

Article

A Hybrid Spatio-Temporal Prediction Model for Solar Photovoltaic Generation Using Numerical Weather Data and Satellite Images

Bowoo Kim  and Dongjun Suh * 

Department of Convergence & Fusion System Engineering, Kyungpook National University, Sangju 37224, Korea; kbw5913@knu.ac.kr

* Correspondence: dongjunsuh@knu.ac.kr; Tel.: +82-54-530-1482

Received: 13 October 2020; Accepted: 6 November 2020; Published: 12 November 2020



Abstract: Precise and accurate prediction of solar photovoltaic (PV) generation plays a major role in developing plans for the supply and demand of power grid systems. Most previous studies on the prediction of solar PV generation employed only weather data composed of numerical text data. The numerical text weather data can reflect temporal factors, however, they cannot consider the movement features related to the wind direction of the spatial characteristics, which include the amount of both clouds and particulate matter (PM) among other weather features. This study aims developing a hybrid spatio-temporal prediction model by combining general weather data and data extracted from satellite images having spatial characteristics. A model for hourly prediction of solar PV generation is proposed using data collected from a solar PV power plant in Incheon, South Korea. To evaluate the performance of the prediction model, we compared and performed ARIMAX analysis, which is a traditional statistical time-series analysis method, and SVR, ANN, and DNN, which are based on machine learning algorithms. The models that reflect the temporal and spatial characteristics exhibited better performance than those using only the general weather numerical data or the satellite image data.

Keywords: solar PV generation; spatio-temporal; prediction; ARIMAX; SVR; ANN; DNN; satellite image

1. Introduction

The issue of climate change caused by carbon emission and the depletion of fossil fuels is emerging worldwide. To address this problem, the Kyoto Protocol aimed at reducing greenhouse gas emissions for the purpose of decarbonization. Furthermore, the Paris Agreement, which strengthened global carbon regulations to suppress the global average temperature within 2 °C before industrialization, was signed [1,2]. While the development of many renewable energies are being sought to replace fossil fuels, South Korea announced its renewable energy generation by 2030 along with denuclearization in line with global trends [3]. In addition, more than 95% of the new renewable energy facilities use clean energy such as solar PV and wind power.

In the past decade, the unit price of solar panels and facilities for power generation systems was decreased, and there is an increase in the number of large-scale solar PV farms worldwide. Typically, the United States, Germany, and China have gigawatt-scale farms, and South Korea established a 467 MW solar PV farms in 2013 [4]. Compared to other renewable energy sources, the solar PV power generation has the advantage of low installation, and maintenance costs and an expected life of more than 20 years [5]. In addition, it is possible to minimize the damages to nature that occur when installing power plants, which is one of the emerging environmental problems. However, solar PV power generation requires a large installation area due to its low energy density, and the amount of

power generated fluctuates with meteorological factors such as a change in irradiance due to clouds or particulate matter (PM) [6,7]. This phenomenon increases the complexity of the plan for stable supply and demand of power systems, it mostly disrupts the schedule for power grid operations. In particular, the solar PV generation technology, which is considered as one of the key components in smart grids, a technology that integrates information and communication technology (ICT) into the power grid, has the advantage of generating clean and unlimited energy, however, for stable systems, there is a demand for accurate predicting technology [8]. Without establishing an accurate power supply and demand plan, here could be huge financial and social losses. For this reason, there is a rapid increase in the need for accurate prediction of the amount of solar PV power generated. Therefore, accurate prediction of the power generation of renewable energy sources is very important in establishing an efficient power supply and demand plan.

Most of the previous studies on the prediction of solar PV generation can be divided into two categories. The first category involves the use of numerical text weather data such as the most common irradiance, temperature, and precipitation [9–13]. This method increases the predictive accuracy of the power generation by taking advantage of the continuity over various times in the data. The second category uses motion vectors or indexes of irradiance and clouds in satellite images [14–18]. However, this approach sometimes ignores the physical information of the solar PV generation system [19]. The most directly affected factor in predicting solar PV power generation is the irradiance, which is greatly influenced by the cloud shadow. The cloud shadow can reflect the increase or decrease in irradiance by detecting the cloud motion vector through the movement of the cloud shown in the satellite image.

Recently, air pollution caused by PM has been another environmental issue in South Korea [20]. The increase in the concentration of PM in the atmosphere not only has a fatal effect on the human body, such as the respiratory organs, but also reduces the irradiance reaching solar panels due to the scattering of the solar radiation [21]. Most of the previous studies analyzed the effects of accumulation various types of dust as well as PM on solar panels [22–26], however, this study aims at investigating the effect of PM concentration distribution in the atmosphere on the solar PV generation. Both clouds and PM vary with time and affect the atmosphere because of the spatial characteristics shifted from the adjacent region to the measurement point in the wind direction. Therefore, using satellite images, we investigated the influence on the spatial characteristics of the cloud and PM. To reflect the spatial characteristics, the area where the solar PV power plant is located was designated as a region of interest (ROI) and then the adjacent area in eight directions was set as the adjacent region of interest (ROI_{adj}). To investigate the effect of the movement of clouds and PM according to the direction of the wind moving from ROI_{adj} to ROI, the amounts of clouds and PM extracted from the satellite image were first predicted and used as variables for the solar PV generation prediction model. We propose a hybrid solar PV generation forecasting model, which combines the numerical weather data composed of texts with spatial information extracted from satellite images. To develop the proposed model that reflects the temporal and spatial characteristics, a comparative analysis was conducted by dividing the experiments into three groups. The first group uses the numerical text weather data from the Korea Meteorological Administration (KMA), the second group uses the data extracted from the satellite images, and the third group combines the numerical weather data from the KMA and those from the satellite images. The proposed model uses the power generation data from a solar PV power plant located in Incheon, South Korea, to predict the amount of PV generation after one hour. The prediction model employs the autoregressive moving average with exogenous input (ARIMAX), which combines external factors in autoregressive moving average (ARIMA) considering the time-series among the traditional statistical analysis methods. It also uses Support Vector Regression (SVR) and artificial neural network (ANN), which employ machine learning algorithms that have been recently used actively in various fields. In addition, the Deep Neural Network (DNN) with additional hidden layers in ANN was analyzed.

Research Framework

In this study, the numerical text weather data from KMA and the data extracted from satellite images were used together to develop a hybrid spatio-temporal model that not only considers the temporal characteristics of the input parameter over the time flow but also reflect the spatial characteristics. The research framework is shown in Figure 1. The first step in the framework is the collection and preprocessing of the data provided by KMA and the Korea National Meteorological Satellite Center (NMSC). The second step involves the extraction of the necessary data from each of the four satellite images collected. The wind direction and wind speed in the Atmospheric Motion Vector (AMV) image, the amount and thickness of the cloud in the cloud optical thickness image, the amount and concentration of PM in the aerosol optical depth image, and the irradiance from the insolation image were extracted. The third step involves setting the ROI of the desired region in the image and then designating the ROI_{adj} of the same size for the 8 directions adjacent to each edge and vertex. A solar PV power plant located in Incheon, South Korea, was analyzed and set as the ROI. Then, based on the wind direction information of the ROI, the effect of cloud and PM movement from the ROI_{adj} to the ROI was analyzed, and the amount of cloud, cloud thickness, amount and concentration of PM in the ROI were predicted. In the fourth step, the data extracted from satellite images were analyzed together with numerical text weather data collected from KMA to perform the preprocessing process for the development of hybrid spatio-temporal models. In the last step, to develop solar PV generation forecasting analysis models that are based on the ARIMAX, SVR, ANN, and DNN methods, the prediction performance of the model was optimized by choosing optimal parameters for each technique.

