https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

From cloud optical thickness to solar radiation: deep spatiotemporal forecasting with geostationary satellite imagery

Theano Drosoglou[1], Athanasios Natsis[1], Dimitris Dimitriadis[2], Kyriakoula Papachristopoulou[3], Stelios Kazadzis[3], Angelos Georgakis[1], Kevin Kilchhofer[3], Grigorios Tsoumakas[2], and Vassilis Amiridis[1]

[1]1Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens (IAASARS/NOA), Athens, 15236, Greece
[2]School of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[3]Physikalisch-Meteorologisches Observatorium Davos, World Radiation Center (PMOD/WRC), Davos 7260, Switzerland

Correspondence: Theano Drosoglou (tdroso@noa.gr)

Abstract. Accurate short-term cloud forecasting is essential for solar radiation and solar energy applications, where rapidly evolving cloud fields introduce substantial variability in surface solar irradiance. In this study, we develop a hybrid ConvLSTM-UNet deep learning model for 15-min-ahead forecasting of cloud optical thickness (COT) from Meteosat/SEVIRI observations over Europe. The model combines multiscale spatial feature extraction with recurrent spatiotemporal learning and employs a combined Charbonnier–SSIM loss to improve both numerical accuracy and the preservation of cloud structures. The proposed forecasting model is evaluated against real COT satellite retrievals and benchmarked against persistence and a dense optical flow cloud motion vector (CMV) approach. For 15-min-ahead forecasts, ConvLSTM-UNet achieves lower COT prediction errors and higher perceptual quality than both benchmarks, with MAE and RMSE of 1.25 and 9.52, respectively, and SSIM of 0.95. The impact of COT forecast errors on solar radiation is further assessed through the cloud modification factor (CMF), for which the proposed model achieves a forecast skill of approximately 30% relative to persistence. Its performance advantage is particularly evident under cloudy and rapidly evolving conditions, and remains evident under recursive forecasts extending to approximately 165 min ahead. These results demonstrate the potential of deep spatiotemporal learning for satellite-based short-term cloud and solar radiation forecasting.

1 Introduction

Accurate forecasting of surface solar irradiance (SSI) is becoming increasingly important as solar energy expands its contribution to global energy production and power-system decarbonization (Edenhofer et al., 2011; IEA, 2022). Reliable forecasts from minutes to several hours ahead are needed to support solar power plant operation and energy management. SSI is controlled by several atmospheric parameters, including aerosols, water vapor, and cloud properties, with clouds constituting the dominant source of variability under all-sky conditions (e.g. Fountoulakis et al., 2021). Their strong attenuation of incoming solar radiation and highly dynamic spatial and temporal behavior make accurate representation and forecasting of cloud evolution essential for solar radiation and solar energy applications.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

Cloud fields vary markedly across broad spatial and temporal scales, influenced by atmospheric dynamics, cloud formation and dissipation, and interactions with local meteorological conditions. Earth observation (EO) satellites provide an ideal platform for monitoring cloud evolution at regional and continental scales. In particular, geostationary orbit (GEO) satellites continuously observe the same portion of the Earth with revisit times of only a few minutes, enabling the tracking of rapidly evolving atmospheric phenomena (e.g. Menzel, 2001). Satellite observations from platforms such as Meteosat have therefore become widely used for weather monitoring, climate applications, and solar radiation assessment. Their high temporal resolution and broad spatial coverage also make GEO observations particularly suitable for cloud nowcasting and short-term forecasting, leading to improved forecasts compared to Numerical Weather Prediction (NWP) models (Perez et al., 2010; Pelland et al., 2013). Satellite image time series have been increasingly exploited for spatiotemporal forecasting, including weather nowcasting, cloud evolution, flood forecasting, land cover, and other environmental applications (e.g. Ravuri et al., 2021; Lammers et al., 2021; Jiang et al., 2024; Park et al., 2024; Saleem et al., 2025; Tahir et al., 2025).

