
HYBRID MODELING APPROACH USING CLOUD DYNAMICS AND DEEP LEARNING FOR SHORT-TERM SOLAR FORECASTING

Jun Sasaki Kenji Utsunomiya Maki Okada Koji Yamaguchi
Environment and Energy Department
Japan Weather Association
{sasaki.jun,utsunomiya,okada.maki,yamaguch}@jwa.or.jp

ABSTRACT

This study introduces a novel hybrid model combining deep learning and Numerical Weather Prediction (NWP), specifically targeting cloud dynamics, for accurate short-term solar irradiance forecasting. This is particularly crucial for optimizing photovoltaic power systems. Utilizing satellite imagery and deep learning techniques, the model accurately estimates initial fields of cloud thickness, wind speed, and atmospheric pressure in three dimensions. It employs atmospheric dynamics for cloud advection and neural networks for cloud microphysics, enhancing cloud formation and dissipation predictions. Compared to traditional methods, including NWP and video prediction methods, our model exhibits superior accuracy in short-term forecasting. This significant advancement contributes to the optimization of photovoltaic power generation.

Keywords solar forecasting · cloud dynamics · physics-informed neural network

1 Introduction

Highly accurate forecasting of solar irradiance is essential for the operation of photovoltaic power generation systems. Particularly, short-term forecasting is utilized for supply and demand adjustment in the electricity intraday market. This study focuses on forecasting of solar irradiance 6 hours ahead and develops a forecast model to improve the accuracy. In general, solar irradiance can be calculated from the spatial distribution of cloud thickness using a radiative transfer model. Therefore, if the spatiotemporal variation of cloud thickness can be accurately predicted, it is possible to compute the forecast values of solar irradiance for each forecast time through the radiative transfer model. The temporal changes in cloud thickness are primarily caused by advection dependent on wind and atmospheric pressure, and by formation and dissipation dependent on temperature and water vapor content. Particularly in short-term forecasting, it is important to predict these processes at a small scale, such as a spatial resolution of 1 km and a temporal resolution of 10 minutes.

Traditionally, for short-term forecasting of solar irradiance, statistical models and high-resolution numerical weather predictions (hereafter referred to as NWP or HRES-NWP) have been utilized. In statistical models, the initial values are typically estimated from satellite observation data, followed by calculating the time evolution using cloud displacement vectors that are estimated through a template matching method [1, 2, 3]. Though the direct utilization of satellite observation data results in high accuracy in the estimation of initial values, the accuracy in the description of time evolution is limited. This limitation arises from difficulties in accurately describing the advection of clouds at a small scale, their vertical structure, and their formation and dissipation processes. In recent years, deep learning techniques for image generation and video prediction have been applied to short-term solar forecasting [4, 5]. However, these methods are considered to be insufficient in accurately describing the complex behaviors of clouds, including formation, dissipation, and advection associated with vertical wind shear, due to the lack of sufficient training data to cover the countless variations of clouds. On the other hand, while HRES-NWP models can describe advection at small scales, formation and dissipation, they have difficulties in directly assimilating satellite images due to their high computational costs, leading to lower accuracy in the initial values of cloud thickness and wind speed. Additionally,

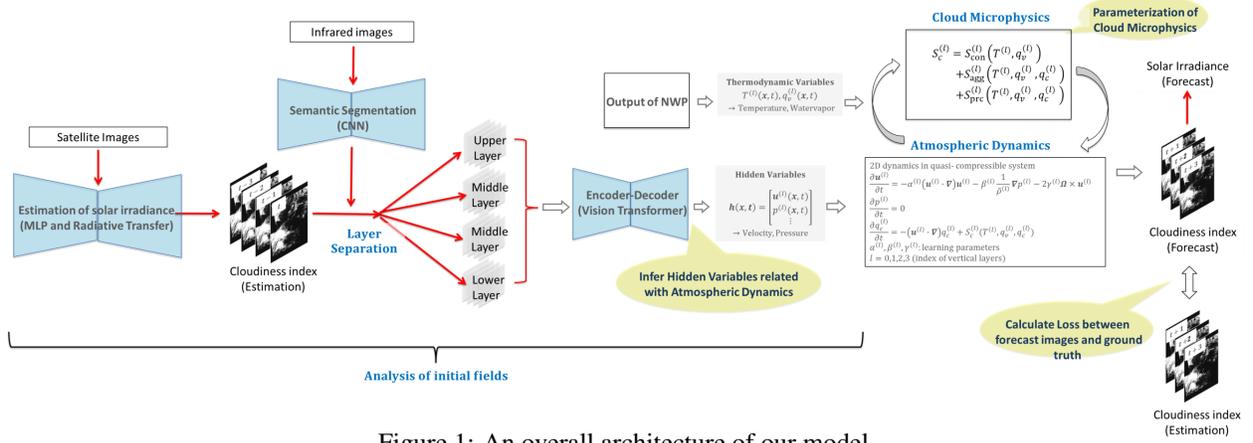


Figure 1: An overall architecture of our model.