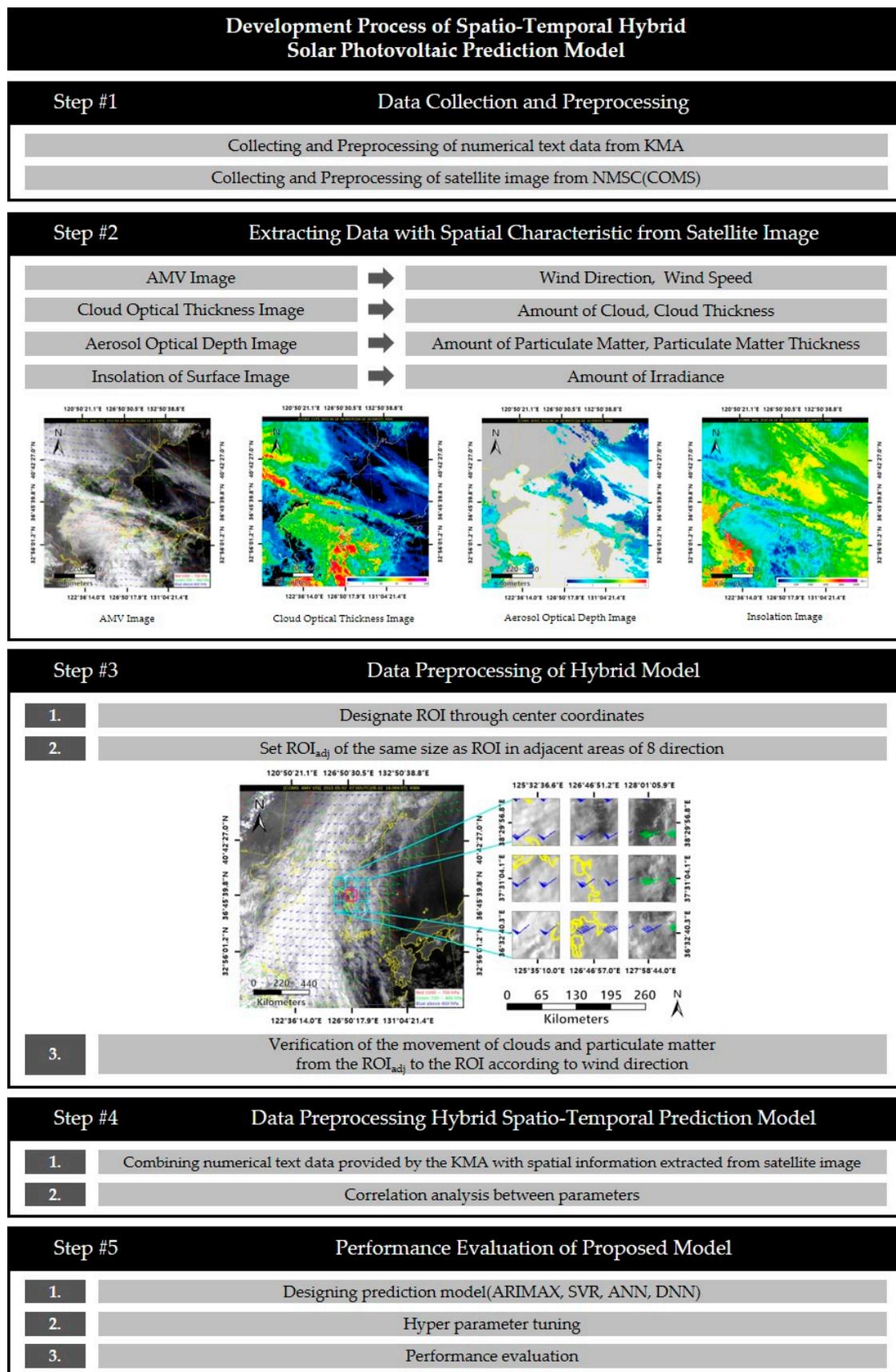


Figure 1. Research framework.

2. Methodology

2.1. Numerical Text Data

To predict one hour ahead solar PV generation, the numerical text data used in the experiment are classified into three categories: meteorological data, air pollution data (such as PM) and solar PV generation. The numerical text weather data were provided by Incheon meteorological station, located at 37.47772 lat. and 126.6249 long. in KMA [27]. KMA provides public data on more than 15 types of time-based weather data, including temperature, precipitation and wind speed, for 103 stations across the country. In this study, temperature, precipitation, wind speed and direction, humidity, amount of sunshine, irradiance, cloudiness, and visibility were used as the input parameters to predict the amount of solar PV generated in Incheon. The air pollution data were provided at an Air Korea air pollutant station located within 3 km from the test location [28]. The rapid increase in the use of fossil fuels has raised the issue of air pollution, and the recent increase in the level of air pollutions has great environmental and economic effects in Korea. As the concentration of PM, such as PM₁₀ and PM_{2.5}, in the atmosphere increases, the solar radiation reaching the Earth from the Sun is scattered in the atmosphere, reducing the visibility and irradiance reaching the Earth surface. To analyze the air pollutants expected to affect the performance of solar PV generation, pollutants such as SO₂, CO, O₃, NO₂, PM₁₀, and PM_{2.5} were considered together with the meteorological data. Lastly, the solar PV generation data was provided by the Open Data Portal [29]. The analyzed solar PV power plant is located in Incheon, and the data on the amount of solar PV power generation were collected on an hourly basis from 0:00 on 1 January to 23:00 on 31 December 2015. K-NN algorithm was used for the interpolation of missing data for all parameters. This algorithm selects the nearest value and does not consider the values of other neighboring points, a constant interpolation can be generated. Therefore, interpolation was performed using 36 h of data before and after the data point at the time of missing. As mentioned above, solar PV generation is highly affected by the irradiance, so only the data from 9:00 to 17:00, which is the daylight time, were used. Table 1 shows the numerical text data samples.

Table 1. The sample of the numerical dataset.

Date	Temperature [°C]	Precipitation [mm]	Wind Speed [m/s]	Wind Direction [0–360 degree]	Humidity [%]	Amount of Sunshine [hr]	Irradiance [MJ/m]	Cloudiness [0–10 level]	Visibility [10m]	SO ₂ [ppm]	CO [µg/m ²]	O ₃ [ppm]	NO ₂ [ppm]	PM ₁₀ [µg/m ²]	PM _{2.5} [µg/m ²]	PV [kW]
1 January 2015 09:00:00	−8.4	0	6.7	340	56	0.8	0.21	0	2000	0.006	0.5	0.017	0.012	145	33	60
1 January 2015 10:00:00	−8.1	0	6.1	226	54	0.0	0.67	1	2000	0.006	0.5	0.019	0.01	117	34	374
1 January 2015 11:00:00	−7.6	0	6.1	340	53	0.0	1.1	1	2000	0.006	0.6	0.019	0.01	98	33	638
1 January 2015 12:00:00	−6.9	0	6.4	340	52	0.0	1.41	1	2000	0.006	0.6	0.021	0.01	90	30	784
1 January 2015 12:00:00	−6.1	0	6.4	340	53	0.0	1.53	1	2000	0.006	0.6	0.023	0.01	85	27	842
.
.
31 December 2015 13:00:00	2.9	0.0	3.3	360	74	1.0	1.38	4	800	0.011	1.2	0.013	0.042	78	66	230
31 December 2015 14:00:00	3.3	0.0	3.1	360	75	1.0	1.24	3	800	0.011	1.2	0.023	0.032	100	72	310
31 December 2015 15:00:00	3.1	0.0	3.4	340	77	1.0	0.93	3	800	0.011	1.2	0.024	0.034	87	67	439
31 December 2015 16:00:00	3.3	0.0	3.2	340	77	1.0	0.67	2	900	0.009	1.2	0.024	0.035	90	68	303
31 December 2015 17:00:00	2.9	0.0	2.1	320	78	1.0	0.26	0	700	0.009	1.2	0.017	0.047	83	65	95

2.2. Satellite Image Data

In this study, satellite images provided by the NMSC were used [30]. The satellite images were provided by the Communication, Ocean, and Meteorological Satellite (COMS), which was launched on 27 June 2010. It is the first geostationary combined satellite in Korea that performs ocean and meteorological observations and communication service mission. Detailed sensor information of the COMS is summarized in Table 2. The COMS provides more than 16 types of image data, including raw images, basic images, and processed images, every 15 min for North-East Asia and the Korean Peninsula. Image data of AMV, including cloud optical thickness image, aerosol optical depth image, and insolation image, were used in this study [31–34]. Each image shows the Korea Peninsula in a size of 1024×1024 and has a ground resolution of 1720.8 m per pixel. Figure 2 shows four types of images used in this experiment, which were provided at 15:00 on 18 April 2015. Detailed information and the data extraction method for each image are described in the subsequent subsections.

Table 2. Basic performance data of COMS.

