

## Article

# Spatiotemporal Optimization for Short-Term Solar Forecasting Based on Satellite Imagery

Myeongchan Oh , Chang Ki Kim, Boyoung Kim , Changyeol Yun, Yong-Heack Kang and Hyun-Goo Kim \* 

New & Renewable Energy Resource Map Laboratory, Korea Institute of Energy Research, Daejeon 34129, Korea; amir117@kier.re.kr (M.O.); ckkim@kier.re.kr (C.K.K.); bykim@kier.re.kr (B.K.); yuncy@kier.re.kr (C.Y.); yhkang@kier.re.kr (Y.-H.K.)

\* Correspondence: hyungoo@kier.re.kr

**Abstract:** Solar forecasting is essential for optimizing the integration of solar photovoltaic energy into a power grid. This study presents solar forecasting models based on satellite imagery. The cloud motion vector (CMV) model is the most popular satellite-image-based solar forecasting model. However, it assumes constant cloud states, and its accuracy is, thus, influenced by changes in local weather characteristics. To overcome this limitation, satellite images are used to provide spatial data for a new spatiotemporal optimized model for solar forecasting. Four satellite-image-based solar forecasting models (a persistence model, CMV, and two proposed models that use clear-sky index change) are evaluated. The error distributions of the models and their spatial characteristics over the test area are analyzed. All models exhibited different performances according to the forecast horizon and location. Spatiotemporal optimization of the best model is then conducted using best-model maps, and our results show that the skill score of the optimized model is 21% better than the previous CMV model. It is, thus, considered to be appropriate for use in short-term forecasting over large areas. The results of this study are expected to promote the use of spatial data in solar forecasting models, which could improve their accuracy and provide various insights for the planning and operation of photovoltaic plants.

**Keywords:** solar forecasting; spatial analysis; satellite images; cloud motion vector (CMV); spatiotemporal; optimization



**Citation:** Oh, M.; Kim, C.K.; Kim, B.; Yun, C.; Kang, Y.-H.; Kim, H.-G. Spatiotemporal Optimization for Short-Term Solar Forecasting Based on Satellite Imagery. *Energies* **2021**, *14*, 2216. <https://doi.org/10.3390/en14082216>

Academic Editor: Jesús Polo

Received: 24 March 2021

Accepted: 13 April 2021

Published: 15 April 2021

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



**Copyright:** © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The growing importance of solar forecasting has been emphasized in many published studies and in review articles comparing different forecasting techniques. However, review papers [1–3] have categorized previous studies according to the method, forecast horizon, and other characteristics. In addition, text mining of literature [4] and those focused on machine learning methods [5]. However, it can be identified that these reviews did not propose one optimal method [6]. It is difficult to evaluate the superiority of models as their performances depend on evaluating their metrics, forecast horizons, and the region studied. Although several objective comparison methods have been proposed [6,7], their performances depend on spatiotemporal features, which need to be thoroughly considered.

Although ensemble or machine learning models have been shown to perform well, models suited to specific forecasting conditions are required. As it is difficult to develop one model that provides optimal results for every geographical region and forecast horizon. Thus, hybrid or ensemble models containing several different models have been introduced, and these have been shown to exhibit better performances than individual models [8]. Many high-performance specific models have been selected and combined in this respect [9]. Machine learning techniques have yielded remarkable performances when using such models as input. However, to improve the results of these models, it is essential to gain physical interpretation and domain knowledge of the specific models.

The aim of this study, therefore, was to develop a forecasting model for specific conditions that can be used as a core component in ensemble methods. Satellite-image-based forecasting was selected here, as this is appropriate for making short-term (i.e., intra-hour or intra-day) predictions over a large area. The relative characteristics of the satellite techniques have been previously discussed [2,5]. The most important advantage of satellite images is that they provide a large amount of information over a large area. Therefore, many studies have, thus, incorporated spatial information from satellite images into forecasting models. A text mining review [4] found that four model terms were frequently referred to: Cloud motion forecast (CMF), cloud motion vector (CMV), cross-correlation method (CCM), and particle image velocimetry (PIV). However, all these terms are methods for CMVs, and most previous studies have focused on CMVs, whereas other methods have rarely been studied.

Cloud tracking using satellite images has been used in various fields [10], and all these tracking models are defined as CMV in this study. The CMV model using block matching has been widely applied in Germany [11–15], and a root mean squared error (RMSE) of 20–40% was yielded over periods of 1–5 h using data collected from a single station [13]. The same model was applied to regions in the USA [8,16], France [17], and Uruguay [18]. Other CMV models using a type of optical flow (OF) have been applied in several regions [19–22], and remarkable performances have been obtained. Many of these studies have shown that CMV models perform better than numerical weather prediction (NWP) models for short forecast horizons (less than a few hours) [12,14,15]. A detailed description and comparison of CMV models are presented in the methodology chapter.

Certain CMV models are applied to solar forecasting with small-scale data using sky camera [23–26] or simulated data [27]. However, only selected studies that employed CMV for solar forecasting using satellite images are considered here. Small-scale studies and satellite-scale studies should be treated differently, as the corresponding cloud characteristics and research scope are different. Previous studies have shown that small-scale CMV forecast is suitable for a few-minute [23,24] forecasts. Additionally, they require more measurement sensors and are difficult to apply to a large area. In contrast, solar forecasting with satellite images can make forecasts for several hours on a national scale.

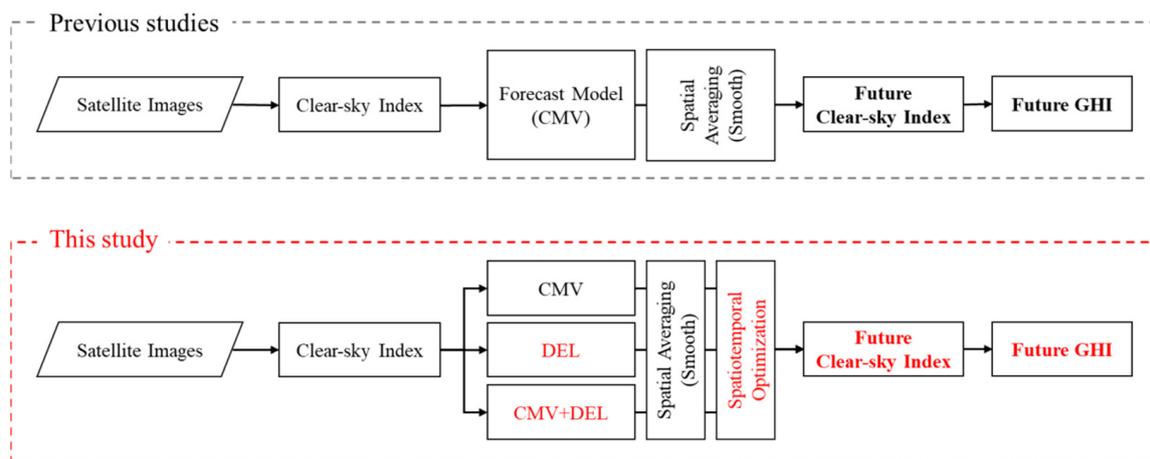
CMV models consider the movement of clouds, and they extract cloud motion information from satellite images and apply it to past images for forecasting. However, CMV models can be considered to be CMV persistence models [28] because the first estimated CMV is applied to all forecast horizons. Although CMV models yield a high performance, they assume that the shape, movement, and intensity of the cloud are constant, but this does not reflect the highly dynamic behavior of clouds. These assumptions are also applied to the extraction methods of CMV, and this limits their accuracy and forecasting steps. Such limitations, thus, significantly lower the performance of the model in areas with complex cloud dynamics. In this respect, the CMV model may deliver a poorer performance than a persistence model (i.e., a simple reference model) in such areas. It is, thus, necessary to conduct a spatial analysis of the models' performances, and the influence of spatial characteristics should be studied.

Studies with additional considerations have been conducted to improve the performance of CMVs. Multiple CMVs according to height were considered by performing cloud classification according to height [29] or by generating three-dimensional cloud detection [30]. Classification according to cloud patterns was also considered for CMV calculation [31]. Some studies on precipitation forecast used a motion of precipitation distribution similar to CMV [32–34]. One of these studies compared several motion extraction methods, including motion extraction points [34]. Nevertheless, there was no solution to the fundamental limitation of CMV from its assumption.

To date, there have been insufficient evaluations of the spatial accuracy and distribution and their correlations with other factors in forecasting studies. Although CMV models use spatial information and improvements have been made, only ambient spatial information is included, and other factors or distributions are not considered. A remark-

able performance was observed for a CMV model using spatial averaging as a smoothing method [11,14], but the smoothing was performed using a simple filter application. Spatiotemporal correlation [27] analyses have been conducted, but analyses were only conducted to enhance CMV extraction in a small-scale area. Some simple comparisons between different regions have been made, such as between different climate zones within the same area [8]; however, most studies have used site-specific or overall-averaging results. As the available solar energy incident on Earth depends on various weather conditions, it must be correlated with the terrain (land or sea, and coastal or mountain regions). Therefore, conducting a spatial analysis using distribution patterns and topography could contribute to error analyses and spatial optimization. In this respect, satellite images provide uninterrupted information about an entire region, compared to specific site-measurement data, and are, hence, suitable for spatial analysis studies.