cloud microphysics are formulated based on empirical parameterizations, which indicates that the representation of formation and dissipation processes may also be incomplete.

In this study, we have developed a hybrid model that combines AI with NWP techniques. Our model utilizes satellite images and a deep learning framework to estimate the three-dimensional initial fields of cloud thickness, wind speed, and atmospheric pressure. For the description of time evolution, the equations of atmospheric dynamics related to cloud advection, which rely on fluid dynamics and thermodynamics with less physical uncertainty, are explicitly provided with minimal parameters. Conversely, for the cloud microphysics related to formation and dissipation, which has relatively more uncertainty, neural network-based modeling is employed for optimal parameterization. The overall architecture of the model is shown in Fig. 1. These approaches enable accurate estimation of initial fields, including vertical structures, and precise description of cloud formation and dissipation processes. This advancement overcomes the limitations of traditional models and is expected to improve the accuracy of solar forecasting. In practice, when the trained model was employed to evaluate forecasting accuracy against ground-measured solar irradiance, our model exhibited superior performance compared to the Conv-LSTM [6], a video prediction method and LFM, a type of HRES-NWP model by the Japan Meteorological Agency.

2 Methodology

Analysis of initial fields The initial fields are created from data obtained by the Himawari-8/9 meteorological satellite, operated by the Japan Meteorological Agency. Initially, the cloudiness index is calculated using the following formula:

$$\text{CloudinessIndex} = 1 - \frac{\text{EstimatedSolarIrradiance}}{\text{ClearskySolarIrradiance}} \quad (1)$$

The cloudiness index indicates the degree of attenuation of incident solar flux from outside the atmosphere and is a variable corresponding to cloud thickness. Additionally, the estimated solar irradiance is calculated using the model that utilizes satellite observation data, neural networks, and radiative transfer theory [7]. The clear-sky solar irradiance can be theoretically calculated from the position of the sun [8]. Satellite observation data provides two-dimensional spatial information, yet clouds are distributed in three dimensions. Infrared images from satellite observations, containing information about cloud heights [9], enable the vertical separation of cloudiness indices through segmentation based on infrared intensity. In this study, a segmentation model based on color segmentation [10] was employed to classify the cloudiness indices into four distinct atmospheric layers. Short-term cloud changes, typically occurring within an hour, are primarily driven by advection, which depends on wind speed and atmospheric pressure. In this context, the spatial distribution of wind speed and atmospheric pressure for each vertical layer can be inversely inferred from sequences of cloudiness index images. The direct use of satellite images, in contrast to data assimilation in NWP models, contributes to improvement of the accuracy in initial field estimations. In this study, a vision transformer [11] was employed to estimate wind speed and atmospheric pressure for the initial fields.

Atmospheric dynamics To describe cloud advection, we introduce equations of atmospheric dynamics. In this paper, the quasi-compressible approximation [12, 13] is applied, and for simplicity, vertical motions and time evolution of atmospheric pressure are assumed to be negligible. The governing equations are as follows:

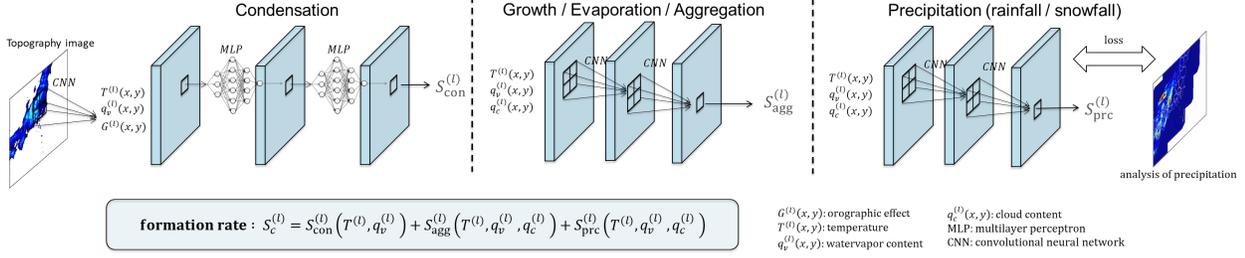


Figure 2: A schematic view of the parameterization of cloud microphysics.