Channel	Center Wavelength (μm)	Wavelength Band (μm)	Spatial Resolution (km)
Visible	0.67	0.55~0.8	1
Shortwave Infrared	3.7	3.5~4.0	4
Water vapor	6.7	6.5~7.0	4
Infrared 1	10.8	10.3~11.3	4
Infrared 2	12.0	11.5~12.5	4

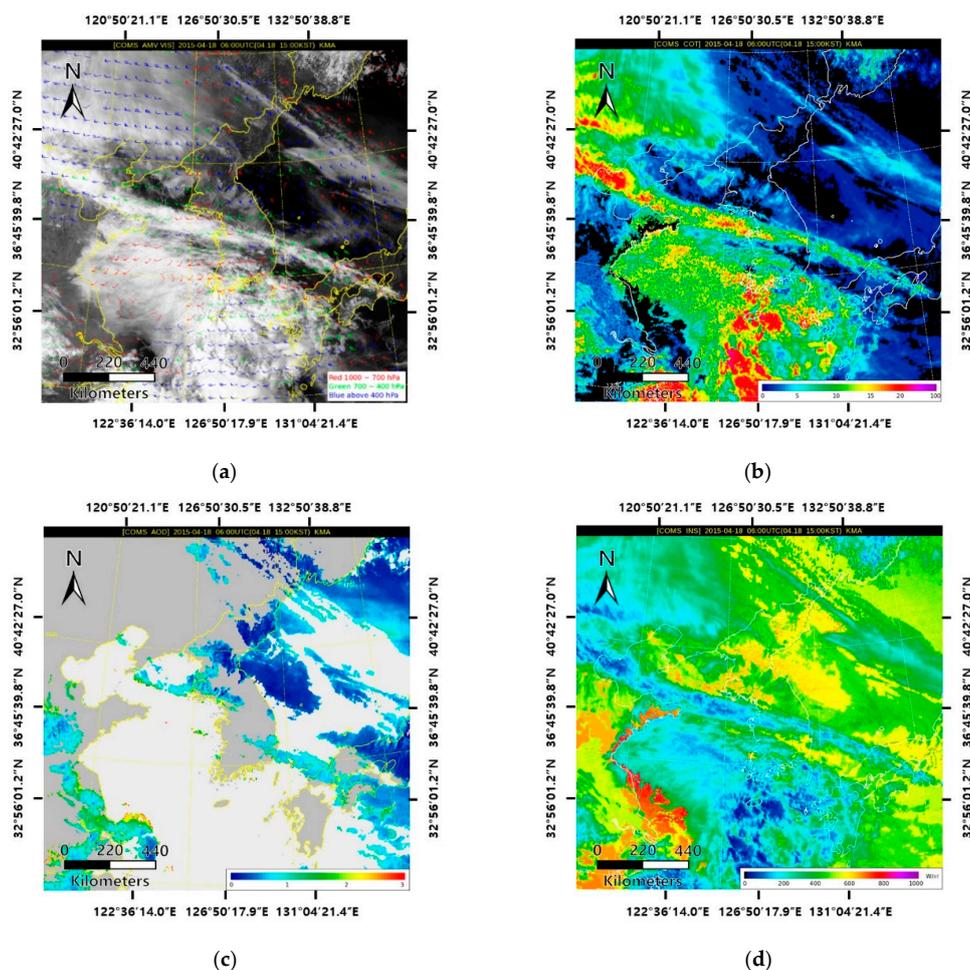


Figure 2. Four types of satellite images: (a) AMV Image; (b) Cloud Optical Thickness Image; (c) Aerosol Optical depth Image; (d) Insolation Image.

2.2.1. Atmospheric Motion Vector (AMV) Image

Figure 2a shows the AMV image, which depicts the information about the movement of the atmosphere, including the information on wind direction and wind speed. The ROI has to be set before extracting the information of the wind direction and speed from the image. The numerous wind direction arrows shown in the AMV image have fixed starting point of the arrow and are referred to as center point in this paper. In order to set the ROI, the center point that exists at the Incheon, South Korea as the target point is selected. As the wind direction changes over time, the angle of the arrow indicating the wind direction in the next time step image is changed, but the coordinates of the center point are fixed because the arrow is rotated based on the center point. Therefore, the ROI was set at 50×50 , a size that does not interfere with the rotation radius of the arrow when the arrow rotates 360° to the center point. Figure 3 shows the set ROI in magenta.

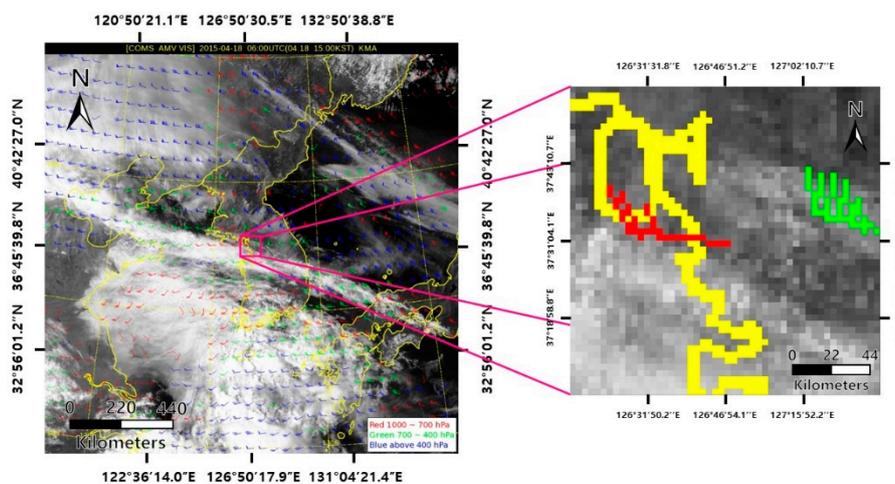


Figure 3. An exemplary satellite image of AMV with ROI.

The wind direction vector information is not provided in the AMV image. Thus, to extract the wind direction information, other previous studies have extracted wind direction vector information from images [14]. In this study, the wind direction arrows are extracted from images and calculated as described in the following sequence. To obtain information on the wind direction, first, the coordinates of the tail point is found using the value of the pixel located farthest from the center point among the pixels that constitute the arrow. Next, the farthest pixel from the center point and tail point at the same time among the pixels that constitute the arrow is designated as the endpoint. Then, the wind direction can be acquired by calculating the angle with respect to the north direction of the straight line connecting the obtained center point and the endpoint. Figure 4 shows a single AMV vector, which includes the information on the wind direction, wind speed, center point, endpoint, and tail point. The process of extracting information on the wind speed is as follows: first, a straight line is drawn parallel to the arrow body in about the distance of the endpoint in the direction of the tail point; secondly, the wind speed is calculated primarily by counting the number of pixels that overlap between the newly drawn straight line and the existing wind direction arrow; thirdly, another new straight line is drawn in the middle of the distance between the arrow body and the tail point; lastly, by calculating the number of overlapping pixels in the same way as in the second step, it is determined whether there is a tail segment of 2 m/s, and the wind speed is finally determined by calculating whether there is a triangle representing 25 m/s.

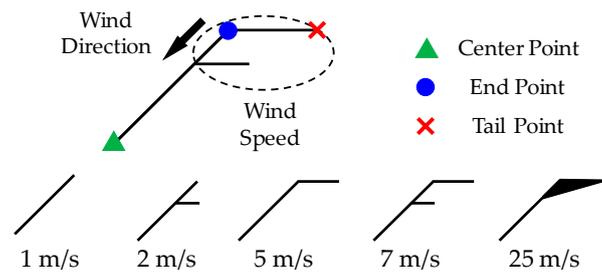


Figure 4. A standard station model for wind direction and speed with the location of each point.

2.2.2. Cloud Optical Thickness, Aerosol Optical Depth, and Insolation Image

Figure 2b shows the Cloud Optical Thickness Image, which presents the cloud thickness in a color index. Various studies are being conducted to analyze the movement and amount of clouds on satellite images through cloud motion vector [35]. However, in this study, the color index of the image was used to predict the amount of clouds. To get information on the amount and thickness of the cloud, the index from 0 to 100 was divided into four levels: clear, partly cloudy, mostly cloudy, and cloudy. Before extracting the data to analyze the spatial characteristics, ROI_{adj} that is of the same size as the ROI was designated for 8 directions of the areas adjacent to the ROI adjacent in the AMV image. The ROI_{adj} were denoted as S_{1-8} in order from the top left. To extract the amount and thickness of clouds present in the ROI_{adj} and ROI, the pixels in each region are divided into four color indexes, and the number of pixels per index was recorded. Figure 5 shows the set ROI_{adj} in cyan.

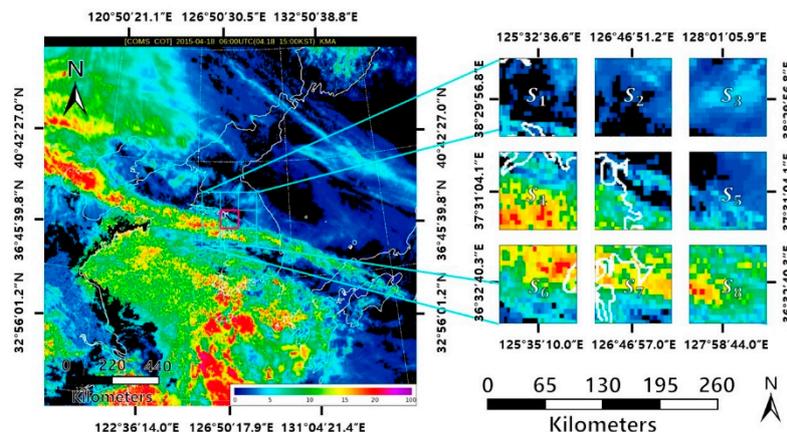


Figure 5. An exemplary satellite image of the Cloud Optical Thickness with sub ROI.

