# Self-Supervised Learning on Multispectral Satellite Data for Near-Term Solar Forecasting

Akansha Singh Bansal [1] Trapit Bansal [1] David Irwin [1]

## Abstract

With the unprecedented increase in distributed photovoltaic (PV) capacity across the globe, there is an increasing need for reliable and accurate forecasting of solar power generation. While PV output is affected by many factors, the atmosphere, i.e., cloud cover, plays a dominant role in determining the amount of downwelling solar irradiance that reaches PV modules. This paper demonstrates that self-supervised learning of multispectral satellite data from the recently launched GOES-R series of satellites can improve near-term (15 minutes) solar forecasting.

We develop deep auto-regressive models using convolutional neural networks (CNN) and long short-term memory networks (LSTM) that are globally trained across many solar sites on the raw spatio-temporal data from GOES-R satellites. This self-supervised model provides estimates of future solar irradiance that can be fed directly to a regression model trained on smaller site-specific solar data to provide near-term solar PV forecasts at the site. The regression implicitly models site-specific characteristics, such as capacity, panel tilt, orientation, etc, while the self-supervised CNN-LSTM implicitly captures global atmospheric patterns affecting a site's solar irradiation. Results across 25 solar sites show the utility of such self-supervised modeling by providing accurate near-term forecast with errors close to that of a model using current ground-truth observations.

## 1. Introduction

Energy generation from renewable sources surpassed coal-fired generation for the first time in the U.S. in 2019 (EIA, 2020). The U.S. Energy Information Administration (EIA) projects that the share of renewable sources, such as wind and solar, in the electricity generation mix will double to almost 42% by 2050, and solar generation is poised to account for almost 80% of this increase (EIA, 2021). A large portion of solar energy comes from small-scale rooftop residential solar installations, including private homes and businesses. This in turn implies that consumers, instead of buying electricity from the grid, are now generating their own energy as well as selling it back to the grid. All of this decentralized and intermittent energy generation by millions of individual users fundamentally alters the grid's supply and demand balance. The underlying problem is the uncertainty associated with solar and that solar has a potentially infinite ramp rate – its output can rise and fall instantaneously. In contrast, the ramp rate for conventional thermal generators is finite – it takes time to ramp up generation. Since the grid must balance electricity's supply and demand at all times, any rise or drop in solar generation must be compensated by using conventional generators. To do so, utilities often keep these generators on "spinning reserve", such that they are active but not connected. Such spinning reserves increases energy's cost and carbon emissions.

Accurate near-term solar forecasts can reduce the need for such spinning reserves by enabling utilities to better plan when to activate generators. Thus, such near-term forecasts can help grid operators plan and balance electricity generation and consumption (Wu et al., 2015; Haupt et al., 2019). Moreover, it can help in energy market optimization (Kaur et al., 2016), help in planning of residential power and grid defection (Bansal & Irwin, 2019), and lower the risks associated with using intermittently available solar power (Quan et al., 2015). While some utility-scale solar sites may invest in developing their own custom models and forecasts, solar is becoming highly distributed with millions of small sites, which motivates a more automated approach.

Solar generation is affected by a number of factors, including a solar site's physical characteristics as well as atmospheric conditions. While physical characteristics, such as solar module area, tilt, orientation, etc., have well-known physical models (Chen & Irwin, 2017), modeling and forecasting the impact of passing clouds and other atmospheric phenomenon are more difficult and often the largest sources

[1]University of Massachusetts, Amherst. Correspondence to: Akansha Singh Bansal <akanshasingh@umass.edu>.

Tackling Climate Change with Machine Learning Workshop at ICML 2021.

of error (Yang et al., 2012; Sanders et al., 2017; Chen et al., 2018; Bansal & Irwin, 2020a; Siddiqui et al., 2019). While, ground-level weather monitoring stations include weather sensors to facilitate such forecasts, they can be sparsely located. Thus, a more accessible approach is to develop machine learning (ML) models from multispectral satellite data that can remotely and uniformly sense large portions of the globe. This approach is especially promising for near-term solar forecasting.