$$\frac{\partial \mathbf{u}^{(l)}}{\partial t} = -\alpha^{(l)}(\mathbf{u}^{(l)} \cdot \nabla)\mathbf{u}^{(l)} - \beta^{(l)}\frac{\nabla p^{(l)}}{\rho^{(l)}} - 2\gamma^{(l)}\mathbf{f} \times \mathbf{u}^{(l)}, \quad \frac{\partial p^{(l)}}{\partial t} = 0$$

$$\frac{\partial q_c^{(l)}}{\partial t} = -(\mathbf{u}^{(l)} \cdot \nabla)q_c^{(l)} - S_c^{(l)}(T^{(l)}, q_v^{(l)}, q_c^{(l)}), \quad l = 0, 1, 2, 3$$

In these equations, $\mathbf{u}^{(l)}$, $p^{(l)}$, $q_c^{(l)}$, $q_v^{(l)}$, $T^{(l)}$, S_c , \mathbf{f} and l represent wind speed, atmospheric pressure, cloud water content, water vapor content, temperature, cloud formation rate, the Coriolis parameter, and indices of the vertical layers, respectively. Additionally, $\alpha^{(l)}$, $\beta^{(l)}$ and $\gamma^{(l)}$ are correction factors accounting for effects of orography and latitude, which are optimized through training. Ideally, to accurately describe phenomena such as cumulus cloud formation, vertical motions should be considered. In this paper, we use predicted values of temperature and humidity from HRES-NWP as input data for estimating the cloud formation rate, thereby indirectly considering vertical motions. Although the accuracy of temperature and humidity predictions is same as NWP, the cloud microphysics are optimized through a neural network model, potentially improving the forecasting accuracy in cloud formation and dissipation.

Cloud microphysics It is essential to formulate cloud microphysics to calculate cloud formation rate S_c , though the processes involved are highly complex [14]. In this paper, for simplicity, cloud microphysics are formulated using three sub-models as illustrated in Fig. 2. The first is a model related to condensation. When a moist air mass rises due to low pressure or orographic convection, clouds are formed by condensation. It is assumed that S_c due to the condensation process can be obtained by inputting temperature, water vapor content, and orographic factors into a multilayer perceptron (MLP). The orographic factor, introduced to represent the influence of terrain, is defined as feature maps generated through the convolution of a topographic image. The second sub-model focuses on aggregation. Cloud particles collide and aggregate, leading to accelerated cloud growth. This aggregation effect is represented by convolutions of the computational grid through a convolutional neural network (CNN) with a narrow receptive field. Temperature, water vapor content, and cloud water content are used as inputs, from which S_c due to the aggregation process can be computed. Here, the input cloud water content is estimated through a linear transformation of the cloudiness index. This model can also represent the growth of cloud particles due to water vapor adsorption and the evaporation process. The last is a model related to precipitation. The architecture of the model is the same as that for the aggregation process, but the output, which corresponds to the precipitation intensity, is optimized against the analysis of precipitation created from radar observations. This enhances the representation of precipitation and prevents the overgrowth of clouds. The outputs of these sub-models are accumulated to calculate the overall cloud formation rate.

Training and inference The input data for the network consist of cloudiness indices and infrared images at 20-minute intervals over the past hour from the initial time. To forecast up to six hours ahead, time evolution calculations are performed in 12 steps at 30-minute intervals, starting from the initial values. At each time step, the cloudiness indices for each vertical layer are cumulated on each grid to derive the final forecast values of the cloudiness index, which are then used as input data for the subsequent time step. During the training, the network is optimized by backpropagating the losses of the cloudiness index and precipitation. The comprehensive loss function is formulated as follows:

$$\mathcal{L} = \frac{1}{N_{\text{step}}} \sum_{t=t_0}^{t_0+N_{\text{step}}} \|\hat{C}(\mathbf{x}, t) - C(\mathbf{x}, t)\|^2 + \frac{\lambda}{N_{\text{step}}/2} \sum_{t=t_0+1}^{t_0+N_{\text{step}}/2} \left\| \sum_{\tau=t}^{t+1} \hat{R}(\mathbf{x}, \tau) - R(\mathbf{x}, t) \right\|^2$$

Here, \hat{C} and C denote the forecasted and estimated value of the cloudiness index, while \hat{R} and R represent the normalized forecasted and analyzed value of precipitation, respectively. The hyperparameter λ is set to 1.0 in this study. Additionally, the forecasted precipitation is time-integrated to match the accumulation time (1 hour) of the analysis.