Figure 2c shows the Aerosol Optical Depth Image, which represents the aerosol components, such as yellow dust and PM, in a color index. As in the Cloud Optical Thickness image, the color index is divided into good, moderate, unhealthy, and very unhealthy, and the number of pixels was recorded.

Figure 2d shows the Insolation Image, which represents the irradiance reaching the ground surface in a color index. To extract the irradiance in the ROI, the amount of insolation in the range of 0–1000, which is the size of the color index, was readjusted to the range of 0–488, which is the pixel size of the color index in the insolation image. After that the average value of the numerical index in the ROI was recorded.

The satellite images taken at the NMSC are provided every 15 min, which involves the time-series continuity. However, to combine with the numerical text weather data from KMA, an image of the same time-scale of 1 h was extracted and used. As with the numerical text weather data, the data extracted from the satellite images had temporal continuity, as shown in Figure 6. Finally, the values of wind direction, wind speed, cloud thickness, concentration of PM, and irradiance, which are parameters every hour extracted from each satellite image, were added to the numerical text weather data set.

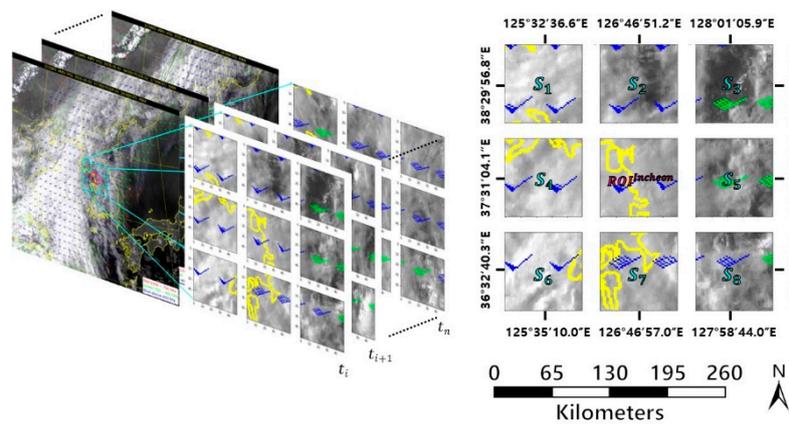


Figure 6. Continuity of satellite images with the change in time.

2.3. Correlation Analysis with the Wind Direction of the Cloud and PM

Clouds and PM in the atmosphere have spatial characteristics that move in the wind direction [36]. Therefore, in order to demonstrate the reflection of spatial characteristics in the prediction of solar PV generation, the mobility of the clouds and PM according to the wind direction from the ROI_{adj} to ROI_{adj} was numerically verified and expressed as Algorithm 1. The sample datasets used for the Algorithm 1 are shown in Tables 3 and 4, and the composition of Algorithm 1 follows the following sequence: the wind direction of the ROI at time t is identified; the amount of cloud and PM at time t of each ROI_{adj} and ROI are compared and analyzed; the increase or decrease of clouds and PM due to the movement of clouds and PM at time $t + 1$ of the ROI according to the wind is determined. For example, assuming that the wind direction at the ROI at time t is northwest and the amount of clear cloud in S_1 is greater than that in the ROI ($ROI^{Incheon}$), then, the amount of clear cloud of ROI increases at time $t + 1$, it is determined as true about movement of clouds and the opposite case as false. Then, the accuracy is derived by calculating the number of true and false for the entire time. PM also uses the same method as Algorithm 1. Tables 5 and 6 list the results of the determined mobilities of both clouds and PM with respect to the wind direction. PM has high accuracy in all the cases, but clouds have high accuracy only in the mostly cloudy and cloudy levels. In the case of partly cloudy and cloudy, the accuracy is considered inferior because there is a possibility of natural extinction over time. Therefore, it can be determined that clouds and PM, which have a great influence on the fluctuations in irradiance and solar PV generation, have mobility according to the movement of the wind.

Algorithm 1. Algorithm of discriminant for movement of clouds and PM by wind direction.

Denotes $WD = \text{Wind Direction}$; $S_{ROI} = ROI^{Incheon}$; $S_i = ROI_{adj}$; $N = \text{number of samples}$; $c_1 = \text{number of 1 in conditional for}$.

1: Determination of i by identification of WD_i in the S_{ROI}

2: $WD = \{NW, N, NE, W, E, SW, S, SE\}$

3: Determination of the S_i through i

4: $S = \{S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8\}$

5: Initialize the time step $t = 1$

6: Comparison of the amount of cloud in $S_{i,t}$ at t and the amount of cloud in the $S_{ROI,t}$

7: Comparison of the amount of cloud in $S_{ROI,t}$ and $S_{ROI,t+1}$

8: For $n = 1$ to N

9: If $S_{i,t} \geq S_{ROI,t}$

10: If $S_{ROI,t+1} \geq S_{ROI,t}$

11: True

12: Else

13: False

14: Else

15: If $S_{ROI,t+1} \leq S_{ROI,t}$

16: True

17: Else

18: False

19: Accuracy (%) = $\frac{c_1}{N} \times 100$

Table 3. The example of the cloud data sets.

Date	Wind Direction	Clear	Partly Cloudy	Mostly Cloudy	Cloudy
18 January 2015 11:00:00	W	0	0	0	0
18 January 2015 12:00:00	W	1608	114	6	0
18 January 2015 13:00:00	SW	1147	935	31	0
18 January 2015 14:00:00	W	0	0	0	0
2015-08-02 08:00:00	SW	0	0	0	0
2015-08-02 09:00:00	W	175	636	165	6
2015-08-02 10:00:00	W	238	332	222	35
2015-08-02 11:00:00	W	52	591	364	0

Table 4. The example of the PM data sets.

Date	Wind Direction	Good	Moderate	Unhealthy	Very Unhealthy
18 January 2015 11:00:00	W	27	131	119	57
18 January 2015 12:00:00	W	0	88	204	78
18 January 2015 13:00:00	SW	0	4	30	20
18 January 2015 14:00:00	W	0	6	0	0
2 August 2015 08:00:00	SW	6	144	12	18
2 August 2015 09:00:00	W	0	12	12	9
2 August 2015 10:00:00	W	0	9	62	39
2 August 2015 11:00:00	W	0	0	0	18

Table 5. The result of the discriminant for movement of clouds by wind direction.

Cloud	Clear	Partly Cloudy	Mostly Cloudy	Cloudy
Accuracy (%)	75.068	75.793	84.044	91.296

Table 6. The result of discriminant for movement of PM by wind direction.

PM	Good	Moderate	Unhealthy	Very Unhealthy
Accuracy (%)	87.489	85.585	89.665	93.382

3. Forecasting Method of Solar PV Generation

3.1. Prediction of Cloud and PM in ROI

Before directly predicting solar PV generation, the concentration of cloud and PM were first predicted from the image data to reflect the temporal and spatial effects of clouds and PM. There are four variables in both clouds and PM (cloud: clear, partly cloudy, mostly cloudy, cloudy; PM: good, moderate, unhealthy, very unhealthy). Next, the data for predicting the amount of cloud and PM in units of an hour about ROI using the above variables. To predict the amount of the clouds and PM, the SVR using the linear kernel was employed, and the data from January to December 2015, which is the entire experimental period, were randomly sampled and trained 60% of each month and predicted 40% of each month to reflect seasonal characteristics.

3.2. Proposed Models for the Prediction of Solar PV Generation in ROI

The solar PV generation was predicted using ARIMAX, which is a time-series analysis method, and SVR, ANN, and DNN, which are machine learning algorithms widely used in various fields. Each method is described in detail as the subsequent subsections.

3.2.1. Autoregressive Moving Integrated Average Exogenous input (ARIMAX)

The ARIMAX is developed by Newsham and Birt [37], and is a method in which external factors act as additional variables in the ARIMA, which is a traditional statistical time-series analysis method [38]. The ARIMA is a method that satisfies both the autocorrelation model (AR) and the moving average model (MA) at the same time. The AR determines whether the past data affect the future data, and the MA identifies the tendency for the average value of any random variable to increase or decrease continuously over time. The ARIMA is a technique in which differences (I) are added in ARMA, and it can eliminate abnormalities by applying the initial differencing step. The ARIMAX is mainly employed when the variables to be predicted have a time series or a periodicity, and is mainly used for short-term solar PV generation forecasting [39,40]. Since the meteorological and air pollutant variables, including solar PV generation, provide as an hour data that satisfies the time-series characteristic. The ARIMAX has the order of p , d , and q , which represent the autoregressive order, difference order, and moving average order, respectively. In this study, the ARIMAX with an order (1, 0, 1) was used.