Forecasting solar PV output is akin to forecasting solar irradiance (which is widely available) since the former strongly correlates with the latter (Raza et al., 2016). Numerical Weather Predictions (NWP) algorithms (Gamarro et al., 2019; Chen et al., 2017; Tiwari et al., 2018; Mathiesen & Kleissl, 2011), that mostly leverage physics-based modeling, are often used for solar irradiance forecasting. These physics-based models are most appropriate for forecast horizons on the scale of hours to days, and not near-term forecasts on the scale of minutes to an hour (Hao & Tian, 2019; Wang et al., 2019). Over long-term horizons, the complex and non-linear evolution of climate patterns can be difficult to model, requiring knowledge of climate processes and the history of many atmospheric events over time that can cause subtle changes. On the other hand, at shorter time scales of 5 to 60 minutes, machine learning can hold the potential to implicitly model local changes directly from observational data (Wang et al., 2019; Rolnick et al., 2019). While there is recent work on analyzing images from ground-based sky cameras (Zhang et al., 2018; Siddiqui et al., 2019; Paletta & Lasenby, 2020) for near-term solar forecasting, it requires installing additional infrastructure at the site. Another alternative is based on estimating cloud motion vectors (Lorenz & Heinemann, 2012) from satellite images, however ML approaches that more directly model solar irradiance tend to perform better (Lago et al., 2018; Bansal & Irwin, 2020b).

In this work, we explore methods using satellite data from GOES-16, a recent geo-stationary satellite that was launched by NOAA and NASA in late 2017 and began releasing data in 2018 (NOAA, 2018). They generate data in 16 spectral channels comprising of different wavelengths of light at a high spatial resolution of 1-2 km$^2$ area and high temporal resolution of 5 minutes. These minute-level satellite images capture many atmospheric phenomenon, such as cloud motion, that are high predictive of solar irradiance at the surface and can be used to infer solar PV output using ML (see Figure 1, Bansal & Irwin (2020a)). We develop self-supervised deep learning models using CNN and LSTM to directly learn to forecast future satellite observations at any site of interest. Since satellite data is available in abundance, such self-supervised objectives can enable training of accurate deep learning models that implicitly capture local statistical patterns over short intervals and provide accurate forecasts of future observations at a site. When evaluated across 25

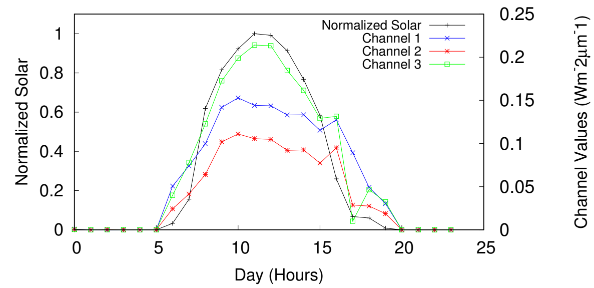


*Figure 1. Relationship between actual solar PV output and first 3 channels from GOES-16 for a single site on a sunny day.*

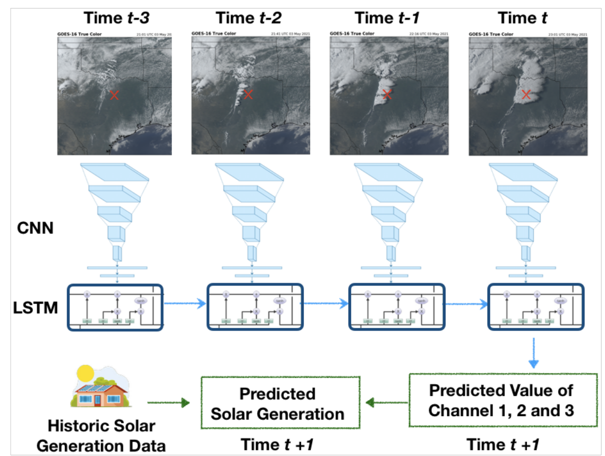


*Figure 2. Overview of our complete solar forecasting approach.*

residential solar sites, spread across the U.S., we find that these predictions prove useful in a simple auto-regressive model to forecast site-specific solar PV output at 15 minutes in the future, yielding errors close to an upper bound that takes the true future satellite observations as input.

## 2. Modeling

In this section, we describe our complete methodology for solar forecasting. We first describe our neural network model for self-supervised learning on raw satellite observations. Then, we describe a simple auto-regressive model for solar generation forecasting that can leverage forecasts from the self-supervised model.