The main aim of this study was to analyze spatial characteristics relevant to the development of solar-forecasting models with the objective of improving forecasting performance. The following steps were applied: (1) The characteristics of satellite-image-based forecasting models containing CMV were compared; (2) a spatial analysis of the performance of these models was conducted; and (3) spatiotemporal optimization was conducted to maximize the forecasting performance. A new spatiotemporal forecasting model was then developed as a hybrid of CMV and additional models considering their spatial accuracy. The structure of this model is shown in Figure 1, where it is also compared to the structure of most previously studied models. The new models in red, such as delta persistence (DEL), are described in detail in the methodology chapter.



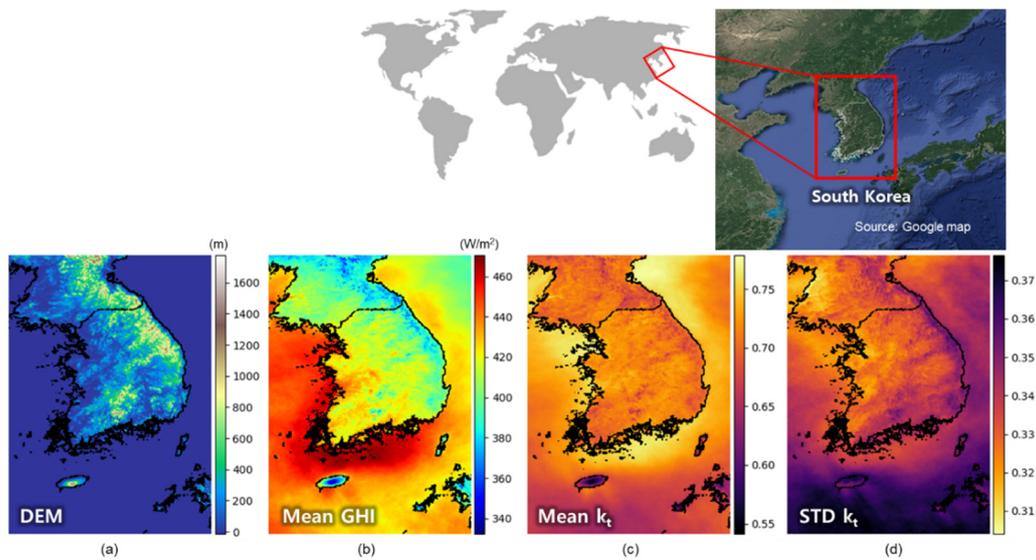
**Figure 1.** Flowchart of satellite-image-based solar forecasting. Top: Forecasting in previous models, bottom: Model in this study. Newly proposed or altered terms are shown in red, and associated details are provided in the methodology chapter.

## 2. Data

### 2.1. Study Area

The study area was the South Korean peninsula, which has complex topography. The location of the study area and associated basic spatial information are shown in Figure 2, and the area corresponds to a region spanning  $636 \times 459$  km.

The digital elevation model (DEM), also shown in Figure 2, was resampled from Shuttle Radar Topography Mission (SRTM) data. The study area experiences complex meteorological phenomena as it is surrounded by sea and mountainous regions (a long mountain range in the east and Jeju Island, which is a large island in the south, are visible on the map in Figure 2). The area lies within a temperate climate zone, experiences four seasons, and is affected by the East Asian monsoon. The mean global horizontal irradiance (GHI) of South Korea is approximately  $3.5 \text{ kWh/m}^2/\text{day}$ .



**Figure 2.** Location of the study area (South Korea) and basic associated spatial information. (a) digital elevation model (DEM), (b) mean global horizontal irradiance (GHI), (c) mean clear-sky index ( $k_t$ ), and (d) standard deviation of  $k_t$  are visualized.

The spatial distribution of GHI and the clear-sky index ( $k_t$ ) are also shown; these were estimated from satellite images, and the process used is described in detail later. Only data from 2018 were used, and it is acknowledged that annual volatility may result in different distributions. Nevertheless, differences relating to the land and sea areas are obvious, regardless of such inter-year volatility. Although all variables (mean GHI, mean  $k_t$ , and standard deviation (STD) of  $k_t$ ) are related to the solar potential, they show slightly different distributions.

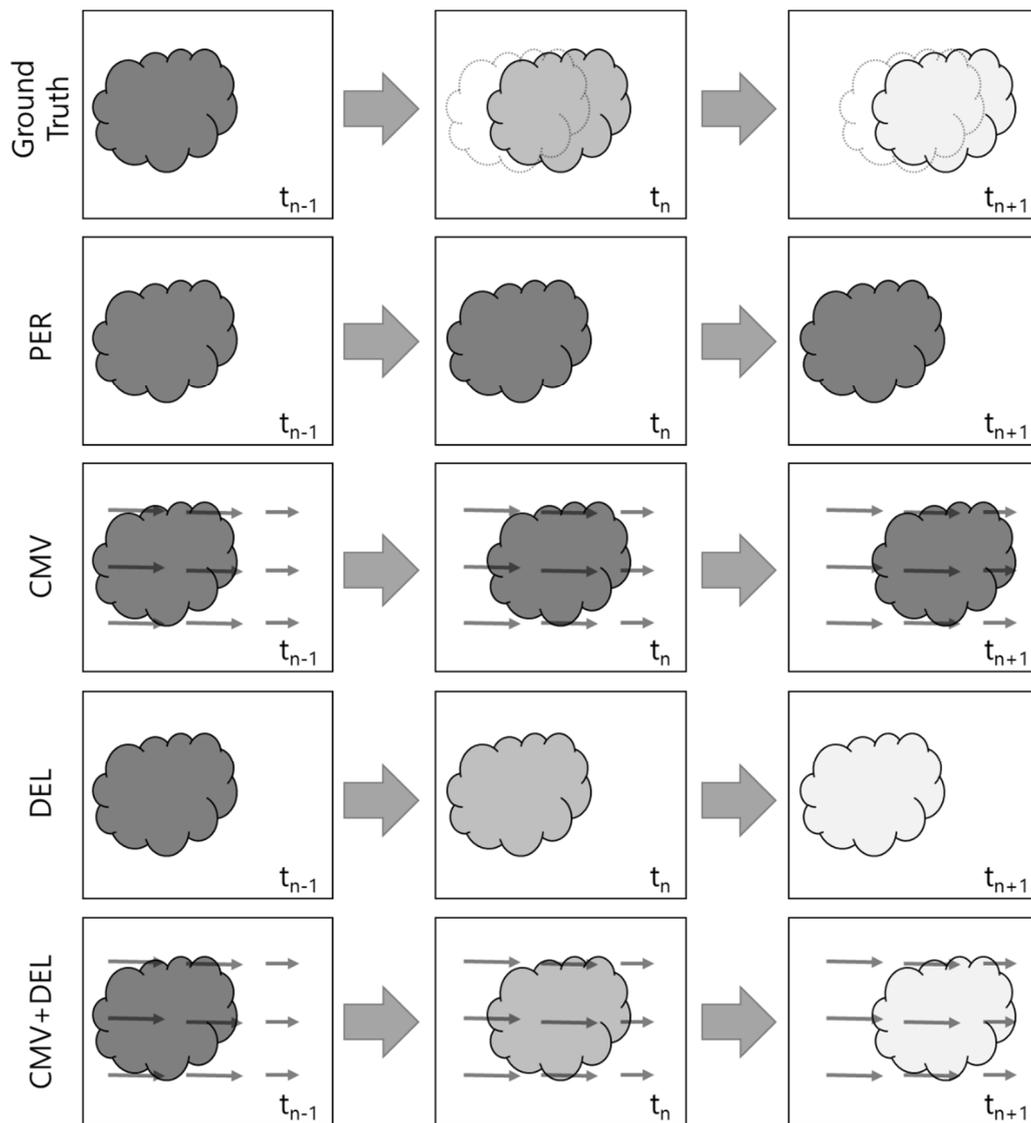
## 2.2. Satellite Images

The Communication, Ocean and Meteorological Satellite (COMS), a geostationary satellite operated by the Korea Meteorological Administration, was used to obtain the satellite images for this study. The satellite images had a spatial resolution of 1 km and a temporal resolution of 15 min. Only processed daytime images (08:30 to 17:00) for the entire period of 2018 were used, and no raw images were employed. The  $k_t$  and clear-sky GHI images were extracted from the COMS images using the University of Arizona Solar Irradiance Based on Satellite—Korea Institute of Energy Research (UASIBS/KIER) model [35], and subsequently used for solar forecasting, and the detailed algorithm and accuracy results can be found in the associated reference. Figure 1 shows that the  $k_t$  images were used in forecasting, and clear-sky GHI images were used to calculate GHI from  $k_t$ .