3.2.2. Support Vector Regression (SVR)

The SVR is an extended method for regression prediction in the Support Vector Machine (SVM) proposed by Vapnik in 1995 [41]. The key algorithm of SVR is to learn the training data and then find the optimal regression function $f(x)$ in which all the predicted values exist within ϵ , a specific deviation called the support vector, and the error values of the predicted result are minimized. In general, the datasets used for the actual prediction cannot be solved with only linear problems of one dimension. To this end, vectors existing in high dimensions can be calculated linearly using kernel functions such as RBF, linear and polynomial kernels. The regression function $f(x)$ can be calculated by the Lagrangian method using the value of the calculated dot product. Compared to other techniques, the SVR is more generalization ability and is widely used to solve regression problems. Also, due to the global minimum value, it has remarkable performance in time-series analysis, so it is widely used in solar PV generation prediction research [42,43]. In this study, a linear kernel was used for both the prior prediction model for predicting the cloud and PM and the prediction model for predicting solar PV generation.

3.2.3. Artificial Neural Network (ANN)

ANN, which is actively used for classification and prediction in various fields, is an artificial neural network that mimics the structure of the human brain. It is composed of input, hidden, and output layers [44,45]. Like the human brain has numerous neurons connected to collect and process data, ANN has interconnected nodes in each layer. The output value for input is predicted by the activation function included in each layer. In the prediction process, the activation function of each layer mainly calculates the function value and the first-order derivative, and the learning time depends on the corresponding calculation process. Herein, the proposed model has one hidden layer, and the Relu function was used as the activation function for the solar PV generation prediction model.

3.2.4. Deep Neural Network (DNN)

DNN is used to expand the number of hidden layers in ANN, as shown in Figure 7 [46,47]. The ANN contains a single hidden layer between input and output layers whereas the basic form of the DNN can have one or more hidden layers. As the number of hidden layers increases, the computational complexity also increases, but a combination of nonlinear transformation techniques can lead to high

prediction accuracy. As in ANN, d nodes $X = (1, x_1, x_2, \dots, x_d)^T$, excluding the bias nodes, are inputs and $O = (O_1, O_2, \dots, O_c)^T$ with c nodes are outputs. Then, the number of nodes on the hl th layer is denoted as n_j [48].

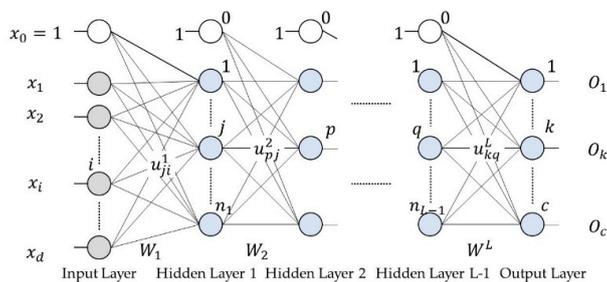


Figure 7. Structure of the DNN model.

The weight W that connects the $(L - 1)$ th layer to the L th layer is given by $(n_{hl-1} + 1)*n_l$, and it is represented by the matrix W^{hl} , as expressed in Equation (1). As mentioned earlier, the number of added hidden layers in ANN can be expressed as Equation (2), and input parameter vector X is represented as an internal parameter Z^0 , as expressed in Equation (3), when applied in DNN.

In the operation of DNN, the sum of the product of the edge weights connected to the Z vector and j th node is first obtained and then stored in the S variables. The values of S and Z variables are stored independently because they are used in the later applied backpropagation algorithm. In this case, the computation of the j th node of the hl th layer is represented by Equation (4), where $n_{hl} - 1$ is the number of nodes in the $(hl - 1)$ th layer. Equation (4) is only for the calculation of a single node, but Equation (5) allows all operations of the hl th layer to be performed simultaneously. The activation function $\tau_1 - \tau_{L-1}$ included in each layer mainly uses such functions as logistic, sigmoid, and ReLU functions, and the softmax function is mainly used in the output layer.

In this study, the tanh, ReLU, and sigmoid functions were applied to the hidden layer and the sigmoid function was applied to the final output layer. Detailed information about each hidden layer is presented in Table 7. In addition, to avoid overfitting, a dropout layer was added, which omits and calculates some of the nodes. Nodes omitted from the dropout do not affect learning, thus, the problem of overfitting can be avoided.

$$W^{hl} = \begin{pmatrix} w_{10}^{hl} & w_{11}^{hl} & \dots & w_{1n_{hl-1}}^{hl} \\ w_{20}^{hl} & w_{21}^{hl} & & w_{2n_{hl-1}}^{hl} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n_{hl}0}^{hl} & w_{n_{hl}1}^{hl} & \dots & w_{n_{hl}n_{hl-1}}^{hl} \end{pmatrix}, \quad hl = 1, 2, \dots, L, 1 \leq hl \leq L \tag{1}$$

$$O = f(x) = f_L(\dots f_2(f_1(x))) \tag{2}$$

$$Z^0 = (z_0, z_1, z_2, \dots, z_{n_0})^T = (1, x_1, x_2, \dots, x_d)^T \tag{3}$$

$$\begin{aligned} z_j^{hl} &= \tau_l(s_j^{hl}), \\ s_j^{hl} &= u_j^{hl} z^{hl-1}, \\ z^{hl-1} &= (1, z_1^{hl-1}, z_2^{hl-1}, \dots, z_{n_{hl-1}}^{hl-1})^T, \\ u_j^{hl} &= (u_{j0}^{hl-1}, u_{j1}^{hl-1}, \dots, u_{jn_{hl-1}}^{hl-1}) \end{aligned} \tag{4}$$

$$z^{hl} = \tau_1(U^{hl} z^{hl-1}), \quad 1 \leq hl \leq L \tag{5}$$

Table 7. The structure of the DNN.

Number of Hidden Layer	1	2	3	4	5	6	7
Number of Nodes	180	0.4	100	0.4	100	0.4	1
Activation Function	tanh	Drop out	Relu	Drop Out	Sigmoid	Drop out	Sigmoid

3.3. Analytic Process for Predicting Solar PV Generation

To predict the solar PV generation, an experiment was conducted by considering three groups for the same period from January to December 2015 with respect to the solar PV power plant located at Incheon as described above. In the entire experiment, after predicting the amount of the clouds and PM of 60% per each month to reflect spatial characteristics in the prediction model. After that, remained sample data of 40% per each month was randomly chosen, and then 70% were used for training and 30% for testing. Group 1 used only numerical text data of the commonly used weather and air pollutants, and forecasting was performed hourly. Group 2 used only satellite image data to identify spatial characteristics. Thereafter, the predicted amounts of clouds and PM were used together with the data extracted from the satellite images, including the wind direction and speed, as the input parameters to predict the solar PV generation. Finally, Group 3 performed prediction by combining the numerical text data and the data extracted from the satellite images to develop a hybrid spatio-temporal prediction model. Group 3 conducted an experiment by first predicting the amount of cloud and PM in the same way as in Group 2, and then combining with numerical text data of the same period.

For all groups, the parameters for the month, day, and time, representing time, were added to reflect the temporal characteristics, and one hour previous solar PV generation data was added as the input parameters for the generation model to reflect the presentity of the prediction model. Also, to understand the effect of each group on the cloud and PM for all the groups, three cases were adopted: case 1, where only the cloud is applied as a parameter; case 2, where only PM is applied as a parameter; case 3 that applies the cloud and PM at the same time. All the groups with the cases are shown in Table 8.

Table 8. Input parameters by each group and case.

	Group 1 (Numerical Text Weather Data)	Group 2 (Satellite Images)	Group 3 (Mixed, G1 + G2)
Common Parameters	Month, Day, Time, PV (previous data)		
Case 1 (Cloud)	Temperature, Precipitation, Wind Speed, Wind Direction, Humidity, Amount of Sunshine, Irradiance, Cloudiness, Visibility	Wind Speed, Wind Direction, Clear, Partly cloudy, Mostly cloudy, Cloudy, Irradiance	Temperature, Precipitation, Wind Speed, Wind Direction, Humidity, Amount of Sunshine, Irradiance, Clear, Partly cloudy, Mostly cloudy, Cloudy, Visibility
Case 2 (PM)	Temperature, Precipitation, Wind Speed, Wind Direction, Humidity, Amount of Sunshine, Irradiance, SO ₂ , CO, O ₃ , NO ₂ , PM ₁₀ , PM _{2.5} , Visibility	Wind Speed, Wind Direction, PM_Good, PM_Moderate, PM_Unhealthy, PM_Very Unhealthy, Irradiance	Temperature, Precipitation, Wind Speed, Wind Direction, Humidity, Amount of Sunshine, Irradiance, PM_Good, PM_Moderate, PM_Unhealthy, PM_Very Unhealthy, Visibility

Table 8. Cont.