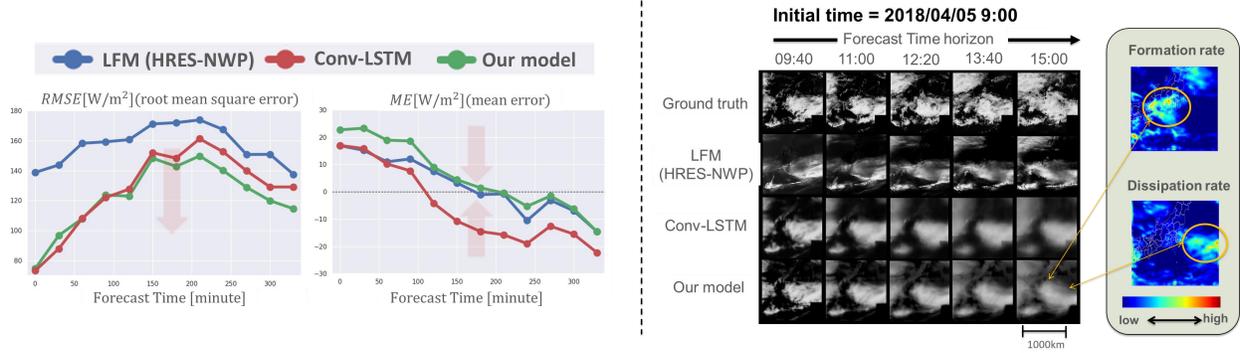


Figure 3: Evaluation of forecast models. Left: comparison of solar irradiance forecasting errors up to 6 hours ahead among our model, ConvLSTM and LFM. Right: comparative visualization of cloudiness Index distribution for forecasts starting from 2018/04/05 at 9:00 JST, with the colormap highlighting the cloud formation and dissipation rates as predicted by our model.

During the inference, the forecasted values of the cloudiness index, combined with the clear-sky solar irradiance, are converted into the solar irradiance forecast values using Eq (1).

3 Results

The proposed model was trained using data from the past four years. To verify its accuracy, forecasts were conducted up to six hours ahead, using 9:30 a.m. as the initial time in Utsunomiya, Japan. The Root Mean Square Error (RMSE) and Mean Error (ME) between the predicted and measured solar irradiance were calculated for each forecast time. The errors of our model, conv-LSTM, and LFM are shown on the left in Fig. 3.

According to the results of RMSE, our model outperformed NWP at all forecast times, particularly in the shorter forecast periods, which could be attributed to the lower accuracy of the initial values in NWP. In fact, the temporal evolution of the cloudiness index for a specific event (on the right in Fig. 3) indicated that HRES-NWP suffered from poor accuracy in initial fields, adversely affecting subsequent forecasts.

In comparison with conv-LSTM, our model exhibited superior performance in terms of RMSE from 2-3 hours onward within the forecast period. Additionally, the significantly lower ME observed in conv-LSTM suggested that its spatial distribution of solar irradiance was blurred, resulting in underestimation during clear sky conditions. As illustrated on the right in Fig. 3, conv-LSTM failed to accurately reproduce cloud formation and dissipation for a specific event. Moreover, in other events (not illustrated in this paper), conv-LSTM inadequately described advection under vertical wind shear and overestimated cloud thickness during precipitation. On the other hand, our model could reproduce these phenomena to some extent. These observations suggest that video prediction methods may struggle to accurately describe complex cloud dynamics, potentially due to insufficient training data. This limitation can lead to inaccuracies and a blurred distribution in the representation of cloud changes.

Overall, our model demonstrated the most superior performance. This can be attributed to the method of accurately estimating the initial fields and the equations for time evolution based on meteorology, which are coherently integrated through a deep learning framework.