Satellite-based observations are widely used to estimate SSI in real time and to generate long-term climatological datasets (e.g. Pfeifroth et al., 2024). Cloud optical thickness (COT), retrieved operationally from multispectral satellite observations, is a key variable that describes the optical attenuation of solar radiation by clouds and serves as an important input for radiative transfer calculations (e.g. Papachristopoulou et al., 2024). A complementary quantity is the cloud modification factor (CMF), which directly expresses the impact of clouds on SSI (e.g. Hou et al., 2022; Papachristopoulou et al., 2024).

Cloud forecasting for solar energy applications has traditionally relied on the persistence approach and optical flow techniques. Persistence assumes that the previous cloud state remains unchanged throughout the forecast horizon and provides a simple but useful baseline (e.g. Pelland et al., 2013; Kallio-Myers et al., 2020). Cloud motion vector (CMV) approaches estimate pixel-level cloud displacement between consecutive satellite images (e.g. Bouguet, 2000; Farnebäck, 2003) and have been widely applied to geostationary satellite imagery for forecasting cloud properties and SSI (e.g. Cros et al., 2014; Nonnenmacher and Coimbra, 2014; Kosmopoulos et al., 2024). Both of the above approaches remain limited in representing complex, nonlinear processes such as cloud formation and dissipation.

The rapid advances in artificial intelligence (AI) and deep learning (DL) have enabled data-driven approaches to learn complex spatiotemporal relationships directly from historical satellite observations (Moskolaï et al., 2021; Miller et al., 2024). DL models can capture dynamic processes that are not explicitly represented in optical flow approaches. Early approaches employed recurrent neural networks (RNNs), including long short-term memory (LSTM) and gated recurrent unit (GRU) architectures, to capture temporal dependencies (e.g. Xiao et al., 2019; Bi et al., 2020). Convolutional neural network (CNN) architectures were subsequently incorporated to account for spatial dependencies (e.g. Weyn et al., 2020), leading to hybrid CNN-RNN approaches that leverage the capabilities of both techniques (e.g. Kamdem De Teyou et al., 2020; Qiao et al., 2021). In particular, the convolutional LSTM (ConvLSTM) architecture introduced by Shi et al. (2015) enables simultaneous spatial feature extraction and temporal modeling and has become a well-established and efficient framework for satellite-based spatiotemporal forecasting. ConvLSTM and related architectures, such as ConvGRU and TrajGRU (Ballas et al., 2016; Shi et al., 2017), have demonstrated strong capabilities for modeling cloud dynamics and other atmospheric processes. Nevertheless,

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

conventional pixel-wise losses used in DL forecasting can generate overly smooth or blurry images, motivating the use of loss functions that better preserve spatial structures (e.g. Moskolaï et al., 2021; Dai et al., 2023).

DL-based cloud forecasting from Meteosat observations over Europe remains limited. For example, Partio et al. (2025) introduced a CNN model based on U-Net architecture (Ronneberger et al., 2015) to predict total cloud cover from Meteosat observations, outperforming traditional NWP models and optical flow approaches. Most existing studies have instead focused on East Asian geostationary platforms, i.e., the FY and Himawari satellite series (e.g. Tan et al., 2019; Xu et al., 2019; Su et al., 2020; Wang et al., 2022; Xia et al., 2024; Perera et al., 2026), while fewer studies have employed GEO satellite images over Americas (e.g. Nguyen et al., 2023; Marchesoni-Acland et al., 2023; Peng et al., 2026). Operational COT forecasting over Europe continues to rely predominantly on optical flow techniques (e.g. Sirch et al., 2017; Papachristopoulou et al., 2024). Furthermore, relatively few studies have investigated how uncertainties in DL-based cloud forecasts propagate to solar radiation estimates (e.g. Perera et al., 2026), despite the importance of this relationship for solar energy applications.

In this study, we develop a hybrid spatiotemporal encoder-decoder model integrating ConvLSTM and U-Net components for short-term COT forecasting using geostationary Meteosat imagery (ConvLSTM-UNet hereafter). The model combines the multiscale spatial feature extraction capability of the U-Net architecture with recurrent ConvLSTM layers that capture temporal cloud evolution, while preserving the spatial continuity of large satellite scenes. To reduce blur effect in the predicted time frames while retaining temporal consistency, a combined loss function was employed that balances robust pixel-wise reconstruction with perceptual similarity. Beyond evaluating COT forecast accuracy, we assess the propagation of cloud prediction errors on solar radiation estimates through CMF and benchmark the proposed model against a widely-used optical flow methodology and the persistence approach.