	Group 1 (Numerical Text Weather Data)	Group 2 (Satellite Images)	Group 3 (Mixed, G1 + G2)
Common Parameters	Month, Day, Time, PV (previous data)		
Case 3 (Cloud + PM)	Temperature, Precipitation, Wind Speed, Wind Direction, Humidity, Amount of Sunshine, Irradiance, SO ₂ , CO, O ₃ , NO ₂ , PM ₁₀ , PM _{2.5} , Cloudiness, Visibility	Wind Speed, Wind Direction, Clear, Partly cloudy, Mostly cloudy, Cloudy, PM_Good, PM_Moderate, PM_Unhealthy, PM_Very Unhealthy, Irradiance	Temperature, Precipitation, Wind Speed, Wind Direction, Humidity, Amount of Sunshine, Irradiance, Clear, Partly cloudy, Mostly cloudy, Cloudy, PM_Good, PM_Moderate, PM_Unhealthy, PM_Very Unhealthy, Visibility
Output Parameter	PV (One hour ahead)		

4. Results and Discussion

To predict solar PV generation, experiments were conducted using five models in three groups, including ARIMAX, SVR_RBF, SVR_Linear, SVR_Poly, ANN, and DNN. The following error analysis methods were employed to evaluate the performance of each model:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i' - y_i| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i' - y_i)^2} \quad (7)$$

$$SMAPE(\%) = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - y_i'|}{|y_i| + |y_i'|} \quad (8)$$

$$MBE(\%) = \frac{\sum_{i=1}^n (y_i' - y_i)}{\sum_{i=1}^n y_i} \quad (9)$$

$$CV(\%) = \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n y_i} \quad (10)$$

where * y' : Predicted value, y : Observed value, n : Number of samples.

In general, the methods of analyzing errors in prediction models can be divided into two: the relative error analysis methods, such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE); absolute error analysis method, which uses percentages, such as Mean Absolute Error (MAPE). In most cases, these error analysis methods are used, but MAPE has the disadvantage of producing distorted results when the actual value is zero or when there are many extreme anomalies. For solar PV generation, MAPE cannot be used for a time other than the daylight time because the power generation converges to zero. To compensate for this problem, the Symmetric Mean Absolute Percentage Error (SMAPE) was employed. The SMAPE can compensate for the above shortcoming because it produces the same values even if the actual and the predicted values change.

In addition, the Mean Bias Error (MBE) and Coefficient of Variation (Cv) were employed according to the criteria of ASHRAE Guideline 14 [49]. These have absolute values like SMAPE, and more objective evaluation is possible because there is a clear standard of ASHRAE Guideline 14. According to

the standard, as shown in Table 9, the hourly prediction is specified within MBE of $\pm 10\%$ and Cv of 30%. The MBE has better performance as it goes closer to zero regardless of the sign. However, in this study, the absolute value was taken for the MBE value for intuitiveness and convenience.

Table 9. Acceptable range of ASHRAE Guideline 14.

Calibration Type	Index	Acceptable Value
Monthly	MBE month	$\pm 5\%$
	Cv (RMSE) month	15%
Hourly	MBE hour	$\pm 10\%$
	Cv (RMSE) hour	30%

Tables 10–12 summarize the error analysis results for each case. In each group, the model with the best performance for each error analysis method is shown in bold. From the analyses of the three cases, the prediction performance of Case 2, in which only PM parameters were used, was the lowest and that of Case 3, in which parameters for both the cloud and PM were used, was the best. The analysis of the prediction results of the three groups in each case showed that Group 3, which combined the other two groups to reflect the temporal and spatial characteristics, exhibited the best performance compared with Group 1 and Group 2. In Case 1 and Case 3, the models of Group 2 had better performance than the models of Group 1. Besides, we found that the cloud information extracted from the satellite images has more influence on solar PV generation prediction than the numerical weather information. On the other hand, in the case of the PM, numerical weather text data has a significant impact on solar PV generation prediction compare to satellite images. Among the SVR models for each kernel, the linear kernel model showed the best performance, and overall, the performance was improved in the order of ANN, ARIMAX, SVR_Linear, and DNN. The performance of the DNN model of Group 3 in Case 3 has the best performance in the entire experiments. The DNN and SVR_Linear models satisfied the criteria of the ASHRAE Guideline 14 in all cases and groups. However, all ANN models except for Group 2 in Case 2, Group 1 and 2 in Case 3, and ARIMAX model of Group 1 in Case 3 did not satisfy the criteria.

Table 10. Solar PV generation prediction result of case 1.

Group	Error	ARIMIX	SVR_RBF	SVR_Linear	SVR_Poly	ANN	DNN
Group 1 (Numerical Text Data)	MAE	81.261	90.686	81.112	102.638	122.285	72.554
	RMSE	101.768	111.289	101.63	128.766	149.354	98.519
	SMAPE	17.845	17.114	17.91	21.16	19.706	14.197
	MBE	0.017	0.394	0.526	0.504	23.958	0.956
	Cv	22.593	24.706	22.562	28.587	33.157	21.872
Group 2 (Satellite Images)	MAE	79.186	101.178	81.147	180.107	92.738	73.496
	RMSE	101.712	124.462	103.107	232.476	117.635	95.934
	SMAPE	15.635	18.178	16.51	28.447	16.786	14.484
	MBE	0.093	0.052	0.455	12.666	13.171	3.142
	Cv	22.58	27.631	22.89	51.61	26.115	21.298
Group 3 (Mixed, G1 + G2)	MAE	78.143	88.868	78.884	85.339	92.323	67.531
	RMSE	101.496	109.402	100.88	116.526	114.274	93.642
	SMAPE	17.792	16.841	17.613	16.58	17.857	15.039
	MBE	0.06	0.592	0.14	3.89	12.21	6.013
	Cv	22.532	24.288	22.396	25.869	25.369	20.789

Table 11. Solar PV generation prediction result of case 2.

Group	Error	ARIMIX	SVR_RBF	SVR_Linear	SVR_Poly	ANN	DNN
Group 1 (Numerical Text Data)	MAE	79.753	94.812	83.152	76.808	86.44	75.432
	RMSE	100.344	115.05	104.319	97.71	106.326	95.944
	SMAPE	17.377	17.417	18	16.241	19.124	14.439
	MBE	0.265	1.8	1.301	1.585	10.807	3.661
	Cv	22.277	25.541	23.159	21.692	23.605	21.3
Group 2 (Satellite Images)	MAE	78.33	96.5	78.103	190.973	82.34	75.727
	RMSE	102.158	118.679	99.672	255.323	105.741	103.094
	SMAPE	16.683	17.757	17.017	28.403	15.489	15.202
	MBE	0.106	1.665	0.009	18.586	3.51	0.303
	Cv	22.679	26.347	22.127	56.682	23.475	22.887
Group 3 (Mixed, G1 + G2)	MAE	78.895	90.663	80.282	77.882	93.974	71.295
	RMSE	102.432	110.424	102.413	104.554	115.703	96.355
	SMAPE	17.631	16.957	18.229	15.217	17.829	14.257
	MBE	0.195	0.493	0.482	1.207	12.481	2.156
	Cv	22.74	24.514	22.736	23.211	25.686	21.391