4 Conclusions

In this study, we have developed a model using information about cloud dynamics and a deep learning framework that accurately estimated initial fields and described cloud advection, and formation / dissipation processes. The developed model demonstrated superior performance compared to HRES-NWP and video prediction method for forecast times up to six hours ahead. The improved accuracy is expected to contribute to the effective operation of photovoltaic power generation systems, enabling better utilization of solar energy.

In the future, we plan to further improve the model by incorporating the temporal evolution of temperature and humidity, as well as vertical motion, into the equations. This will enable the construction of a model completely independent from NWP, facilitating data assimilation processes and ensemble forecasting. These enhancements are expected to lead to further improvements in forecasting accuracy and the creation of probabilistic forecasting information.

Acknowledgments

This study is based on results obtained by the project “development of solar radiation forecasting technologies for short-term forecast of photovoltaic power (Research on short-term forecast of solar radiation)”, funded by the New Energy and Industrial Technology Development Organization (NEDO).

References

- [1] Ping Wang, Rudolf van Westrhenen, Jan Fokke Meirink, Sibbo van der Veen, and Wouter Knap. Surface solar radiation forecasts by advecting cloud physical properties derived from Meteosat second generation observations. *Solar Energy*, 177, pages 47–58, 2019.
- [2] Isabel Urbich ,Jörg Bendix, and Richard Müller. The seamless solar radiation (SESORA) forecast for solar surface irradiance method and validation. *Remote Sensing*, 11, pages 2576, 2019.
- [3] Atsushi Hashimoto ,Akira Usami, and Hiromu Kobayashi. Development of a solar irradiance estimation and forecasting system using Himawari-8 - Assessing the accuracy over Kyushu district for one year -. Central Research Institute of Electric Power Industry, N18003, 2019 (in Japanese).
- [4] Akshay Punjabi and Pablo Izquierdo-Ayala Efficient spatio-temporal weather forecasting using U-Net. In *Proceedings of the CIKM 2021 Workshops*, 3052, 2021.
- [5] Chibuzor N Obiora and Ahmed Ali Effective Implementation of Convolutional Long Short-Term Memory (ConvLSTM) Network in Forecasting Solar Irradiance. In *IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society*, pages 1–5, 2021.
- [6] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong and Wang-chun Woo Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In *Advances in neural information processing systems*, pages 802–810, 2015.
- [7] Jun Sasaki, Kenji Utsunomiya, Maki Okada, Shigeyuki Yoshikawa and Koji Yamaguchi Estimation of solar radiation using satellite images and a data-driven radiative transfer model. In *33rd International Photovoltaic Science and Engineering Conference*, 2022
- [8] Junsei Kondo Meteorology of water environment. Tokyo:Asakura Publishing, pages 55–91, 1994.
- [9] Akihiro Shimizu , Koutarou Saito and Mikito Yamamoto Image characteristics of the 16 bands of Himawari-8’s AHI. Technical Note 62, Tokyo: Meteorological Satellite Center, <https://www.data.jma.go.jp/mscweb/technotes/msctechrep62-3.pdf>, 2017.
- [10] Naofumi Akimoto, Huachun Zhu, Yanghua Jin, and Yoshimitsu Aoki Fast soft color segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8277–8286, 2020.
- [11] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer,Georg Heigold, Sylvain Gelly, Jakob Uszkoreit and Neil Houlsby An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021.
- [12] Joseph B. Klemp and Robert B. Wilhelmson The simulation of three-dimensional convective storm dynamics. *Journal of the Atmospheric Sciences*, 35, pages 1070–1096, 1978.
- [13] Kazuhisa Tsuboki and Atsushi Sakakibara Large-scale parallel computing of cloud resolving storm simulator. In *High Performance Computing:Zima*, H.P., Jou, K., Sato, M., Seo, Y., Shimasaki, M., Eds.; Springer: Berlin/Heidelberg, Germany, pages 243–259, 2002.
- [14] Hugh Morrison, Marcus van Lier-Walqui, Ann M. Fridlind, Wojciech W. Grabowski, Jerry Y. Harrington, Corinna Hoose, Alexei Korolev, Matthew R. Kumjian, Jason A. Milbrandt, Hanna Pawlowska, Derek J. Posselt, Olivier P. Prat, Karly J. Reimel, Shin-Ichiro Shima, Bastiaan van Dierenhoven and Lulin Xue Confronting the challenge of modeling cloud and precipitation microphysics. *Journal of Advances in Modeling Earth Systems*,12,e2019MS001689,<https://doi.org/10.1029/2019MS001689>, 2020.