### 2.1. Self-supervised modeling of satellite data

Short-term forecasting of solar power needs to take into account the recent changes to solar irradiation at the surface and how it will evolve in a short span of time. For example, cloud cover and movement is a key component that determines the amount of down-welling solar irradiation. Indeed, one of the prominent approach to solar modeling and forecasting is through using coarse measurements of cloud cover, called Okta (okt, 2019; Chen et al., 2018). This is also the major source of error in current models as these measurements are coarse and imprecise (Chen et al., 2018; Bansal & Irwin, 2020a;b). Bansal & Irwin (2020a) instead showed that GOES visible bands are highly correlated with solar irradiance at the surface, allowing accurate inference of solar power through machine learning models trained on historic generation data with GOES visible bands as inputs.

We seek to utilize this relationship (Fig. 1) for forecasting by using deep neural networks to implicitly model the short-term changes in the values of the visible bands of GOES. We will use this model to forecast future values of the visible channels which in turn will help in predicting solar generation, owing to this relationship between channel and solar generation. Figure 2 shows an overview of our modeling approach. Given a set of solar sites of interest, we consider an area of $w \times h$ around the site and extract the 3 visible channels of satellite data from the GOES-16 satellite. We get a temporal sequence of $w \times h$ images over time from the satellite, with the target site always at the center of the image as this is a geo-stationary satellite. We consider sequence of upto 5 such images at a time with an interval of 15 minutes between images, spanning an hour of satellite imagery prior to the last image. Each 3D satellite image is processed through a multi-layer convolutional neural network (CNN) into a fixed $d$-dimensional representation at each time-step. The sequence of CNN representations from each step are then fed into a long short-term memory network (LSTM) to capture the temporal evolution relevant for predicting observations at future time-steps.

The final step output is from an LSTM, and, after processing 4 sequences of images (1 hour), is passed through a dense layer with sigmoid output units to predict the value of the visible channels at the site location after 15 minutes. The CNN-LSTM model is trained end-to-end using mean-squared error, with respect to the true future satellite values, as the loss function. Note that we only train one global model by combining satellite data across multiple sites. This enables modeling of shared statistical properties rather than overfitting to the peculiar characteristics of an isolated site. Moreover, this also provides abundant data for learning a useful CNN-LSTM model which typically don't work well with small datasets.

### 2.2. Solar generation forecasting

Given a trained CNN-LSTM model that can generate future satellite observations at a given site, we aim to leverage these predictions in a model for solar power forecasting at any solar installation site of interest. Bansal & Irwin (2020a) showed that the current channel value can be used to infer the site-specific solar generation by training a simple regression model on historic data which helps correlate satellite observations with solar installation specific characteristics. We build on that finding by considering the following auto-regressive model for forecasting near-term solar output:

$$P_{t+1} = f(P_t, C_{t+1}, T_t) \tag{1}$$

where $P_t$ is the solar power generated, $C_t$ are the satellite visible channel values and $T_t$ is the temperature at time $t$. $f(\cdot)$ is a regression model, such as support vector regression, that models the relationship between the input and output variables using historical data. Temperature is an important component of solar generation as solar panel efficiency is sensitive to the surrounding temperature (Chen et al., 2018). Note that we use $C_{t+1}$ instead of $C_t$ in (1). A major component of change in $P_{t+1}$ from $P_t$ is captured in the change in $C_{t+1}$ from $C_t$. This complex relationship is modeled using our CNN-LSTM model, described above, which allows us to predict $\hat{C}_{t+1}$, an estimate of true channel values at $t+1$ to use in the auto-regressive model in (1).

## 3. Evaluation

### 3.1. Implementation and Evaluation Details

We use data for 25 solar sites across two years. Satellite data for the entire continental U.S. is extracted for the 2019 year. We restrict modeling to a $10 \times 10$ window around the 25 solar sites which constitutes the training data for the CNN-LSTM model for compute efficiency. We average observations in a 15-minute window, this helps reduce the sequence length for modeling and noise in the data by reducing the number of missing observations and sensor errors. This yields more than *3 Million* 5-step sequences of $10 \times 10$ images (with 3 channels) at intervals of 15 minutes. CNN model comprises of 2 blocks of convolutions, where each block contains 2 convolution layers with 32 filters of size $3 \times 3$ and ReLU activation followed by a max-pooling layer of size $2 \times 2$. This is followed by two dense layers with hidden dimension $d = 256$ and ReLU non-linearity between layers. We use a one layer LSTM that takes these 256 dimensional inputs and has a hidden dimension of 64.