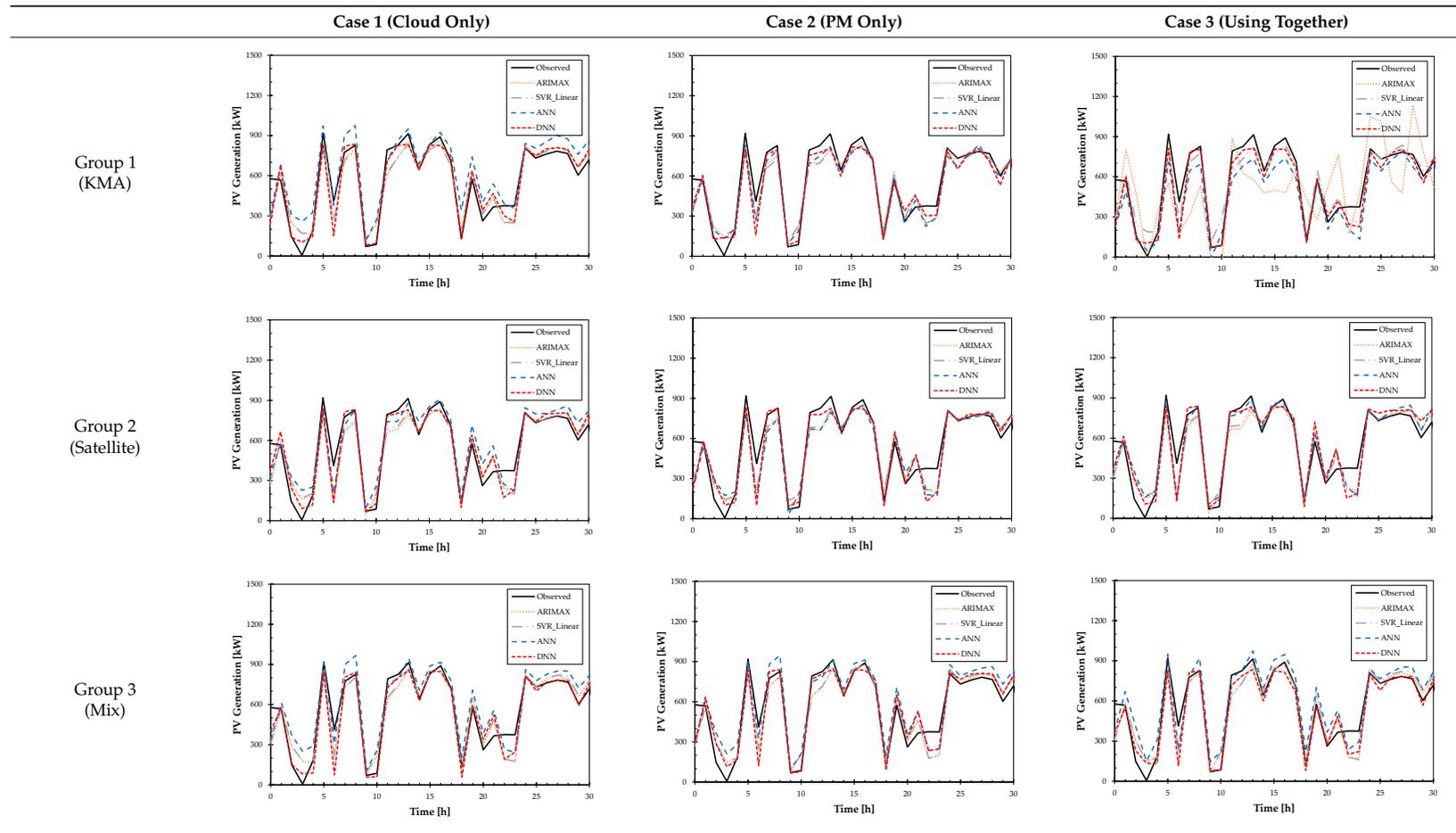
Table 12. Solar PV generation prediction result of case 3.

Group	Error	ARIMIX	SVR_RBF	SVR_Linear	SVR_Poly	ANN	DNN
Group 1 (Numerical Text Data)	MAE	196.129	90.096	81.154	79.128	95.97	79
	RMSE	235.048	111.22	102.268	100.929	119.607	104.059
	SMAPE	31.626	17.041	17.596	16.49	20.808	15.123
	MBE	11.816	2.076	0.171	1.219	6.619	8.376
	Cv	52.181	24.691	22.704	22.406	26.553	23.101
Group 2 (Satellite Images)	MAE	78.444	93.155	81.448	126.299	82.911	77.713
	RMSE	100.831	115.8223	102.696	169.343	107.843	102.005
	SMAPE	15.783	17.539	17.04	21.657	16.08	14.815
	MBE	0.146	0.883	0.744	4.915	8.321	5.343
	Cv	22.385	25.713	22.799	37.604	23.941	22.646
Group 3 (Mixed, G1 + G2)	MAE	77.554	87.587	80.123	83.163	101.233	71.532
	RMSE	101.943	108.229	102.889	107.278	122.508	92.938
	SMAPE	17.669	16.66	17.91	17.373	18.776	14.107
	MBE	0.342	0.498	0.039	0.134	16.572	4.986
	Cv	22.632	24.027	22.842	23.816	27.197	20.633

Table 13 lists the prediction results of the SVR_Linear model, which have the best performance among the SVR models, and those of the ARIMAX, ANN, and DNN models. These graphs are represented by extracting 30 h from the entire test periods. All the nine experimental group models yielded similar observed values, and the DNN models had the best performance.

To further improve the performance of the proposed models, it is necessary to identify and improve the factors that affect the prediction performance. As an impediment factor, the temporal continuity of the temporal characteristics may be interrupted during the removal of the missing values to satisfy the same conditions between each group. Also, there are cases where the clouds disappear naturally and the wind direction is not observed for a long time due to the weakness of the wind. Therefore, there is a need for more precise data interpolation methodologies to complement this. Moreover, the performance of the proposed model could be improved by collecting more data for analysis and optimizing the machine learning techniques used in this study.

Table 13. Solar PV generation prediction result of each group.



5. Conclusions

In this study, a hybrid spatio-temporal prediction model, which combines numerical weather text data and satellite image data, is proposed to develop an accurate solar PV generation prediction model, which is the most popular model in the field of renewable energies. Conventional meteorological data are composed of numerical text data and have a continuity of time, but it is difficult to reflect the effects of spatial characteristics, such as the movement of clouds and PM moving by the wind direction, as raw data in the prediction model. Therefore, numerical weather text data, satellite images, and time series-based solar PV generation data were employed to develop the prediction model. Data from a solar PV power plant located in Incheon, South Korea, was used as the test target, and the entire test was conducted on an hourly basis from January to December 2015. To develop the optimal prediction model, machine learning algorithms and statistical time-series analysis methods were employed, including ARIMAX, SVR_RBF, SVR_Linear, SVR_Poly, ANN, and DNN.

To reflect spatial characteristics in the proposed prediction model, ROI was designated in the satellite image, and then, ROI_{adj} was designated for eight adjacent directions around the ROI. After that, the data extracted from ROI_{adj} and ROI were used to predict the amount of cloud and PM parameters to be used in the prediction model for solar PV generation in the ROI. In addition to the amount of cloud and PM predicted in the ROI, other numerical weather text data and an hour previous solar PV generation data were combined to predict the solar PV generation. The experiment was conducted in three cases and three groups to determine the impact of clouds and PM on the prediction of solar PV generation. Among the different cases, the DNN model of Group 3 in Case 3, one of the hybrid models, yielded MAE of 71.532, RMSE of 92.938, SMAPE of 14.107%, MBE of 4.986%, Cv of 20.633%, indicating the best performance among all the models.

We propose a hybrid spatio-temporal DNN model that reflects the movement of clouds and PM using numerical weather text data and satellite images. From nine different experiments, we found that the spatial characteristics of clouds and PM affect the solar PV generation. Through the deepen data analysis, we confirmed that a proposed prediction model based on both satellite image data of clouds and numerical weather text data of PM improves the performance of solar PV generation prediction. In this study, a hybrid spatio-temporal prediction model supporting a notable performance was developed by utilizing satellite images and numerical weather data in predicting solar PV generation with time-series characteristics. This would be a useful guideline for the development of precise solar PV generation prediction models required to improve the efficient and stable power supply of renewable energy generation systems.

Author Contributions: Conceptualization & Methodology, B.K., D.S.; Writing—original draft, B.K., D.S.; Writing—review & editing, D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Korea Electric Power Corporation (grant number R19XO01-04) and results of a study on the “HPC Support” Project, supported by the ‘Ministry of Science and ICT’ and NIPA.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cullen, D. Climate change. *Nature* **2011**, *479*, 267–268. [[CrossRef](#)]
2. Horowitz, C.A. Paris Agreement. *Int. Leg. Mater.* **2016**, *55*, 740–755. [[CrossRef](#)]
3. Ministry of Trade, Industry and Energy. *Renewable Energy 3020 Plan. 3020 Plan*; Ministry of Trade, Industry and Energy: Sejong, Korea, 2017.
4. Korea Ministry of Trade, Industry and Energy. *Renewable Energy Statistics 2013. 2014*. Available online: <http://www.motie.go.kr> (accessed on 17 December 2017).
5. Choi, H.; Zhao, W.; Ciobotaru, M.; Agelidis, V.G. Large-scale PV system based on the multiphase isolated DC/DC converter. In *Proceedings of the 2012 3rd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, Aalborg, Denmark, 25–28 June 2012; pp. 801–807. [[CrossRef](#)]