The main contributions of this work are: **(a)** a hybrid framework combining ConvLSTM and U-Net architectures for short-term COT forecasting from Meteosat observations, designed to jointly capture multiscale spatial structure and temporal cloud evolution while handling large satellite scenes; **(b)** a combined loss formulation that improves the balance between quantitative accuracy and structural realism in cloud predictions; and **(c)** an assessment of the downstream impact of COT forecast errors on cloud-induced solar attenuation through CMF, linking cloud forecasting performance directly to solar radiation forecasting applications.

2 Data and Methodology

This section describes the data, forecasting approaches, and evaluation framework adopted in this study. First, the Meteosat observations and the retrieval of COT are presented, followed by the computation of the CMF used to assess the impact of cloud forecasting errors on SSI. The persistence and CMV approaches are then introduced as baseline forecasting methods, together with the quantitative and perceptual metrics used for model evaluation. Finally, the proposed ConvLSTM-UNet forecasting framework is described, including its architecture, loss function, training procedure, and implementation details. An overview of the data processing, forecasting, and evaluation workflow is provided in Fig. 1.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

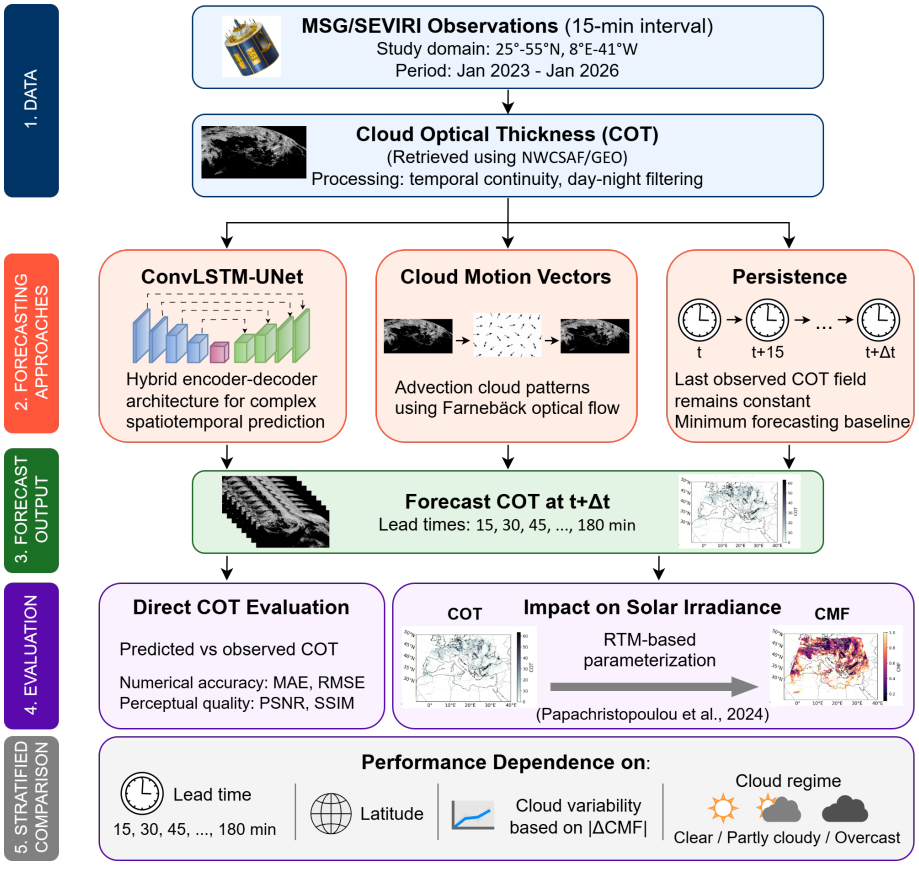


Figure 1. Schematic overview of the forecasting and evaluation workflow.

