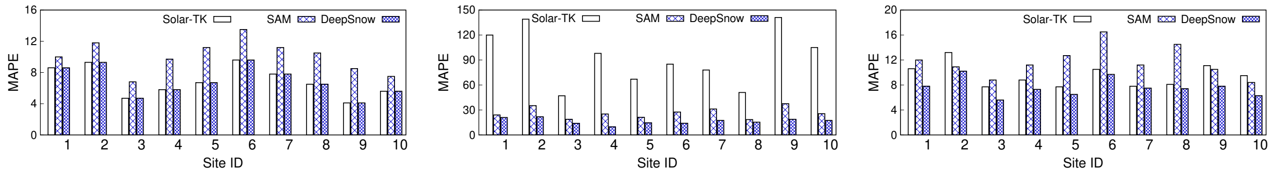
Figure 8: Percentage improvement in accuracy from using shade adjustment factor. Sites 2, 3, 9 and 10 don't experience shading and do not see any improvement as a result.

5.2 Local versus Global Model

The effect of snow, as we have illustrated, depends on variables that are listed in Table 2. Some of these variables are not dependent upon the solar sites, i.e., snow depth, humidity, precipitation, and wind speed. That is, while the value of these variables vary across space and time, they are not dependent upon the configuration of the solar site itself. In contrast, other variables such as the shading factor, tilt, and orientation depend upon the site's configuration. Each of these site-dependent parameters will determine how snow accumulates and melts over time. Therefore, ideally, building a local snow model for each site seems like an attractive option. However, not all of the sites will have enough data to train a local model. In this case, a global snow model is a reasonable alternative, as the accuracy of different machine learning models have a different relationship with the amount of data available, i.e., NNs typically require more training data than the Random Forest model.

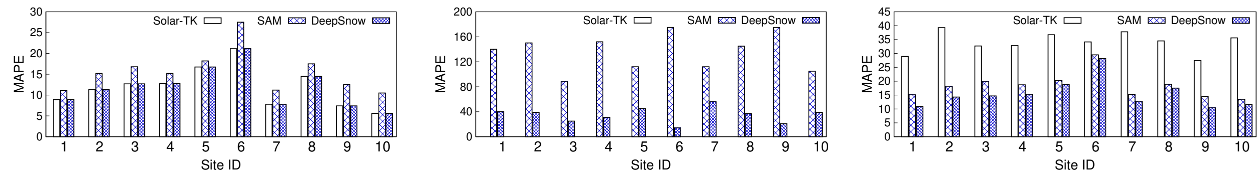
We next evaluate the accuracy of DeepSnow using the linear regression, SVM, RF regression and NN models that we discussed in §3. Figure 6 shows the accuracy of each model when building a local model and a global model. For this analysis, we choose a site for which we have 11 years of data. For simplicity, we will refer to this site as the local site. When training a local model, we use only the data available for the local site. We keep the last 3 years of data as the test dataset. For the global model, we use data from all the sites, which adds up to a total of 56 years. To train the global model, we use all data from the remaining 9 sites used for evaluation. While training the global model, we do not use any data from the local site. As a result, we measure the accuracy of the global model on an unseen site whose data was not used for training.

The linear regression and SVM models show a high error for both the local and global models. The RF model performs well for the local model, as it is able to achieve a good fit for the model. A well-trained local model with 8 years of data achieves a very high accuracy. This happens because the characteristics of the local site do not change over time. However, the model's accuracy degrades as we train a global model with data coming from many sites. NNs are not suitable for building a local model because there is a limited



(a) Summer (b) Winter (c) Overall

Figure 9: Accuracy for Solar-TK, SAM, and DeepSnow in estimating energy during (a) summer, (b) winter, and (c) overall.



(a) Summer (b) Winter (c) Overall

Figure 10: Accuracy for Solar-TK, SAM, and DeepSnow in estimating power during (a) summer, (b) winter, and (c) overall.

amount of data available for a single site. However, a global model trained using NNs achieves higher accuracy. But, its accuracy is still less than a well-trained local model. This presents a trade-off between scalability and accuracy. A global model is less accurate but applies to many sites. A local model can be accurate, but it is often not possible to train such a local model due to a lack of data.

As we saw in Figure 6, a well-trained local model can outperform the global model. However, the local model that outperformed the global model was trained on 8 years of data. The data for 8 years or more is probably not going to be available for most solar sites. Therefore, we also evaluate how the accuracy of the local model improves with the amount of data available. Figure 7 presents the accuracy of the local model as more and more training data is used to train the local model. The local model does not have high accuracy if there is less than 5 years of data available. The accuracy of the local and global models is comparable at 6 years and the local model outperforms the global model with additional data. However, the gain in accuracy becomes marginal even with more data.