6. Javier, M.; Luis, M.; Eduardo, L.; David, A.; Izco, E. Power output fluctuations in large scale PV plants: One year observations with one second resolution and a derived analytic model. *Prog. Photovolt.* **2011**, *19*, 218–227.
7. Won, J.-M.; Doe, G.-Y.; Heo, N.-R. Predict Solar Radiation According to Weather Report. *J. Korean Navig. Port Res.* **2011**, *35*, 387–392. [[CrossRef](#)]
8. Fang, X.; Misra, S.; Xue, G.; Yang, D. Smart Grid—The New and Improved Power Grid: A Survey. *IEEE Commun. Surv. Tutor.* **2011**, *14*, 944–980. [[CrossRef](#)]
9. Bae, K.Y.; Jang, H.S.; Sung, D.K. Hourly Solar Irradiance Prediction Based on Support Vector Machine and Its Error Analysis. *IEEE Trans. Power Syst.* **2016**, *32*, 935–945. [[CrossRef](#)]
10. Shi, J.; Lee, W.-J.; Liu, Y.; Yang, Y.; Wang, P. Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines. *IEEE Trans. Ind. Appl.* **2012**, *48*, 1064–1069. [[CrossRef](#)]
11. Yang, H.-T.; Huang, C.-M.; Huang, Y.-C.; Pai, Y.-S. A Weather-Based Hybrid Method for 1-Day Ahead Hourly Forecasting of PV Power Output. *IEEE Trans. Sustain. Energy* **2014**, *5*, 917–926. [[CrossRef](#)]
12. Lee, D.; Kim, K. Deep Learning Based Prediction Method of Long-term Photovoltaic Power Generation Using Meteorological and Seasonal Information. *Soc. e-Bus. Stud.* **2019**, *24*, 1–16.
13. Lee, K.; Kim, W.-J. Forecasting of 24_hours Ahead Photovoltaic Power Output Using Support Vector Regression. *J. Korean Inst. Inf. Technol.* **2016**, *14*, 175–183. [[CrossRef](#)]
14. Jang, H.S.; Bae, K.Y.; Park, H.-S.; Sung, D.K. Solar Power Prediction Based on Satellite Images and Support Vector Machine. *IEEE Trans. Sustain. Energy* **2016**, *7*, 1255–1263. [[CrossRef](#)]
15. Jang, H.S.; Bae, K.Y.; Park, H.-S.; Sung, D.K. Effect of aggregation for multi-site photovoltaic (PV) farms. In Proceedings of the 2015 IEEE International Conference on Smart Grid Communications (SmartGridComm), Miami, FL, USA, 2–5 November 2015; pp. 623–628. [[CrossRef](#)]
16. Hammer, A.; Heinemann, D.; Lorenz, E.; Lücke, B. Short-Term Forecasting of Solar Radiation. *1999 ISES Sol. World Congr.* **2000**, *67*, 411–418. [[CrossRef](#)]
17. Peng, Z.; Yoo, S.; Yu, D.; Huang, D. Solar irradiance forecast system based on geostationary satellite. In Proceedings of the 2013 IEEE International Conference on Smart Grid Commun. SmartGridComm, Vancouver, BC, Canada, 21–24 October 2013; pp. 708–713. [[CrossRef](#)]
18. Kim, I.-J.; Lee, S.-K. A Study on the Design of Testable CAM using MTA Code. *Trans. Korean Inst. Electr. Eng.* **2019**, 106–111. [[CrossRef](#)]
19. Zhang, X.; Li, Y.; Lu, S.; Hamann, H.F.; Hodge, B.-M.; Lehman, B. A Solar Time Based Analog Ensemble Method for Regional Solar Power Forecasting. *IEEE Trans. Sustain. Energy* **2018**, *10*, 268–279. [[CrossRef](#)]
20. Kang, H. An Analysis of the Causes of Fine Dust in Korea Considering Spatial Correlation. *Environ. Resour. Econ. Rev.* **2019**, *28*, 327–354. [[CrossRef](#)]
21. Peters, I.M.; Karthik, S.; Liu, H.; Buonassisi, T.; Nobre, A. Urban haze and photovoltaics. *Energy Environ. Sci.* **2018**, *11*, 3043–3054. [[CrossRef](#)]
22. Darwish, Z.A.; Kazem, H.A.; Sopian, K.; Al-Goul, M.; Alawadhi, H. Effect of dust pollutant type on photovoltaic performance. *Renew. Sustain. Energy Rev.* **2015**, *41*, 735–744. [[CrossRef](#)]
23. Maghami, M.R.; Hizam, H.; Gomes, C.; Hajjighorbani, S.; Rezaei, N. Evaluation of the 2013 Southeast Asian Haze on Solar Generation Performance. *PLoS ONE* **2015**, *10*, e0135118. [[CrossRef](#)]
24. Sarver, T.; Al-Qaraghuli, A.; Kazmerski, L.L. A comprehensive review of the impact of dust on the use of solar energy: History, investigations, results, literature, and mitigation approaches. *Renew. Sustain. Energy Rev.* **2013**, *22*, 698–733. [[CrossRef](#)]
25. Bergin, M.H.; Ghoroi, C.; Dixit, D.; Schauer, J.J.; Shindell, D. Large Reductions in Solar Energy Production Due to Dust and Particulate Air Pollution. *Environ. Sci. Technol. Lett.* **2017**, *4*, 339–344. [[CrossRef](#)]
26. El-Shobokshy, M.S.; Hussein, F.M. Degradation of photovoltaic cell performance due to dust deposition on to its surface. *Renew. Energy* **1993**, *3*, 585–590. [[CrossRef](#)]
27. Korea Meteorological Administration. Available online: <https://data.kma.go.kr/> (accessed on 22 October 2020).
28. Air Korea. Available online: <https://www.airkorea.or.kr/> (accessed on 22 October 2020).
29. Open Data Portal. Available online: <https://www.data.go.kr/> (accessed on 22 October 2020).
30. National Meteorological Satellite Center. Available online: <https://nmssc.kma.go.kr/> (accessed on 22 October 2020).
31. National Meteorological Satellite Center. *AMV (AMV: Atmospheric Motion Vector) Algorithm Theoretical Basis Document*; National Meteorological Satellite Center: Jincheon, Korea, 2012.

32. National Meteorological Satellite Center. *COT Algorithm Theoretical Basis Document*; National Meteorological Satellite Center: Jincheon, Korea, 2012.
33. National Meteorological Satellite Center. *AOD Algorithm Theoretical Basis Document*; National Meteorological Satellite Center: Jincheon, Korea, 2012.
34. National Meteorological Satellite Center. *INS Algorithm Theoretical Basis Document*; National Meteorological Satellite Center: Jincheon, Korea, 2012.
35. Wolff, B.; Kühnert, J.; Lorenz, E.; Kramer, O.; Heinemann, D. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. *Sol. Energy* **2016**, *135*, 197–208. [[CrossRef](#)]
36. Li, L.; Gong, J.; Zhou, J. Spatial Interpolation of Fine Particulate Matter Concentrations Using the Shortest Wind-Field Path Distance. *PLoS ONE* **2014**, *9*, e96111. [[CrossRef](#)] [[PubMed](#)]
37. Newsham, G.R.; Birt, B.J. Building-level occupancy data to improve ARIMA-based electricity use forecasts. In Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building, Zurich, Switzerland, 3–5 November 2010; pp. 13–18. [[CrossRef](#)]
38. Dong-hyun, L.; A-hyun, J.; Jin-young, K.; Chang-Ki, K.; Hyun-goo, K.; Yung-seop, L. Solar Power Generation Forecast Model Using Seasonal ARIMA. *Korean Sol. Energy Soc.* **2019**, *39*, 59–66.
39. Paulescu, M.; Badescu, V.; Brabec, M. Tools for PV (photovoltaic) plant operators: Nowcasting of passing clouds. *Energy* **2013**, *54*, 104–112. [[CrossRef](#)]
40. Alsharif, M.H.; Younes, M.K.; Kim, J. Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea. *Symmetry* **2019**, *11*, 240. [[CrossRef](#)]
41. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
42. Liu, W.; Liu, C.; Lin, Y.; Ma, L.; Xiong, F.; Li, J. Ultra-Short-Term Forecast of Photovoltaic Output Power under Fog and Haze Weather. *Energies* **2018**, *11*, 528. [[CrossRef](#)]
43. Kim, K.; Jin, H. Photovoltaic Power Forecasting and Analysis of Forecasting Error for Model Learning Periods Using SVR. In Proceedings of the Korean Institute of Electrical Engineers, Pyeongchang, Korea, 11–13 July 2018; pp. 324–325.
44. McCulloch, W.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biol.* **1990**, *52*, 99–115. [[CrossRef](#)]
45. Saberian, A.; Hizam, H.; Radzi, M.A.M.; Ab Kadir, M.Z.A.; Mirzaei, M. Modelling and Prediction of Photovoltaic Power Output Using Artificial Neural Networks. *Int. J. Photoenergy* **2014**, *2014*, 1–10. [[CrossRef](#)]
46. Teoh, E.J.; Tan, K.C.; Xiang, C. Estimating the Number of Hidden Neurons in a Feedforward Network Using the Singular Value Decomposition. *IEEE Trans. Neural Netw.* **2006**, *17*, 1623–1629. [[CrossRef](#)]
47. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [[CrossRef](#)]
48. Oh, I.-S. *Machine Learning*; HANBIT Academy, Inc.: Seoul, Korea, 2017.
49. ANSI/ASHRAE. ASHRAE Guideline 14-2002 Measurement of Energy and Demand Savings. *ASHRAE* **2002**, *8400*, 170.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).