2.1 Satellite Cloud Observations

Meteosat satellites follow the Earth's rotation on a geostationary orbit at 36,000km above the equator providing continuous monitoring of around one-third of the Earth's surface. The Meteosat First Generation (MFG) series began its operation in November 1977 and had been providing observations until 2017. With the launch of the first satellite of the Meteosat Second Generation (MSG) mission in August 2002, improved geostationary images of higher quality are available to further support medium and short-term weather forecasts and climate studies. Following the legacy of Meteosat's first and second-generation missions, the first Meteosat Third Generation (MTG) satellite was successfully launched in December 2022 and it has been providing images of higher spatial and temporal resolution since October 2024.

In this study, observations from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard MSG satellites have been employed. The MSG SEVIRI data are freely accessed on the EUMETSAT Data Store. The spectral capabilities of MSG allow for accurate cloud cover analysis. SEVIRI is the primary instrument of MSG and has the capacity to observe the Earth and its atmosphere in 12 spectral channels, 11 of which provide measurements with a spatial resolution of 3km at nadir with 8

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

channels in the thermal infrared, providing data about the temperatures of clouds among other information. The last channel, which is known as the High Resolution Visible (HRV) channel, provides measurements with a higher sampling distance of 1km at the sub-satellite point.

COT fields are retrieved using the visible and near-infrared bands centered on 0.6μm and 1.6μm, respectively, and the thermal infrared bands centered on 8.7μm and 10.8μm as input to the software package NWCSAF/GEO v2021.3 provided by the EUMETSAT Satellite Application Facilities for Nowcasting and Very Short Range Forecasting, NWC SAF (Météo-France, 2016; Derrien and Gléau, 2005). Each available time frame (every 15 minutes) was cropped to a domain that covers the area with latitudes 25°-55°N and longitudes 8°E-41°W, resulting to high resolution images of size 1536×768. Frames where the domain of interest was not lit at least partially by the sun were filtered out by calculating the solar zenith angle (SZA) at the boundaries of the domain.

For the proposed DL model, the sequence samples were generated using sliding temporal windows over chronologically ordered satellite-derived COT images. A sequence was retained only if all adjacent image pairs satisfied strict temporal continuity constraints corresponding to the expected 15-minute acquisition interval ($\Delta t = 15min$). Sequences containing missing frames or temporal discontinuities, e.g., due to gaps in data acquisition or transitions from one day to another, were discarded to ensure physically consistent temporal dynamics during training. The dataset was partitioned into two subsets; one used for training and validation (fine-tuning) and one used for the final testing, comprising 90% and 10% of the data, respectively. The training and validation sequence samples were subsequently split based on an 80%/20% splitting strategy and, together, they cover a period of over 2.5 years between 01/01/2023 and 10/09/2025 comprising 47781 frames in total. The testing subset corresponds to the subsequent 4-month period extending from 11/09/2025 until 14/01/2026 and consists of 5310 frames.

2.2 Cloud Modification Factor (CMF)

Clouds are one of the major factors controlling the amount of solar radiation reaching the Earth's surface. Thus, uncertainties in cloud forecasting directly affect the accuracy of SSI prediction and, therefore, solar energy forecasting applications. COT is a satellite-derived cloud property that can be retrieved directly in near real time from geostationary observations without requiring extensive post-processing. COT quantifies the optical attenuation of solar radiation by clouds and is strongly related to the amount of downwelling solar irradiance reaching the surface (e.g. Taylor et al., 2016; Qu et al., 2017). However, its relationship with surface irradiance is nonlinear and depends on additional atmospheric and geometric conditions.