We use 5-fold validation in all the experiments, splitting by day so that test sets have entire days hold-out for evaluation. Solar generation data from the energy meters for 25 sites and temperature data from the weather station are from years 2018-19. We restrict generation data to be from 9 am to 3pm every day, which is the peak duration of solar generation. The metrics used are Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE). MAPE, often used to quantify the performance in prior work (Wang et al., 2019), is an intuitive metric and is comparable across solar sites of different installation size and configurations. However, it is sensitive to periods to low absolute solar generation and can be significantly affected by small absolute errors. We also used MAE to quantify the error in channel modeling given that the first three channels are reflectance values in the range of 0 to 1.

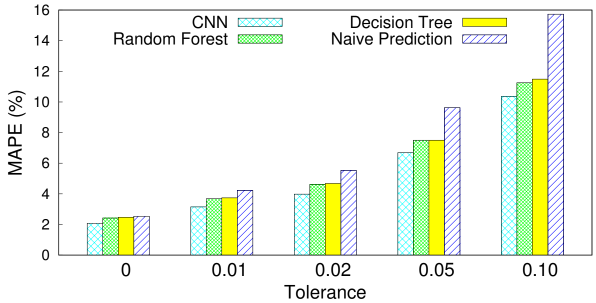


*Figure 3. MAPE of multiple ML models to predict (t+1)th value.*

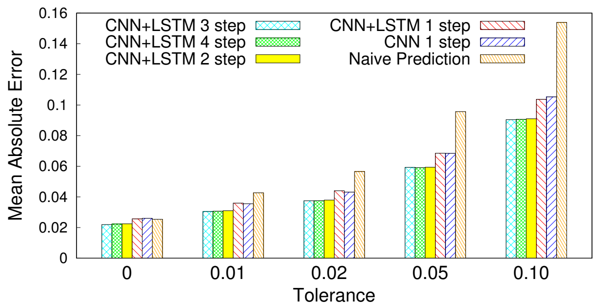


*Figure 4. MAPE of CNN and CNN-LSTM models for 15-minute forecasts of first 3 satellite channels. Here step means the number of time instants of the past values used.*

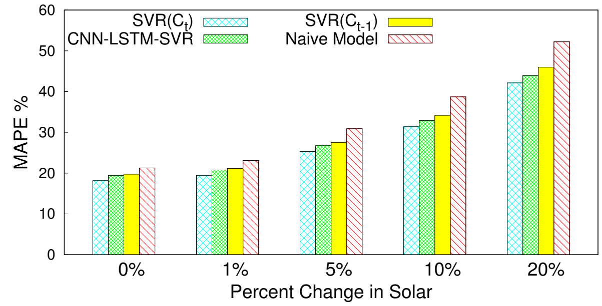


*Figure 5. End to end solar forecasting on 10x10 km area, averaged over 25 solar sites over 15 mins, summer months (May-September).*

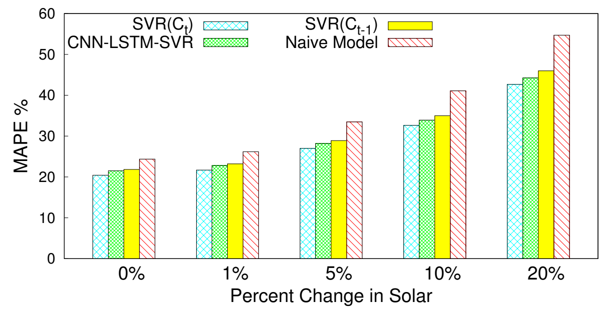


*Figure 6. End to end solar forecasting on 10x10 km area, averaged over 25 solar sites over 15 mins, full-year.*

### 3.2. Results

The first step to solar forecasting is channel forecasting using the CNN-LSTM model. Since the inputs are 3D images of the area around a site, we first evaluate whether CNN models will be more suited for this modeling compared to simpler ML models. We compare using decision tree, random forest and convolutional neural networks for modeling single static images to forecast the next instant channel values. This is compared to a naive model which assumes there will be no change to the channel values. In Figure 3, we can see all the models' performance as a function of the amount of change in the absolute channel values between successive time instants (x-axis). It is important to note that a simple past predicts the future works the best when there are not enough changes, but there is a need for accurate prediction models when there is substantial change in the time series data. In this case, CNN performs the best over all the other models and will be the basis of subsequent modeling.