## 3. Methodology

### 3.1. Satellite-Image-Based Solar Forecasting Models

Four models were used for satellite-image-based solar forecasting. The persistence (PER) and CMV models have been reported in previous studies and were used as the reference and standard models, respectively. The delta persistence (DEL) model and a hybrid model of the CMV and DEL (CMV+DEL) were initially proposed in this study to overcome the limitations of the existing CMV models. Simple diagrams of all the models are shown in Figure 3. The models were analyzed using  $100 \times 100$  km sample images over a period of 5 h from processed satellite images, and the forecasting models were implemented using Python code and open-source libraries, such as OpenPIV, OpenCV, and TensorFlow.



**Figure 3.** Diagram of satellite-image-based solar forecasting models used in the study. PER, CMV, and DEL represent the persistence model, cloud motion vector model, and delta persistence model, respectively. The lightness of the cloud represents the clear-sky index of the cloud.

### 3.1.1. Persistence Model

The PER model is commonly used as a reference model [3]. It has several variations, but we used the “clear-sky persistence” model [6], which uses the last-measured  $k_t$  for all periods. The PER model is described by Equation (1),

$$k_{t\_PER}(x, y, t_{n+1}) = k_t(x, y, t_n), \quad (1)$$

where  $x, y$  denotes the location of  $k_t$ , and  $t_n$  denotes the  $n$ th time point of  $k_t$ . The calculated  $k_t$  images at  $t_0$  are applied to images at  $t_1$  to  $t_n$  with different clear-sky GHI values. This model assumes that cloud covers remains unchanged, and its expected accuracy is, therefore, low. Compared to previous site-specific studies, the spatial analysis of the PER model can show the reference accuracy in each location. In addition, the distribution pattern and distinct locations can be identified.

### 3.1.2. Cloud Motion Vector Models

CMVs are based on the extraction of movement information from sequential images, and they have been used in many applications. Block matching is the most popular CMV model used for solar forecasting [11,14,15], and it identifies the vectors by estimating the displacement of similar areas in the before and after images. It is commonly used in particle image velocimetry (PIV) [36] for determining fluid dynamics. PIV uses a cross-correlation for evaluation, whereas the CMV model applied in Germany [14] uses the mean squared error (MSE). The MSE is simple and easy to implement, but it requires extremely high computational power; therefore, a low spatial resolution of 43 km was used in the previous study. In contrast, the cross-correlation method (CCM), which calculates cross-correlation using Fourier transforms, is more appropriate for high-resolution images. However, it has limitations in similar pixel groups [37] and shows a lower performance than the OF method [20].

The OF method extracts movement using an image processing technique and is being actively researched for computer vision. Although block matching can also be categorized as OF, we use these two terms separately here. OF, thus, refers to differential methods, such as the Horn–Schunck method [38] or deep-learning-based models, which show smoother CMVs. Many associated models have been developed recently. Some studies have used OF methods, such as the Lucas–Kanade (LK) method [19,20] and Classic++ [21] for solar forecasting. However, as these models were optimized using different images and have many hyperparameters, they need to be objectively compared. Therefore, we compared several CMV extraction methods using processed satellite images: CCM, Farneback [39], Dual TVL1 [40], Deep Flow [41], Brox [42], and PWC-Net [43]. All of these methods are OF methods, except for CCM and PWC-Net, which are deep learning models. The optimized parameters for the Dual TVL1 were taken from similar previous studies [44].

The extracted CMV is applied to  $k_t$  images using the warp method for image processing. When the time of the first image is  $t_0$ , the CMV is estimated using images at  $t_0$  and  $t_1$ . First, forecasting is conducted using the image at  $t_1$ , followed by forecasting using the forecasted image. Data from multiple locations can be considered for actual CMV processing. However, at a particular area at the boundary of the image, no data are recorded, due to no coming CMV data, and the PER model has been applied to such areas. Equation (2) describes the CMV model,

$$k_{t\_CMV}(x, y, t_{n+1}) = \begin{cases} k_t(x - u, y - v, t_n) & \text{if } n = 0 \\ k_{t\_CMV}(x - u, y - v, t_n) & \text{if } n > 0 \end{cases} \quad (2)$$

where  $u$  and  $v$  denote the two vectors of CMV.

### 3.1.3. Delta Persistence (DEL) Model

A new type of persistence model, named delta persistence, was developed, which assumes that the change in  $k_t$  is constant (see Equation (3)). The concept is similar to ramp persistence [28], but ramp persistence uses a summation term, whereas the DEL model uses a product term. The product term for the change in  $k_t$  is referred to as the delta ( $\Delta$ ) term in this study. As the rate of change is constant, the magnitude of the change continuously decreases over time. This model can consider the change in cloud thickness, but it cannot consider cloud movements. As  $k_t$  is used as a denominator, the rate of change can be extremely large when  $k_t$  is close to 0. Thus, the rate of change is set to be less than 1 to form an inverse index with increasing  $k_t$ ,

$$k_{t\_DEL}(x, y, t_{n+1}) = \begin{cases} k_t(x, y, t_n) \times \Delta_{dec} & \text{if } k_t(x, y, t_1) < k_t(x, y, t_0) \\ 1 - \{1 - k_t(x, y, t_n)\} \times \Delta_{inc} & \text{if } k_t(x, y, t_1) > k_t(x, y, t_0) \end{cases} \quad (3)$$

$$\left[ \Delta_{dec} = \frac{k_t(x, y, t_1)}{k_t(x, y, t_0)}, \Delta_{inc} = \frac{1 - k_t(x, y, t_1)}{1 - k_t(x, y, t_0)} \right]$$

### 3.1.4. CMV + DEL Model

The hybrid model of CMV and DEL model uses  $\Delta$  terms estimated from images from the CMV model. As shown in Equation (4), the difference between this model and the DEL is the association of  $k_{t\_CMV}$  with some terms,

$$k_{t\_DEL}(x, y, t_{n+1}) = \begin{cases} k_{t\_CMV}(x-u, y-v, t_n) \times \Delta_{dec\_CMV} & \text{if } k_{t\_CMV}(x, y, t_1) < k_t(x, y, t_0) \\ 1 - \{1 - k_{t\_CMV}(x-u, y-v, t_n)\} \times \Delta_{inc\_CMV} & \text{if } k_{t\_CMV}(x, y, t_1) > k_t(x, y, t_0) \end{cases} \quad (4)$$

$$\left[ \Delta_{dec\_CMV} = \frac{k_{t\_CMV}(x, y, t_1)}{k_t(x, y, t_0)}, \Delta_{inc\_CMV} = \frac{1 - k_{t\_CMV}(x, y, t_1)}{1 - k_t(x, y, t_0)} \right]$$

This model uses the CMV to accurately forecast future  $k_t$  images, and the  $\Delta$  term is then used to consider the formation and disappearance of clouds. The  $\Delta$  terms are estimated using actual images at  $t_1$  and forecasted images using the CMV. Theoretically, the model should provide a higher performance than the CMV and DEL separately, but its overall accuracy was not significantly improved because of errors in the extracted CMV and  $\Delta$  terms. A detailed error analysis was, thus, conducted to verify the accumulation of errors.

### 3.2. Spatial Averaging

Spatial averaging was applied as a smoothing method. This is not strictly a part of the forecasting model but is rather a post-processing technique for increasing the performance. Previous studies have shown the importance of smoothing in PER and CMV models [14]. During smoothing, the average value of nearby locations was used to reduce the spatial uncertainty. Rectangular filters with different sizes according to the forecast horizon [14] were used in this study. The size of the filter was selected based on the resulting accuracy, and was specifically 10, 20, 30, and 50 km for 0.25 h, 0.5 h, 0.75 h, and >1 h, respectively. Equation (5) is the simple equation used for smoothing,

$$k_{t\_FORs}(x, y, t_{n+1}) = \frac{\sum_i \sum_j k_{t\_FOR}(x+i, y+j, t_{n+1})}{N}, \quad (5)$$

where  $FOR$  denotes all forecasting models,  $s$  denotes smoothing, and  $N$  denotes the number of nearby pixels during smoothing.

Smoothing was also applied to the  $\Delta$  term in the DEL and hybrid models, as this term also has spatial uncertainty. The equation for smoothing is similar to Equation (5), but it uses the  $\Delta$  term instead of the  $k_{t\_FOR}$  term. The size of the filter was 150 km for both models.

### 3.3. Spatial Accuracy Analysis

Three accuracy metrics were used in the analysis: The root mean squared error (RMSE), mean bias error (MBE), and the skill score (SS). The relative value, which divides the metric into a mean value was used for normalization. The SS was used to effectively visualize and compare the performances of the models. The simplified equation in Equation (6) [3] was used to calculate SS,

$$SS \approx 1 - \frac{RMSE}{RMSE_p}, \quad (6)$$

where  $RMSE$  is the RMSE of the model being analyzed, and  $RMSE_p$  is the RMSE of the persistence model. Although a combination of climatology and persistence models is recommended for SS calculations [6], we used the persistence model as a reference, as it is considered to be more accurate than a climatology model for short-term forecasting.

The spatial accuracy of the models was visualized and analyzed by temporally averaging the metric in each location. However, the influence of weather conditions was greater than the difference between each model, therefore, the SS effectively showed the difference between the accuracy of the models and the reference model. A similar large distribution in the forecast accuracy was identified in the actual results. Accordingly, the RMSE and

MBE maps of the PER model were analyzed, and then the SS and MBE maps of the other models were analyzed. To quantify the similarity of the distributions or interconnections, the correlations between several variables and metrics were considered in this analysis.