The global model presented in Figure 6 uses the shading index feature as an input, which means that we need some data for each site to train the shading module. Therefore, this model cannot be used for sites that have no prior data available or even for the sites that have very little data available. To remove this condition, we trained the global model without the shading feature and compared its accuracy with the global model with the shading feature. Figure 8 compares the accuracy of a global model with the shading index as an input feature and without it. The shading index feature is not important for the sites that do not have any shading present, as with sites 2, 3, 9 and 10 in this case. For the sites with shading present, adding the shading factor as the feature improves the accuracy of the global model by 10-15% with an average improvement of 12%. We use a global model with shading feature for all subsequent experiments.

5.3 Energy

The accuracy in inferring energy output is important for estimating a site's revenue. A higher accuracy allows the users to make better financial projections, which affects the value of a site and possible financing. Figure 9 shows the error in inferring solar energy production during summer, winter, and overall during a year. Since DeepSnow uses Solar-TK's solar-performance model and simply adds a module to incorporate the snow effect, both Solar-TK and DeepSnow yield the same accuracy in the absence of snow. SAM on the other hand has lower accuracy than DeepSnow for all 10 the sites. For some of the sites, i.e., site 7, the difference is exacerbated by the shading effect present at the site. The shading effect is not modeled in SAM as shading is a site-specific property which must be learned from data. DeepSnow uses the shading module of the Solar-TK, and is able to achieve a higher accuracy.

The accuracy of all approaches decreases during the winter season, as can be seen in Figure 9(b). Solar-TK has the worst accuracy as it does not incorporate the effect of snow in its modeling approach. The simple physical model used by SAM improves the accuracy of the model during the winter where it significantly outperforms Solar-TK. The improvement in accuracy is significant as Solar-TK predicts solar power to be near its maximum expected, especially if it is a sunny day. However, in presence of snow even on a sunny day, the actual power generated will be closer to 0 than the maximum power. Therefore, even a simple model that predicts the power to be zero under snow would likely perform better job than Solar-TK on a sunny day with a large amount of snow. DeepSnow improves the accuracy of inference beyond SAM as it not only considers the simple snow model used by SAM, but also incorporates the effect of shading and orientation. It also learns additional information from the variables gathered through NOHSRC dataset.

Even after incorporating the snow model, both SAM and DeepSnow have a fairly high error during the winter season. However, the winter season in our analysis only includes time periods when the snow depth is greater than 0. The total number of days is around 2 months per year on average for all of the sites, which means that the overall effect of inaccuracy during winter is reduced when viewed over the whole year. In addition, the energy generated during the winter is also less, which further lessons the impact of winter inaccuracy. However, this does not mean that incorporating the effect of snow is not valuable. Figure 9(c) shows the overall MAPE for all the three approaches over the course of a year. We can see that DeepSnow improves the accuracy of Solar-TK considerably as it incorporates the snow effect. However, while a snow model improves the winter accuracy for SAM, the better accuracy of Solar-TK in solar modeling for non-snow periods means that it has an overall better accuracy than SAM even for the sites with less snow.

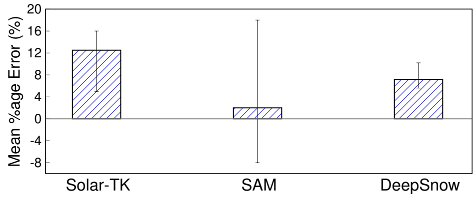


Figure 11: Error analysis of different modeling approaches.

5.4 Power

The accuracy in inferring the instantaneous power output is important for the grid operators, as higher accuracy allows them to plan ahead and manage their supply resources better. Figure 10(a) shows the error in inferring solar power during the summer. As with the case of energy, Solar-TK and DeepSnow have the same accuracy during the summer because their underlying performance model in the absence of snow is the same. DeepSnow, however, achieves better accuracy than SAM for all the sites. The shading effect on site 7 worsens the accuracy of the SAM model, as was the case for energy.

The accuracy of all approaches during the winter is worse than the summer season, as shown in Figure 10(b). Solar-TK is not present in this figure as the MAPE values for it across sites range from 1700-2500%. The high MAPE for Solar-TK is expected as it does not incorporate snow effects. In the event of snow, as the actual power approaches 0, the MAPE of Solar-TK approaches infinity. Therefore, Solar-TK alone would not provide good accuracy during the winter for the sites that experience snow. SAM outperforms Solar-TK with its simple physical model of snow, but the error across sites ranges from 100-180%, which limits its use during winters. DeepSnow yields better accuracy than SAM, as it not only incorporates the SAM model but also models the effect of additional variables.

The overall accuracy for all the approaches lies between the accuracy during summer (high) and winter (low). However, the actual magnitude of the error for different approaches is interesting. While Solar-TK does better during the summer, it has poor accuracy during the winter, which makes itsoverall error quite high (28-40%, average ~35%). However, the addition of DeepSnow reduces its MAPE range to 8-25% with an average value of 12%. Thus, DeepSnow reduce sthe MAPE of Solar-TK by ~65%. The accuracy improvement over SAM is also significant at 25%. SAM's average MAPE is 15% as compared to 12% for DeepSnow. However, even this 3% difference is significant as it translates to revenue that compounds.