We next evaluate the performance of CNN-LSTM variants in forecasting next time instant channel values. We explore the following variants: CNN using 1-step static image, CNN-LSTM using 1-step static image, CNN-LSTM using 2/3/4 steps in the past. Figure 4 shows our results compared with the naive predictions. Incorporating multiple-steps of information in CNN-LSTM is better than using the current static image for forecasting, showing the utility of a deep auto-regressive approach. We find that using 3 or 4 steps, i.e., 45 minutes or 60 minutes in the past, perform comparably.

The end goal of this work is to use these channel predictions and translate them into an end to end solar forecasting model. We use the SVR auto-regressive model, discussed in 2.2, to forecast 15-minute ahead solar generation. We consider 4 different models to evaluate forecast at time $t$: (1) Naive: this is again a simple past predicts the future baseline; (2) SVR($C_t$): this is an upper-bound that uses the ground-truth observation at the future instant and is not a feasible forecast as $C_t$ is unavailable ahead of time; (3) CNN-LSTM-SVR: the same model however using the forecasted channel values from past values $(C_{t-1}$,$C_{t-2}$,$C_{t-3})$ through the CNN-LSTM model (4) SVR($C_{t-1}$): using $C_{t-1}$ as a naive forecast in instead of CNN-LSTM, this should be a lower-bound if the CNN-LSTM model produces useful forecasts. Results are in Figure 5, considering only summer months, and in Figure 6 for the whole year. The performance of forecasting solar using CNN-LSTM is close to using the actual channel values in the model, an upper-bound, hence showing that the approach is useful and accurate for solar forecasting. We have further split the performance of these models into percent changes between successive solar generation values as shown as on the x-axis where 0 means any change and includes all the values whereas as 5% means a change of at least 5% in subsequent values and so on. We can also compare Figure 5 and 6 in that they both show similar trends but only differ in the MAPE% which is higher for full year and a little lower for only summer months.

## 4. Conclusion

This paper provides initial results for applying deep learning to satellite data to perform site-specific solar forecasting. We use deep auto-regressive models that combine CNNs and LSTM applied to spatio-temporal data from GOES-R satellites. Our results are promising and show that 15 minute forecasts have an error near that of a solar model using current weather. We plan to extend our results by evaluating over larger areas and longer time horizons.

## References

NASA, Make a Sky Mirror to Observe Clouds and Contrails. https://mynasadata.larc.nasa.gov/science_projects/make-a-sky-mirror-to-observe-clouds-and-contrails/, June 2019.

Bansal, A. S. and Irwin, D. On the feasibility, cost, and carbon emissions of grid defection. In *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 1–7, 2019. doi: 10.1109/SmartGridComm.2019.8909691.

Bansal, A. S. and Irwin, D. See the light: Modeling solar performance using multispectral satellite data. In *Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys)*, pp. 1–10, 2020a. ISBN 9781450380614. URL https://doi.org/10.1145/3408308.3427610.

Bansal, A. S. and Irwin, D. Exploiting satellite data for solar performance modeling. In *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 1–7, 2020b. doi: 10.1109/SmartGridComm47815.2020.9302984.

Chen, D. and Irwin, D. Black-box Solar Performance Modeling: Comparing Physical, Machine Learning, and Hybrid Approaches. In *Greenmetrics*, June 2017.

Chen, D., Breda, J., and Irwin, D. Staring at the Sun: A Physical Black-box Solar Performance Model. In *BuildSys*, November 2018.

Chen, K., He, Z., Chen, K., Hu, J., and He, J. Solar energy forecasting with numerical weather predictions on a grid and convolutional networks. In *2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1–5, 2017. doi: 10.1109/EI2.2017.8245549.

EIA. U.S. electricity generation from renewables surpassed coal in april. https://www.eia.gov/todayinenergy/detail.php?id=42336, 2020.