To overcome these limitations, the impact of cloud prediction uncertainty can be quantified in terms of cloud modification factor (CMF), which is often used as a more direct representation of the cloud effect on SSI (Calbó et al., 2005). CMF is expressed as the ratio between the global horizontal irradiance (GHI) under actual atmospheric conditions and the corresponding clear-sky (cloudless) global horizontal irradiance (GHI_{clr}):

$$\mathrm{CMF} = \frac{\mathrm{GHI}}{\mathrm{GHI_{clr}}} \tag{1}$$

The clear-sky irradiance is commonly estimated using radiative transfer models (RTMs) by simulating the propagation of solar radiation through the atmosphere in the absence of clouds.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

CMF typically ranges between 0 and 1, where values close to 1 correspond to cloud-free conditions and lower values indicate increasing cloud attenuation. Values approaching 0 are associated with optically thick clouds and overcast conditions. In reality, CMF can exceed 1 in relatively rare cases of solar enhancement induced by clouds (e.g. Natsis et al., 2024) or due to uncertainties in the clear-sky irradiance estimation. Since CMF directly represents the relative decrease caused by clouds in SSI, it serves as a physically meaningful quantity for assessing how cloud prediction errors propagate into SSI forecasting uncertainty. In particular, CMF is highly sensitive to variations in low and moderate COT values, where even small changes in COT can lead to significant variations in SSI.

In operational solar forecasting applications, empirical relationships between COT and CMF can be employed to efficiently estimate cloud radiative effects without the need for computationally expensive RTM simulations. In this study, we apply the CMF parameterizations developed by Papachristopoulou et al. (2024) using RTM-simulated GHI values for different COT conditions along with the corresponding clear-sky irradiance values for the same SZA. The derived CMF values for each SZA were subsequently fitted as a function of COT using a hyperbolic tangent expression:

$$\mathrm{CMF} = 1 - [tanh(\mathrm{COT}^{a})]^{b} \tag{2}$$

where the parameters a and b are expressed as polynomial functions of SZA:

$$a = 2.24 \times 10^{-1} + 2.81 \times 10^{-4}\,\mathrm{SZA} - 2.18 \times 10^{-5}\,\mathrm{SZA}^2 + 3.71 \times 10^{-7}\,\mathrm{SZA}^3 - 2.65 \times 10^{-9}\,\mathrm{SZA}^4 \tag{3a}$$

$$b = 12.2 + 5.27 \times 10^{-3}\,\mathrm{SZA} - 2.24 \times 10^{-3}\,\mathrm{SZA}^2 + 8.33 \times 10^{-6}\,\mathrm{SZA}^3 + 3.94 \times 10^{-8}\,\mathrm{SZA}^4 \tag{3b}$$

This parameterization is used to estimate the cloud-related irradiance attenuation directly from MSG COT observations and model predictions while accounting for solar geometry effects through the SZA dependence.

2.3 Persistence approach

Persistence forecasting is one of the simplest and most widely used baseline approaches in time-series prediction. The underlying assumption is that the current state of the system will remain unchanged throughout the prediction horizon. Thus, the forecast at a future time $t + \Delta t$ is assumed to be equal to the previous observation at time t:

$$\hat{x}(t + \Delta t) = x(t) \tag{4}$$

where $x(t)$ denotes the observed value at current time and $\hat{x}(t + \Delta t)$ is the predicted value after a lead time Δt. When the persistence model is recursively applied for multiple time steps, the predicted value at each time step is assumed to remain equal to the last observation $x(t)$.

Despite its simplicity, persistence often provides surprisingly competitive performance for short forecasting horizons and stable atmospheric conditions (e.g. cloudless skies with low or constant aerosol load). However, its accuracy decreases rapidly as the forecast horizon increases because clouds are highly dynamic and undergo continuous advection, deformation, growth, and decay driven by atmospheric processes. It is commonly used as a benchmark model to evaluate the performance of more sophisticated forecasting techniques, e.g., in solar radiation and cloud forecasting (Kallio-Myers et al., 2020). In this study,

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

we use the persistence approach as the minimum baseline, since demonstrating an improvement over persistence is generally considered a minimum requirement to establish the practical value of a predictive model.