### 3.4. Spatiotemporal Optimization

Previous studies have considered temporal patterns according to the models and proposed ensemble models. The terms “intra-hour”, “intra-day”, and “a day ahead” have been introduced to represent these temporal patterns. However, as the majority of satellite-image-based forecasting models used in previous studies were CMV, temporal optimization has not been previously applied. It was applied here to compare the four models in the current study, and the additional consideration of spatial characteristics was expected to improve the model performance. Both temporal and spatial accuracy patterns were analyzed and used in the optimization process, and the temporal patterns were also analyzed by averaging the spatial values.

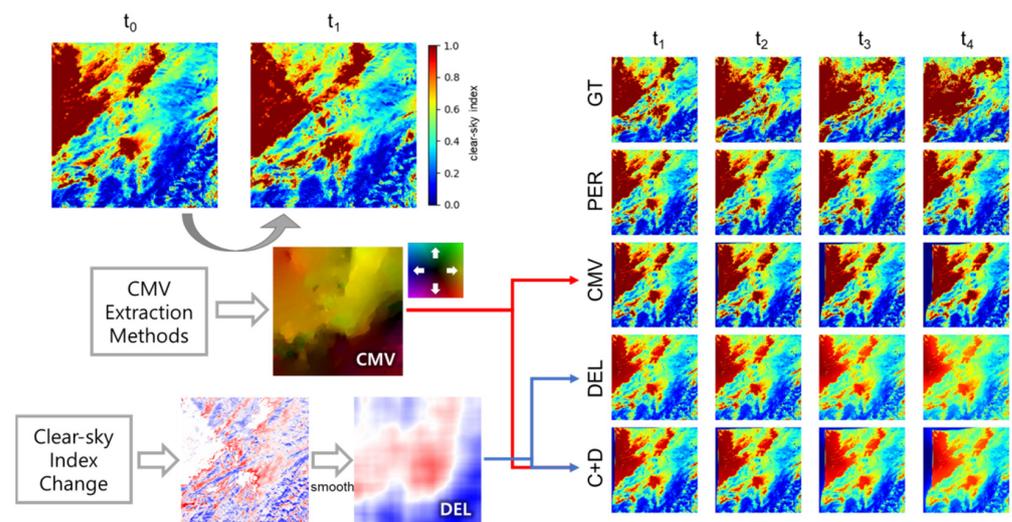
A new spatiotemporal optimization method was developed by mapping the best models according to the location. The best-model map was calculated by comparing the RMSE map for each model. The temporal best-model maps were estimated in each forecast horizon and used for the actual optimized (OPT) model. When calculating the best-model maps, the training and test data were separately used for validation, and two-fold cross-validation was conducted to estimate the results in all periods. As the data were a time-series, they were separated daily. A physical interpretation of the best-model map was made considering the models’ characteristics and other spatial information. Only the land area of South Korea was used for model comparison considering the practical usage.

## 4. Results and Discussion

### 4.1. Performance of Forecasting Models

The six CMV extraction models in the methodology chapter were compared. The differences between the models for extracting CMV values were found to be insignificant, probably because they all assumed that the shape or brightness of an object was constant. The maximum RMSE difference between the OF models was up to 3% in  $k_t$  image forecasting, and the RMSE was 5% lower in the CCM model compared to the best OF models. This was evidenced by the fact that the difference between the CMV persistence forecasting and CMV updating for each image was marginal, and only a 2% difference in the RMSE was observed for  $k_t$  image forecasting. Therefore, one of the most popular OF models, Dual TVL1, was selected as the CMV in this study, as it performed slightly better than the other models.

Figure 4 shows an example of the results of the forecasting models using sample images. This example shows the evolution of complex weather phenomena in the study area with 15-min intervals. The low- $k_t$  regions, shown in blue, represent the clouds. The extracted CMV and  $\Delta$  terms are also identified in this figure. In general, the extracted CMV or DEL was applied to the image at  $t_1$ ; however, in this figure, it was applied to the image at  $t_0$  to analyze the performance of the model. The additional processing for the area with no data in the CMV model was not complemented in the images. Smoothing was applied to the  $\Delta$  terms for the DEL model, but it was not applied to the forecasted images in this figure.

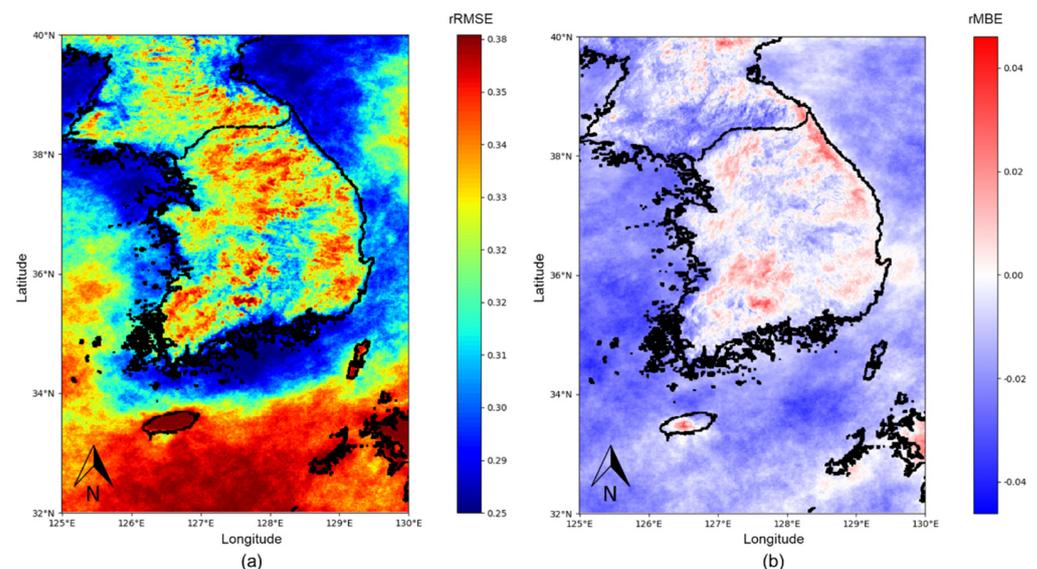


**Figure 4.** Visualization of the forecasting models' processes using  $100 \times 100$  km sample images. GT represents the ground truth, and other terms represent the forecasting models.

The forecast images clearly show the limitations and characteristics of the forecasting models for these complex weather conditions. The CMV model considers the movement of clouds, but the forecast shape is different from the ground truth (GT) because of the errors propagated by the CMV model and the rapid changes in cloud conditions. For the DEL model, the shape of the clouds differs from the GT, but the use of the  $\Delta$  term enables the overall trend of clearing clouds to be visualized. The characteristics of both the CMV and DEL models were observed for the hybrid model. Although the hybrid model combines the advantages of both models, it also contains accumulated errors from both models, and this could make it difficult to accurately predict clouds in some cases. In the dataset used in this study, the clouds rarely moved or changed while maintaining their shape.

#### 4.2. Spatial Analysis

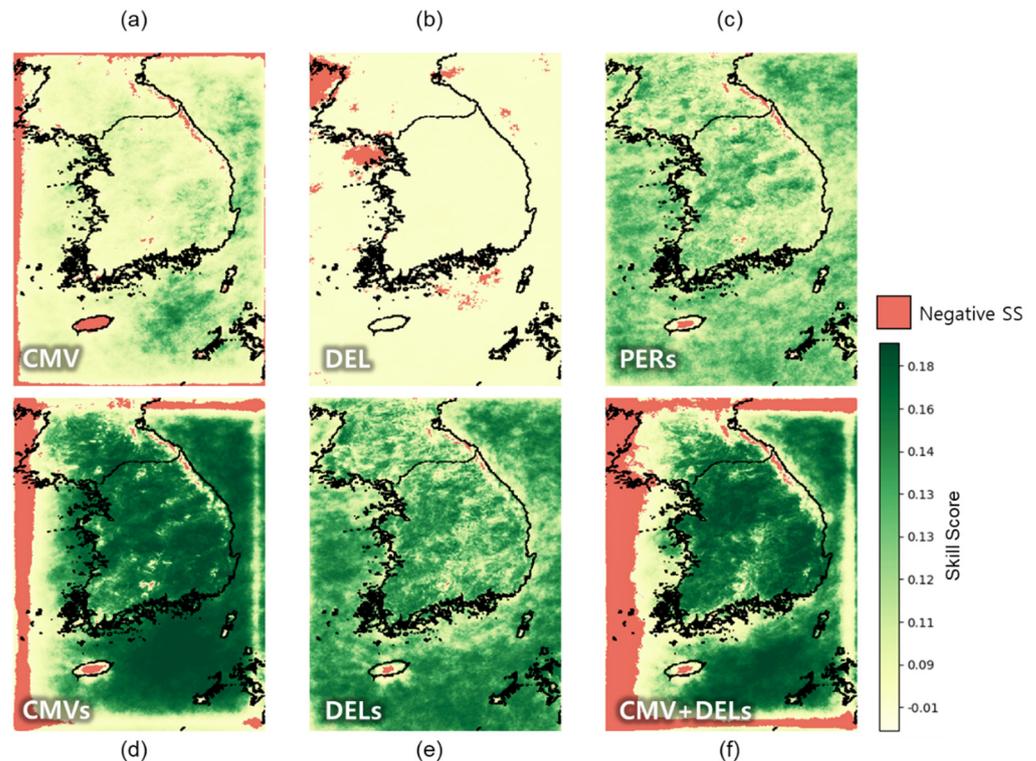
The spatial distributions of the RMSE and MBE of the PER model are shown in Figure 5. The metrics were calculated over 4 h using 16 images. There were no significant changes in the spatial distributions of the error as a function of the forecast horizon for the PER model.