EIA. EIA projects renewables share of U.S. electricity generation mix will double by 2050. https://www.eia.gov/todayinenergy/detail.php?id=46676, 2021.

Gamarro, H., Gonzalez, J. E., and Ortiz, L. E. On the Assessment of a Numerical Weather Prediction Model for Solar Photovoltaic Power Forecasts in Cities. *Journal of Energy Resources Technology*, 141(6), 03 2019. ISSN 0195-0738. doi: 10.1115/1.4042972. URL https://doi.org/10.1115/1.4042972. 061203.

Hao, Y. and Tian, C. A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting. *Applied energy*, 238:368–383, 2019.

Haupt, S. E., Casado, M. G., Davidson, M., Dobschinski, J., Du, P., Lange, M., Miller, T., Mohrlen, C., Motley, A., Pestana, R., et al. The use of probabilistic forecasts: Applying them in theory and practice. *IEEE Power and Energy Magazine*, 17(6):46–57, 2019.

Kaur, A., Nonnenmacher, L., Pedro, H. T., and Coimbra, C. F. Benefits of solar forecasting for energy imbalance markets. *Renewable Energy*, 86:819–830, 2016. ISSN 0960-1481. doi: https://doi.org/10.1016/j.renene.2015.09.011. URL https://www.sciencedirect.com/science/article/pii/S0960148115302901.

Lago, J., De Brabandere, K., De Ridder, F., and De Schutter, B. Short-term forecasting of solar irradiance without local telemetry: A generalized model using satellite data. *Solar Energy*, 173:566–577, 2018.

Lorenz, E. and Heinemann, D. Prediction of solar irradiance and photovoltaic power. 2012.

Mathiesen, P. and Kleissl, J. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental united states. *Solar Energy*, 85(5):967–977, 2011. ISSN 0038-092X. doi: https://doi.org/10.1016/j.solener.2011.02.013. URL https://www.sciencedirect.com/science/article/pii/S0038092X11000570.

NOAA. Geostationary Operational Environmental Satellites-R Series. https://www.goes-r.gov/, 2018.

Paletta, Q. and Lasenby, J. Convolutional neural networks applied to sky images for short-term solar irradiance forecasting, 2020.

Quan, H., Srinivasan, D., Khambadkone, A. M., and Khosravi, A. A computational framework for uncertainty integration in stochastic unit commitment with intermittent renewable energy sources. *Applied Energy*, 152:71–82, 2015. ISSN 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2015.04.103. URL https://www.sciencedirect.com/science/article/pii/S0306261915005668.

Raza, M. Q., Nadarajah, M., and Ekanayake, C. On recent advances in pv output power forecast. *Solar Energy*, 136: 125–144, 2016.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., et al. Tackling climate change with machine learning. *arXiv preprint arXiv:1906.05433*, 2019.

Sanders, S., Barrick, C., Maier, F., and Rasheed, K. Solar radiation prediction improvement using weather forecasts. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 499–504, 2017. doi: 10.1109/ICMLA.2017.0-112.

Siddiqui, T. A., Bharadwaj, S., and Kalyanaraman, S. A deep learning approach to solar-irradiance forecasting in sky-videos, 2019.

Tiwari, S., Sabzehgar, R., and Rasouli, M. Short term solar irradiance forecast using numerical weather prediction (nwp) with gradient boost regression. In *2018 9th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, pp. 1–8, 2018. doi: 10.1109/PEDG.2018.8447751.

Wang, H., Lei, Z., Zhang, X., Zhou, B., and Peng, J. A review of deep learning for renewable energy forecasting. *Energy Conversion and Management*, 198:111799, 2019.

Wu, J., Botterud, A., Mills, A., Zhou, Z., Hodge, B.-M., and Heaney, M. Integrating solar pv (photovoltaics) in utility system operations: Analytical framework and arizona case study. *Energy*, 85:1–9, 2015.

Yang, D., Jirutitijaroen, P., and Walsh, W. M. Hourly solar irradiance time series forecasting using cloud cover index. *Solar Energy*, 86(12):3531–3543, 2012.

Zhang, J., Verschae, R., Nobuhara, S., and Lalonde, J.-F. Deep photovoltaic nowcasting. *Solar Energy*, 176:267–276, 2018.