2.4 Cloud motion vectors (CMV)

Optical flow is a fundamental concept in computer vision that enables the analysis of motion within image sequences. It describes the apparent displacement of objects, surfaces, edges, or intensity patterns between consecutive frames, caused by relative motion between the sensor and the observed scene by assigning a motion vector to each pixel of the image (or to selected features). Optical flow-based CMV techniques have been widely used for both sky-image and space-borne cloud forecasting (Pelland et al., 2013), since they offer several advantages such as dense motion estimation at pixel level, relatively low computational cost, and the ability to produce sharp and visually consistent predictions. However, optical flow approaches exhibit also important limitations, including inability to model complex atmospheric processes such as cloud growth and decay, dependence on brightness constancy, which may not hold for multispectral satellite data, accumulation of errors in multi-step forecasting, and sensitivity to noise and illumination changes. In this study, we calculated CMVs using the well-established dense optical flow algorithm provided by Farnebäck (2003) to benchmark our proposed DL-based model. For this, we utilized the OpenCV python library (Bradski, 2000) and followed an implementation of the Farnebäck CMV approach very similar to the one described in detail by Kosmopoulos et al. (2020).

Since the CMV algorithm accepts only 8-bit grayscale images within the value range [0, 255], the actual COT values were normalized to this grayscale range. COT values can vary significantly within a wide range between 0 and very high values (»255). However, COT values, that effectively interact with downwelling solar irradiance and cause significant variation in the amount reaching the ground, lay within the lower values [0, 3]. In addition, the observation frequency of COT decreases as its value increases, with values larger than 50 being quite rarely observed. To account for these characteristics of COT, we applied a modified normalization function that expands the low but radiatively important portion of the COT range into a wider grayscale interval, thereby producing finer grayscale quantization for COT values below 10 (Kosmopoulos et al., 2020):

$$\mathrm{COT}_{\mathrm{norm}} = \sqrt[8]{128^8\,\mathrm{COT}} \tag{5}$$

To transform the forecasted normalized images into actual COT fields in the original value range, we use the inverse function:

$$\mathrm{COT} = \left(\frac{\mathrm{COT}_{\mathrm{norm}}}{128}\right)^8 \tag{6}$$

2.5 Quantitative Metrics

We adopted several commonly used metrics for model optimization and accuracy assessment. The evaluation of spatiotemporal forecasting requires quantitative metrics capable of assessing both numerical prediction accuracy and perceptual image quality. In satellite image forecasting applications, pixel-wise agreement alone is often insufficient since realistic predictions must also preserve textures and spatial patterns. For this reason, multiple complementary evaluation metrics are commonly employed.

Regarding numerical accuracy, we use several common metrics of the pixel-wise difference between predictions and actual observations. The Pearson correlation coefficient R measures the strength and direction of the linear relationship between

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

prediction $\hat{y}$ and actual observation y:

$$R = \frac{\mathrm{Cov}(y,\hat{y})}{\sigma_y \sigma_{\hat{y}}} = \frac{n\sum y\hat{y} - (\sum y)(\sum \hat{y})}{\sqrt{\left[n\sum y^2 - (\sum y)^2\right]\left[n\sum \hat{y}^2 - (\sum \hat{y})^2\right]}} \tag{7}$$

Mean bias (MB) quantifies the average signed difference between predicted and true values:

$$\mathrm{MB} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i) \tag{8}$$

where $\hat{y}_i$ and y_i denote the predicted and observed values, respectively, for the i-th sample and N is the total number of samples (or pixels in the case of imaging data). Mean absolute error (MAE) measures the average absolute difference between the predicted and reference values:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{9}$$

Mean squared error (MSE) computes the average squared difference between predicted and reference values, while root mean squared error (RMSE) is defined as the square root of the MSE:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{10}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{11}$$

Forecast Skill (FS) is a relative performance metric that evaluates the predictive capability of a forecasting model with respect to a reference or baseline method. In this study, the persistence model was used as a baseline and the FS is computed based on the rRMSE as:

$$\mathrm{FS} = 1 - \frac{\mathrm{rRMSE}_{\mathrm{model}}}{\mathrm{rRMSE}_{\mathrm{pers}}} \tag{12}$$

where $rRMSE_{model}$ is the relative RMSE of the forecasting technique to be evaluated against the persistence forecast and $rRMSE_{pers}$ is the corresponding relative RMSE obtained using the persistence approach. The rRMSE was computed with respect to the mean of the observed values. FS is a dimensionless quantity that provides an intuitive interpretation of model performance. A positive FS indicates that the forecasting model outperforms the persistence baseline, with FS = 1 corresponding to perfect forecast ($rRMSE_{model} = 0$), whereas negative values of FS indicate that the persistence approach is more accurate than the proposed model. An FS equal to 0 denotes that the model performs equivalently to persistence.