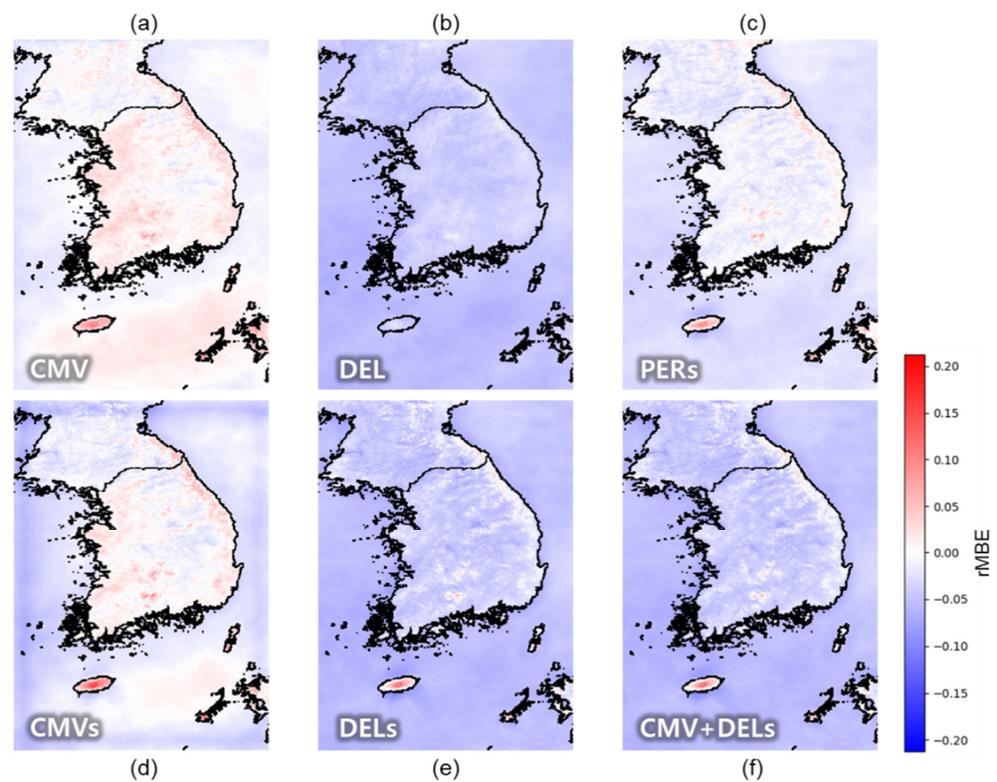
**Figure 5.** Spatial distribution of (a) RMSE and (b) MBE of the PER model.

The RMSE map is seen to follow the geographical map well, in which the land and sea regions were clearly different. The maximum difference in the RMSE was approximately 26%, which means that the RMSE of the area with the lowest accuracy is more than twice that of the area with the highest accuracy area. The color bar in the figure was normalized according to the quantile; thus, the actual minimum and maximum value may differ from the color bar. The mountain on Jeju Island shows the highest RMSE (>50%), but the area in the peninsula excluding Jeju has RMSE values of up to 40%. In contrast, the lowest RMSE of 25% is found in the sea near the peninsula. In the case of the MBE, values up to 5% were observed, and there was a tendency to overestimate the values in the inland highlands and underestimate the values in the remaining areas. The MBE map suggests that there was considerable cloud formation in the high-altitude areas, and less cloud cover in the low-altitude areas.

The SS and MBE values of the six forecasting models are shown in Figures 6 and 7, respectively; the differences between each model are clearly identifiable in these figures. Smoothing increased the overall accuracy, as shown by the SS map of the smoothed model. However, this improvement was small, or even negative, for the eastern mountain range and Jeju Island, and this trend was confirmed for all the other models. The simple PER model was found to be better than the other models for forecasting complex weather at high altitude and at the coast. The performance of the CMVs model was improved by smoothing in many regions, but negligible or negative improvements were observed at the edge of the image. This was expected because there were no incoming clouds at the boundary, but the western region was more affected by this due to the influence of the westerly wind in the study area. The DELs and hybrid model generally underestimated the MBE values. This performance was improved by modifying the delta terms of the DELs model, although this effect was not significant for the RMSE values.

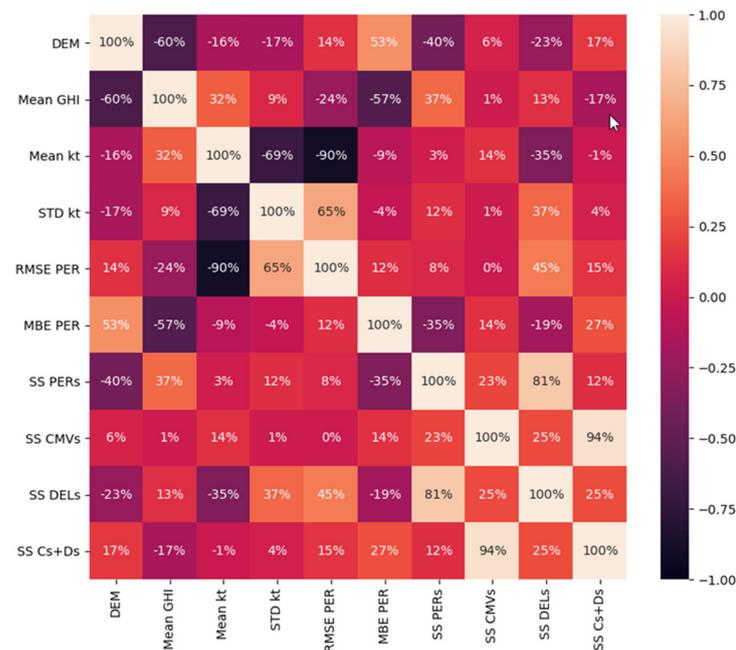


**Figure 6.** Spatial distribution of SS for the (a) CMV, (b) DEL, (c) PERs, (d) CMVs, (e) DELs, and (f) CMV + DELs models. The SS was calculated using the PER model as a reference. The ‘s’ in the model names represents smoothing.

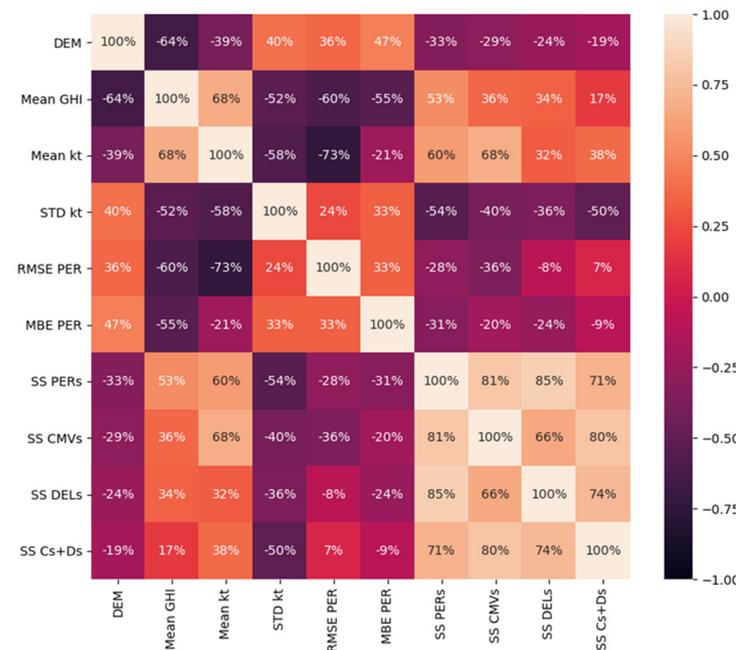


**Figure 7.** Spatial distribution of MBE of (a) CMV, (b) DEL, (c) PERs, (d) CMVs, (e) DELs, and (f) CMV + DELs models. The relative values that are divided by mean values are used. The ‘s’ in the model names represents smoothing.