5.5 Distribution of Errors

MAPE captures both underestimates and overestimates. However, underestimates and overestimates have different impacts. If a model underestimates the annual energy yield, the solar deployment would appear to be a less attractive financial investment, and causes the solar site to have a lower value. In contrast, overestimating overvalues a site, causing the owner to expect more revenue than they will actually earn. Ideally, a model would have a zero error, but a model that constantly overestimates is better than the one that underestimates. Figure 11 shows the distribution of absolute average error for all the modeling approaches. Solar-TK has a higher error than both SAM and DeepSnow, as expected. However, it is interesting to note that while SAM has a higher MAPE than DeepSnow, its average absolute error is low with a very high error range. This means that SAM underestimates for some sites while overestimates for others. Thus, using SAM, these sites have no certainty over whether the model is overestimating or underestimating the power. In contrast, DeepSnow has a slightly average absolute error with a narrow range and always over-predicts. Thus, DeepSnow is much more likely to be an overestimate within a narrow range. That is, DeepSnow's errors tend to be much more consistent and predictable than SAM.

6 CONCLUSIONS

There has been decades of research on solar modeling and forecasting using physical and empirical models. DeepSnow builds on and leverages much of this work. DeepSnow's novelty lies in its empirical modeling of the effect of snow on solar output. None of the prior work does well under snow, and most does not even attempt to model the effect of snow, even though it is significant at higher latitudes. DeepSnow's data-driven approach leverages existing solar modeling frameworks, and uses publicly available snow data. Our data analysis quantifies the effect of different snow variables on solar power using 4 million hourly readings from 40 solar sites. We then evaluate our approach on 10 solar sites, and show that it yields a higher accuracy than the current approach for modeling snow effects used by the the U.S. Department of Energy's System Advisor Model (SAM), a popular solar modeling framework.

Acknowledgements. Research supported by NSF grant #1645952.

REFERENCES

[1] 2019. NumPy. https://www.numpy.org/.
[2] 2019. Python Data Analysis Library. https://pandas.pydata.org/.
[3] 2019. www.nohrsc.noaa.gov. https://www.nohrsc.noaa.gov.
[4] 2020. National Operational Hydrologic Remote Sensing Center (NOHRSC) Snow Data. https://www.nohrsc.noaa.gov/interactive/html/map.html
[5] 2020. Snow Data Assimilation System (SNODAS) Data Products at NSIDC. https://nsidc.org/data/g02158
[6] N. Bashir, D. Chen, D. Irwin, and P. Shenoy. 2019. Solar-TK: A Data-driven Toolkit for Solar PV Performance Modeling and Forecasting. In *MASS*.
[7] N. Blair, A.P. Dobos, J. Freeman, T. Neises, M. Wagner, T. Ferguson, P. Gilman, and S. Janzou. 2014. *System Advisor Model, SAM 2014.1. 14: General Description*. Technical Report. National Renewable Energy Lab. (NREL).
[8] D. Chen, J. Breda, and D. Irwin. 2018. Staring at the Sun: A Physical Black-box Solar Performance Model. In *BuildSys*.
[9] J.M. Freeman, N.A. DiOrio, N.J. Blair, T.W. Neises, M.J. Wagner, P. Gilman, and S. Janzou. 2018. *System Advisor Model (SAM) General Description (Version 2017.9. 5)*. Technical Report. National Renewable Energy Lab. (NREL).
[10] JC Giddings and E LaChapelle. 1961. Diffusion Theory Applied to Radiant Energy Distribution and Albedo of Snow. *Journal of geophysical research* (1961).
[11] W. Holmgren, C. Hansen, and M. Mikofski. 2018. PVlib Python: A Python Package for Modeling Solar Energy Systems. *JOSS* (2018).
[12] O. Järvinen and M. Leppäranta. 2013. Solar radiation transfer in the surface snow layer in Dronning Maud Land, Antarctica. *Polar Science* 7, 1 (2013), 1–17.
[13] B. Marion, R. Schaefer, H. Caine, and G. Sanchez. 2013. Measured and Modeled Photovoltaic System Energy Losses from Snow for Colorado and Wisconsin Locations. *Solar Energy* 97 (2013), 112–121.
[14] ADJ. O'Neill and D.M. Gray. 1972. Solar Radiation Penetration Through Snow. In *The Role of Snow and Ice in Hydrology*, Vol. 107. 227–240.
[15] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc.
[16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* (2011).
[17] D.K. Perovich. 2007. Light Reflection and Transmission by a Temperate Snow Cover. *Journal of Glaciology* 53, 181 (2007), 201–210.
[18] W. Shockley and H.J. Queisser. 1961. Detailed Balance Limit of Efficiency of p-n Junction Solar Cells. *Journal of Applied Physics* 32 (1961).