Although the above metrics quantify pixel-wise accuracy, they do not explicitly evaluate the perceptual quality or structural consistency of predicted images. The structural similarity index measure (SSIM), which is a widely used metric in computer

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

vision, is employed to address this limitation by comparing images in terms of brightness, contrast, and structural similarity. For two image patches x and y, SSIM is defined as:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{13}$$

where μ_x and μ_y are local mean intensities, σ_x^2 and σ_y^2 are local variances of pixel intensities, σ_{xy} is the covariance between the two images, and C_1, C_2 are constants to prevent instability when the denominators are close to zero. SSIM values range between -1 and 1, with values closer to unity indicating higher structural similarity.

Another commonly used image quality metric is the peak signal-to-noise ratio (PSNR), which is derived using the MSE:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}}\right) \tag{14}$$

where MAX_I is the maximum possible pixel value. PSNR evaluates the ratio between the maximum signal and the reconstruction error. Higher PSNR values indicate better image quality and lower distortion.

2.6 ConvLSTM-UNet design and implementation

2.6.1 Model architecture

A DL forecasting framework was developed for short-term prediction of COT fields from geostationary satellite imagery, based on a hybrid spatiotemporal encoder-decoder architecture integrating ConvLSTM and U-Net components (Fig. 2). The proposed model combines the spatial feature extraction capability of U-Net architectures with the temporal sequence modeling capability of recurrent ConvLSTM layers, enabling simultaneous learning of cloud spatial patterns and cloud evolution dynamics. The model receives as input a sequence of past satellite-derived COT images sampled at 15-minute intervals and predicts the subsequent COT field at the next forecast step (15 min ahead). The input tensors can be defined as:

$$X \in \mathbb{R}^{T\times H\times W\times 1} \tag{15}$$

where T denotes the temporal sequence length, H and W represent the spatial dimensions of the image, and the final dimension corresponds to the single normalized COT channel. Images are normalized to the range [0, 1], as it helps neural network learn faster and more accurately compared to the [0, 255] interval utilized for the CMV approach. The model performs many-to-one sequence forecasting, mapping a multitemporal input sequence $X_{t-T+1:t}$ into a single predicted future COT image $\hat{Y}$:

$$\hat{Y} = f\left(X_{t-T+1:t}\right) \tag{16}$$

The encoder follows a hierarchical multi-scale design inspired by the classical U-Net architecture. The spatial feature extraction is performed independently for each time frame using TimeDistributed convolutional operations, preserving temporal ordering throughout the encoding process. The encoder consists of four downsampling blocks containing two 3×3 convolutional layers with same padding, Rectified Linear Unit (ReLU) activation, batch normalization after each convolution, and 2×2