The correlations between several variables and metrics were calculated, as shown in Figure 8, where part Figure 8a shows data for the entire study region (land and seas) and Figure 8b shows data for the land region only. The correlations were made using spatial data, without considering the time-series variability. There was no significant correlation between the DEM and other variables, unlike the visualized pattern. The mean GHI showed a correlation of 60% with the DEM. The influence of the terrain was correlated with the sea; therefore, a high correlation was not seen for the DEM alone (a meaningful correlation is expected when complex phenomena that are correlated with the sea or altitude are considered). The RMSE of the PER model had a high correlation with mean  $k_t$  and was close to 0.9 in all regions. Other similar variables showed low correlations, and these results were similar for the other models. Although the correlation was lower when only the land area was considered, the correlation was still over 0.7. This result can be used to estimate the reference forecasting accuracy. In terms of the SS of the forecasting models, PERs and DELs, the CMVs and hybrid models showed different distributions over the study areas because of the boundary effect of the CMV. This tendency was weaker inland; however, the difference between the CMVs and DELs models was still clear with respect to the low correlation value. Other models also showed similar but different SS distributions for the inland regions.



(a)



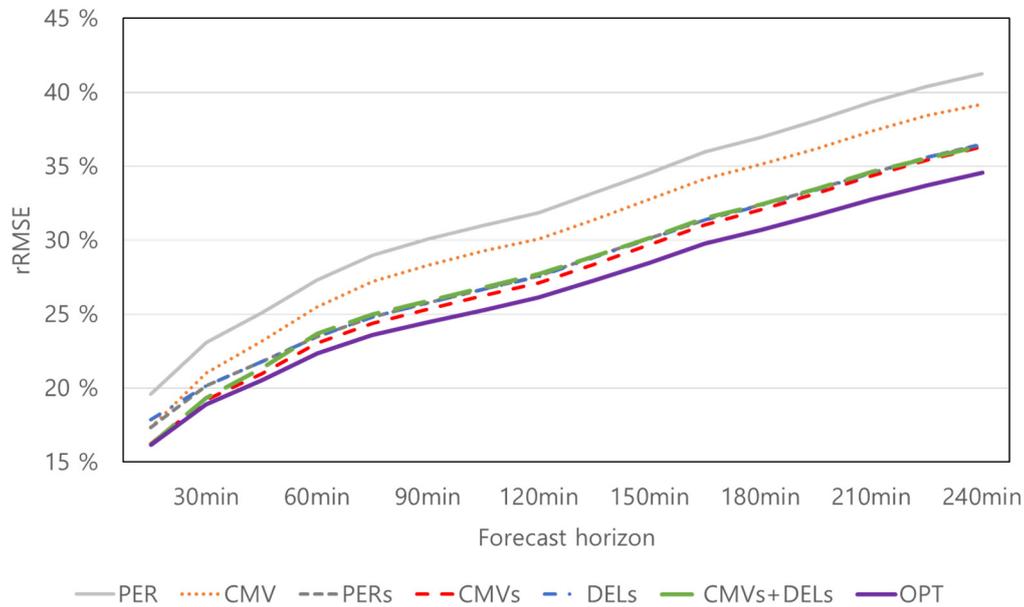
(b)

Figure 8. Correlation between several variables and metrics in (a) all study area and (b) inland of South Korea.

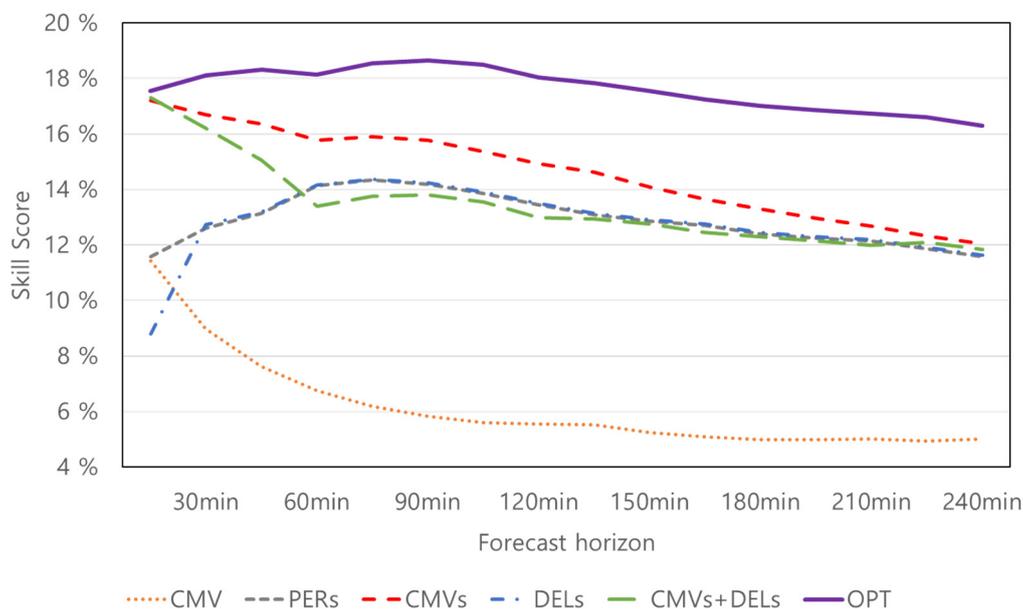
### 4.3. Optimized Model

The effects of the forecast horizon on the performance results of the models are shown in Figures 9 and 10, and the mean accuracy of each forecasting model as a function of the forecast horizon is shown in Figure 9. As the mean value of only the inland area of South Korea is considered, the influence of the CMV boundary effect is insignificant. The RMSE continuously increases with the increasing forecast horizon because of the accumulation of image forecasting errors. However, the RMSE is expected to converge at approximately 50%, which is the RMSE of the model using the constant annual average  $k_t$ .

As the effect of the forecast horizon on the RMSE is larger than the difference between any of the models, the SS in Figure 10 is more effective in the analysis. The SS curves clearly show the differences in accuracy compared to the reference PER model.



**Figure 9.** Mean accuracy of each forecasting model according to the forecast horizon, where the mean value of the inland area of South Korea is used. The OPT model represents the optimized spatiotemporal model.

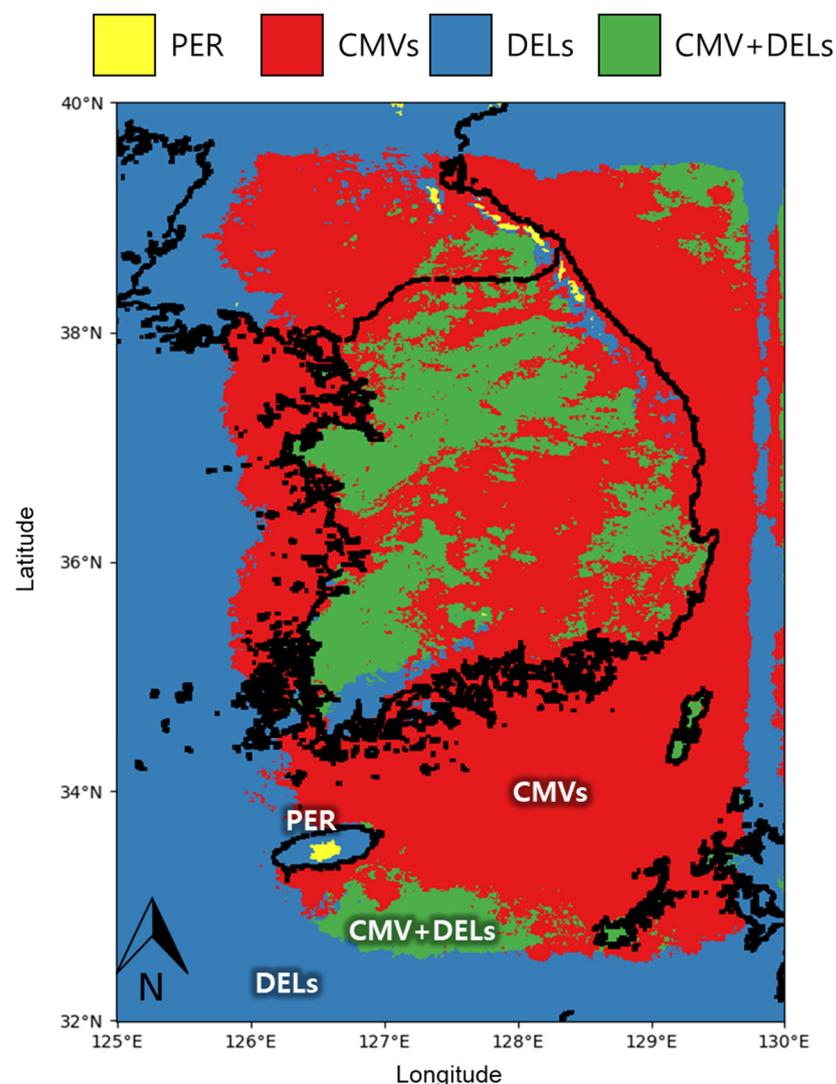


**Figure 10.** SS of each forecasting model according to the forecast horizon, where the mean value of the inland area of South Korea is used. The OPT model represents the spatio-temporal optimized model.

The performances of the PERs and CMVs models improved after smoothing (~10% increase in SS). A comparison between the PERs, CMVs, and DELs models showed that the CMVs model provided a higher accuracy than other models for most of the forecast horizons. The performance of the PER model was lower than all the other models. The hybrid model showed the mean accuracy was worse than that of CMVs. In the case of the temporally optimized model, which uses the best model for each forecast horizon, the

accuracy of that model will be the highest line in these models. This study suggests that additional factors, such as spatial optimization, can be used to further improve the model, as seen in the OPT model.

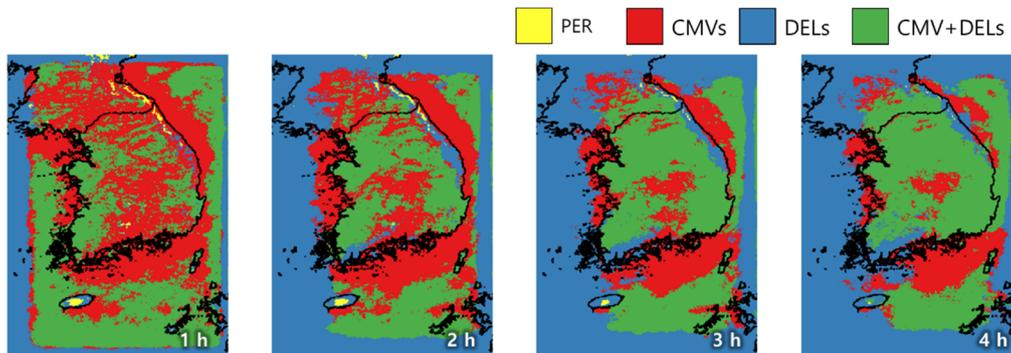
Each forecasting model exhibited a different spatiotemporal behavior and performance for specific regions. Therefore, a best-model map produced using the RMSE values of the PER, CMVs, DELs, and hybrid models for a 4-h forecast horizon is shown in Figure 11. Although the PER model provided low accuracy, it was the best model for solar forecasting in some highland regions near the sea. The DELs model was selected as the best model in areas near those of the PER model and for most of the boundary area; the DELs model was also uniquely selected for the area near the southwest coast. The CMVs model was the best model for most of the central area, while the hybrid model was selected for some inland areas. In particular, the inland lowlands near the sea in the west were selected for the CMVs model.



**Figure 11.** Map of best forecasting models considering 4-h forecast horizon.

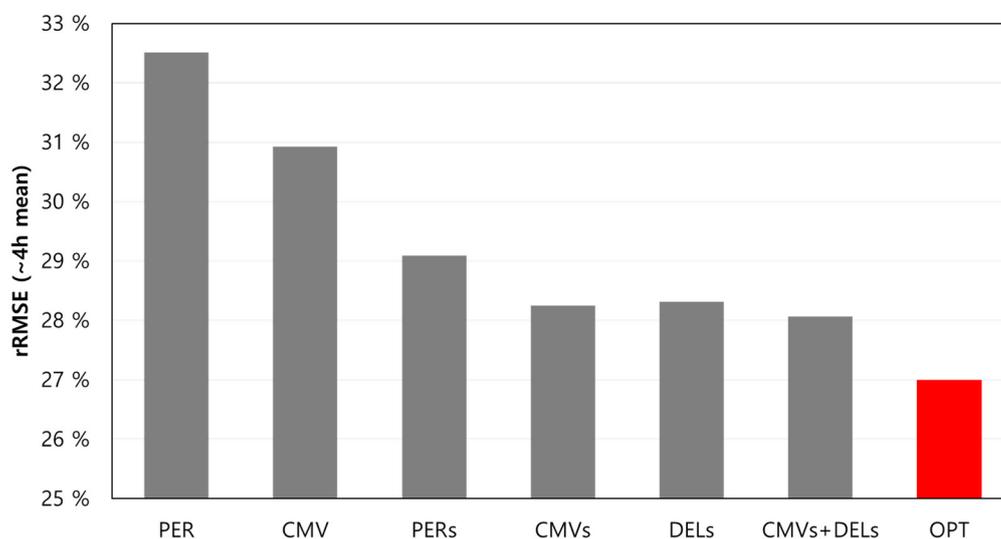
As the best forecasting model depends on the chosen forecast horizon, best-model maps for various forecast horizons (1, 2, 3, and 4 h) are also shown in Figure 12. With an increasing forecast horizon, the CMVs and hybrid models were initially the best models in many areas, but the DELs model later has expanded its area. In addition, there was an increase in the total area in which the hybrid model was the best model compared to the CMVs model when only the land area was considered. Therefore, the influence of the delta

terms increased for longer forecast horizons. A larger boundary effect was also identified when the clouds moved. The PER and DELs models continuously outperformed in some areas in the eastern mountain range and Jeju Island.



**Figure 12.** Maps for best forecasting models considering 1-, 2-, 3-, and 4-h forecast horizons.

Spatiotemporal optimization was conducted using the temporal best-model maps. The best-model maps were estimated for each forecast horizon with 15 min intervals. The OPT model had an RMSE of 27% and SS of 18%, and the latter was 21% higher than that of the best previous model (CMVs). The RMSE of all models is shown in Figure 13. Only the accuracy of the inland region was considered. The CMVs, DELs, and hybrid models provided similar accuracies, but with respect to the temporal pattern, the OPT model outperformed the highest performance models for all forecast horizons, as shown in Figure 10, particularly for forecast horizons longer than 1 h. This result can be attributed to the consideration of the  $\Delta$  terms in areas with highly dynamic cloud activity.



**Figure 13.** Mean RMSE of all forecasting models.

## 5. Conclusions

This study compared satellite-image-based solar forecasting models and their spatial characteristics. The error distributions of the CMVs and DELs models were significantly different, and the simple PER model showed the highest accuracy for some highland areas near the sea. The best model was determined with respect to the forecast horizon and location. Overall, the proposed OPT model, which includes spatiotemporal optimization, outperformed all other models, particularly for forecast horizons of more than 1 h. These results indicate that it is necessary to consider spatial data when conducting solar forecast-

ing. However, as it is difficult to consider spatial information in site-specific studies, the findings of this study are expected to be useful for complementing previous studies.

In addition to improving the forecasting performance, other significant information was derived from conducting spatial analyses. The accuracy of the forecasting models was related to the elevation of terrain and proximity to the sea, and it was highly correlated with the mean  $k_t$ . This effect should, thus, be considered when calculating and mapping solar resources. Accuracy maps could be used to identify areas that need to be analyzed by models supplemented for complex weather, or areas that are advantageous for forecasting. These maps could also be used to select sites for future photovoltaic power plants, or for designing future forecasting models. In this respect, spatial data obtained from satellite images are highly valuable. The availability of such data and spatial analysis methods are expected to make continuous contributions to solar forecasting research and to the wider photovoltaic industry.

**Author Contributions:** Conceptualization, M.O.; methodology, M.O. and B.K.; software, M.O. and C.K.K.; validation, M.O. and B.K.; investigation, C.K.K. and C.Y.; resources, C.K.K. and B.K.; data curation, C.K.K., B.K. and C.Y.; writing—original draft preparation, M.O.; writing—review and editing, H.-G.K.; visualization, B.K.; supervision, Y.-H.K. and H.-G.K.; project administration, H.-G.K.; funding acquisition, H.-G.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 2019371010006B).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

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

## References

1. Das, U.K.; Tey, K.S.; Seyedmahmoudian, M.; Mekhilef, S.; Idris, M.Y.I.; Van Deventer, W.; Horan, B.; Stojcevski, A. Forecasting of photovoltaic power generation and model optimization: A review. *Renew. Sustain. Energy Rev.* **2018**, *81*, 912–928. [[CrossRef](#)]
2. Sobri, S.; Koochi-Kamali, S.; Rahim, N.A. Solar photovoltaic generation forecasting methods: A review. *Energy Convers. Manag.* **2018**, *156*, 459–497. [[CrossRef](#)]
3. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de-Pison, F.J.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Sol. Energy* **2016**, *136*, 78–111. [[CrossRef](#)]
4. Yang, D.; Kleissl, J.; Gueymard, C.A.; Pedro, H.T.C.; Coimbra, C.F.M. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Sol. Energy* **2018**, *168*, 60–101. [[CrossRef](#)]
5. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* **2017**, *105*, 569–582. [[CrossRef](#)]
6. Yang, D.; Alessandrini, S.; Antonanzas, J.; Antonanzas-Torres, F.; Badescu, V.; Beyer, H.G.; Blaga, R.; Boland, J.; Bright, J.M.; Coimbra, C.F.M.; et al. Verification of deterministic solar forecasts. *Sol. Energy* **2020**, *210*, 20–37. [[CrossRef](#)]
7. Perez, R.; Perez, M.; Pierro, M.; Schlemmer, J.; Kivalov, S.; Dise, J.; Keelin, P.; Grammatico, M.; Swierc, A.; Ferreira, J.; et al. Perfect operational solar forecasts: A scalable strategy toward firm power generation. In Proceedings of the International Energy Agency Solar World Congress, Santiago, Chile, 4–7 November 2019; pp. 2237–2247. [[CrossRef](#)]
8. Perez, R.; Schlemmer, J.; Kivalov, S.; Dise, J.; Keelin, P.; Grammatico, M.; Hoff, T.; Tuohy, A. A New Version of the SUNY Solar Forecast Model: A Scalable Approach to Site-Specific Model Training. In Proceedings of the 2017 IEEE 45th Photovoltaic Specialists Conference (PVSC), Washington, DC, USA, 25–30 June 2018.
9. Ren, Y.; Suganthan, P.N.; Srikanth, N. Ensemble methods for wind and solar power forecasting—A state-of-the-art review. *Renew. Sustain. Energy Rev.* **2015**, *50*, 82–91. [[CrossRef](#)]
10. Menzel, W.P. Cloud Tracking with Satellite Imagery: From the Pioneering Work of Ted Fujita to the Present. *Bull. Am. Meteorol. Soc.* **2001**, *82*, 33–47. [[CrossRef](#)]
11. Lorenz, E.; Hammer, A.; Heinemann, D. Short term forecasting of solar radiation based on satellite data. In *EUROSUN2004 (ISES Europe Solar Congress)*; PSE GmbH, Solar Info Center: Freiburg, Germany, 20–23 June 2004; pp. 841–848.

12. Lorenz, E.; Kühnert, J.; Heinemann, D. Short term forecasting of solar irradiance by combining satellite data and numerical weather predictions. In Proceedings of the 27th European Photovoltaic Solar Energy Conference and Exhibition, Frankfurt, Germany, 14 September 2012; pp. 4401–4405. [\[CrossRef\]](#)
13. Lorenz, E.; Heinemann, D. *Prediction of Solar Irradiance and Photovoltaic Power*; Elsevier Ltd.: Amsterdam, The Netherlands, 2012; Volume 1, ISBN 9780080878737.
14. Kühnert, J.; Lorenz, E.; Heinemann, D. *Satellite-Based Irradiance and Power Forecasting for the German Energy Market*; Academic Press: Cambridge, MA, USA, 2013; ISBN 9780123971777.
15. 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. [\[CrossRef\]](#)
16. Perez, R.; Kivalov, S.; Schlemmer, J.; Hemker, K.; Renné, D.; Hoff, T.E. Validation of short and medium term operational solar radiation forecasts in the US. *Sol. Energy* **2010**, *84*, 2161–2172. [\[CrossRef\]](#)
17. Cros, S.; Badosa, J.; Szantái, A.; Haeffelin, M. Reliability predictors for solar irradiance satellite-based forecast. *Energies* **2020**, *13*, 5566. [\[CrossRef\]](#)
18. Giacosa, G.; Alonso-Suárez, R. Performance assessment of intra-day solar irradiation forecast in Uruguay using satellite cloud motion vectors. In Proceedings of the ISES Solar World Congress 2019, Santiago de Chile, Chile, 4–7 November 2019; pp. 2083–2090. [\[CrossRef\]](#)
19. 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 Communications (SmartGridComm), Vancouver, BC, Canada, 21–24 October 2013; pp. 708–713. [\[CrossRef\]](#)
20. Cros, S.; Sébastien, N.; Liandrat, O.; Schmutz, N. Cloud pattern prediction from geostationary meteorological satellite images for solar energy forecasting. In *Remote Sensing of Clouds and the Atmosphere XIX; and Optics in Atmospheric Propagation and Adaptive Systems XVII*; International Society for Optics and Photonics: San Diego, CA, USA, 17–21 August 2014; Volume 9242, p. 924202. [\[CrossRef\]](#)
21. Nonnenmacher, L.; Coimbra, C.F.M. Streamline-based method for intra-day solar forecasting through remote sensing. *Sol. Energy* **2014**, *108*, 447–459. [\[CrossRef\]](#)
22. Kosmopoulos, P.; Kouroutsidis, D.; Papachristopoulou, K.; Raptis, P.I.; Masoom, A.; Saint-Drenan, Y.-M.; Blanc, P.; Kontoes, C.; Kazadzis, S. Short-Term Forecasting of Large-Scale Clouds Impact on Downwelling Surface Solar Irradiation. *Energies* **2020**, *13*, 6555. [\[CrossRef\]](#)
23. Chow, C.W.; Urquhart, B.; Lave, M.; Dominguez, A.; Kleissl, J.; Shields, J.; Washom, B. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Sol. Energy* **2011**, *85*, 2881–2893. [\[CrossRef\]](#)
24. Chow, C.W.; Belongie, S.; Kleissl, J. Cloud motion and stability estimation for intra-hour solar forecasting. *Sol. Energy* **2015**, *115*, 645–655. [\[CrossRef\]](#)
25. Du, J.; Min, Q.; Zhang, P.; Guo, J.; Yang, J.; Yin, B. Short-term solar irradiance forecasts using sky images and radiative transfer model. *Energies* **2018**, *11*, 1107. [\[CrossRef\]](#)
26. Tiwari, S.; Sabzehgar, R.; Rasouli, M. Short term solar irradiance forecast based on image processing and cloud motion detection. In Proceedings of the 2019 IEEE Texas Power Energy Conference TPEC 2019, College Station, TX, USA, 7–8 February 2019; pp. 1–6. [\[CrossRef\]](#)
27. Jamaly, M.; Kleissl, J. Robust cloud motion estimation by spatio-temporal correlation analysis of irradiance data. *Sol. Energy* **2018**, *159*, 306–317. [\[CrossRef\]](#)
28. Lipperheide, M.; Bosch, J.L.; Kleissl, J. Embedded nowcasting method using cloud speed persistence for a photovoltaic power plant. *Sol. Energy* **2015**, *112*, 232–238. [\[CrossRef\]](#)
29. Escrig, H.; Batlles, F.J.; Alonso, J.; Baena, F.M.; Bosch, J.L.; Salbidegoitia, I.B.; Burgaleta, J.I. Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast. *Energy* **2013**, *55*, 853–859. [\[CrossRef\]](#)
30. Peng, Z.; Yoo, S.; Yu, D.; Huang, D.; Kalb, P.; Heiser, J. 3D cloud detection and tracking for solar forecast using multiple sky imagers. In Proceedings of the 29th Annual ACM Symposium on Applied Computing, Gyeongju, Korea, 24–28 March 2014; pp. 512–517. [\[CrossRef\]](#)
31. Zhen, Z.; Pang, S.; Wang, F.; Li, K.; Li, Z.; Ren, H.; Shafie-Khah, M.; Catalao, J.P.S. Pattern Classification and PSO Optimal Weights Based Sky Images Cloud Motion Speed Calculation Method for Solar PV Power Forecasting. *IEEE Trans. Ind. Appl.* **2019**, *55*, 3331–3342. [\[CrossRef\]](#)
32. Germann, U.; Zawadzki, I. Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Mon. Weather Rev.* **2002**, *130*, 2859–2873. [\[CrossRef\]](#)
33. Woo, W.C.; Wong, W.K. Operational application of optical flow techniques to radar-based rainfall nowcasting. *Atmosphere* **2017**, *8*, 48. [\[CrossRef\]](#)
34. Ayzel, G.; Heistermann, M.; Winterrath, T. Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1). *Geosci. Model Dev. Discuss.* **2019**, *12*, 1387–1402. [\[CrossRef\]](#)
35. Kim, C.K.; Kim, H.G.; Kang, Y.H.; Yun, C.Y. Toward Improved Solar Irradiance Forecasts: Comparison of the Global Horizontal Irradiances Derived from the COMS Satellite Imagery Over the Korean Peninsula. *Pure Appl. Geophys.* **2017**, *174*, 2773–2792. [\[CrossRef\]](#)

36. Raffel, M.; Willert, C.E.; Wereley, S.T.; Kompenhans, J.; Willert, S.; Wereley, S.T.; Kompenhans, J. *Particle Image Velocimetry: A Practical Guide*; Springer: Berlin/Heidelberg, Germany, 2007; ISBN 9783540723073.
37. Dazhi, Y.; Walsh, W.M.; Zibo, D.; Jirutitijaroen, P.; Reindl, T.G. Block matching algorithms: Their applications and limitations in solar irradiance forecasting. *Energy Procedia* **2013**, *33*, 335–342. [[CrossRef](#)]
38. Horn, B.K.P.; Schunck, B.G. Determining optical flow. *Artif. Intell.* **1981**, 185–203. [[CrossRef](#)]
39. Farneback, G. Two-frame motion estimation based on polynomial expansion. In *Scandinavian Conference on Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2003. [[CrossRef](#)]
40. Zach, C.; Pock, T.; Bischof, H. A Duality Based Approach for Realtime TV-L<sup>1</sup> Optical Flow. In *Proceedings of the Pattern Recognition: 29th DAGM Symposium, Heidelberg, Germany, 12–14 September 2007*; pp. 214–223.
41. Weinzaepfel, P.; Revaud, J.; Harchaoui, Z.; Schmid, C. DeepFlow: Large displacement optical flow with deep matching. In *Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013*.
42. Brox, T.; Bruhn, A.; Papenberger, N.; Weickert, J. High accuracy optical flow estimation based on a theory for warping. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3024, pp. 25–36. [[CrossRef](#)]
43. Sun, D.; Yang, X.; Liu, M.Y.; Kautz, J. PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018*; pp. 8934–8943. [[CrossRef](#)]
44. Urbich, I.; Bendix, J.; Müller, R. A novel approach for the short-term forecast of the effective cloud albedo. *Remote Sens.* **2018**, *10*, 955. [[CrossRef](#)]