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

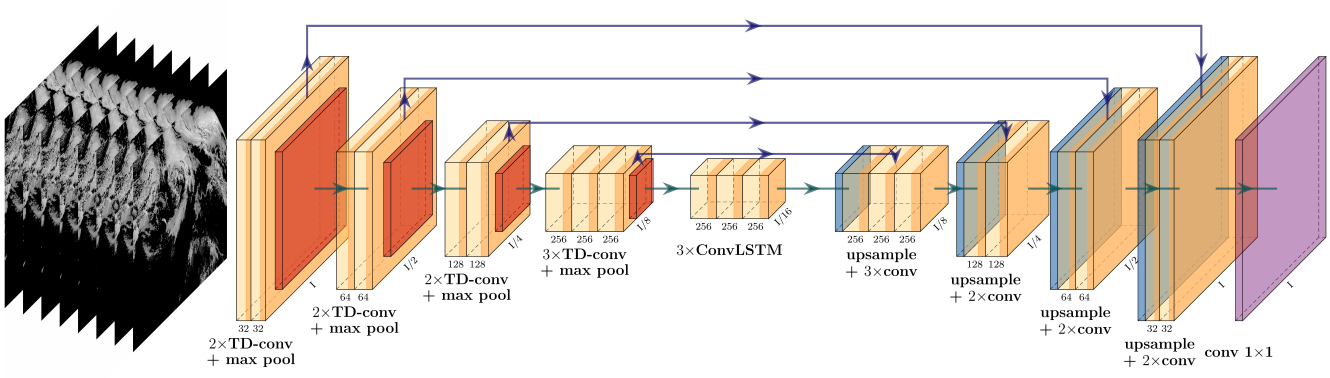


Figure 2. The proposed model architecture for satellite-based cloud nowcasting integrating ConvLSTM and U-Net components.

max-pooling for spatial downsampling. The number of convolutional filters at the encoder levels are 32, 64, 128, and 256, consecutively.

For each encoder level l and time step t, the convolutional block produces feature maps:

$$F_t^{(l)} \in \mathbb{R}^{H_l \times W_l \times C_l} \tag{17}$$

where H_l and W_l denote the spatial dimensions after pooling, and C_l denotes the number of filters at level l. At the deepest encoder level, an additional dilated convolution with dilation rate equal to 2 is introduced to enlarge the receptive field and improve the representation of mesoscale cloud structures without further reducing spatial resolution. Skip connections are retained at each encoder level to facilitate high-resolution feature fusion during decoding. For this, the encoder preserves the complete temporal stack of feature maps at each scale.

The encoded feature maps are forwarded to a stacked ConvLSTM bottleneck responsible for modeling temporal cloud evolution and motion dynamics. The bottleneck contains three stacked ConvLSTM layers with 3×3 convolution kernels, same padding, ReLU activation, and batch normalization. Unlike conventional LSTM layers that use fully connected transformations, ConvLSTM replaces dense matrix multiplications with convolutional operations, enabling simultaneous spatial and temporal modeling (Shi et al., 2015). The ConvLSTM state transitions are defined as follows:

$$f_t = \sigma\left(W_f * X_t + U_f * H_{t-1} + b_f\right) \tag{18a}$$

$$i_t = \sigma\left(W_i * X_t + U_i * H_{t-1} + b_i\right) \tag{18b}$$

$$g_t = \tanh\left(W_g * X_t + U_g * H_{t-1} + b_g\right) \tag{18c}$$

$$o_t = \sigma\left(W_o * X_t + U_o * H_{t-1} + b_o\right) \tag{18d}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot g_t \tag{18e}$$

$$H_t = o_t \odot \tanh\left(C_t\right) \tag{18f}$$

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

where X_t is the input feature map at time t, H_{t-1} is the previous hidden state, C_{t-1} is the previous cell state, $*$ denotes convolution, $\odot$ denotes element-wise multiplication, and σ denotes the sigmoid activation function. The forget gate f_t controls the amount of information retained from previous temporal states. Values close to 1 preserve prior memory, while values near 0 discard outdated information. The candidate cell state g_t introduces new information into memory, enabling the model to adapt to evolving cloud structures. The implementation follows the TensorFlow ConvLSTM formulation, which omits peephole connections originally proposed by Shi et al. (2015). The final ConvLSTM layer collapses the temporal dimension into a latent forecast representation, effectively encoding cloud motion patterns and temporal dependencies into a compact spatiotemporal embedding.

The decoder mirrors the encoder structure through progressive upsampling and feature reconstruction. Its purpose is to recover high-resolution spatial information from the latent ConvLSTM representation. Each decoder stage contains 2×2 upsampling, concatenation with corresponding encoder skip connections, two 3×3 convolutional layers with same padding, ReLU activation, and batch normalization. A dilated convolution is also included at the deepest decoder level to preserve large-scale contextual information during reconstruction.

A notable architectural feature of the proposed framework is the hybrid temporal skip connection mechanism. During encoding, skip tensors preserve the complete temporal structure:

$$S^{(l)} = \left\{ F_1^{(l)}, F_2^{(l)}, \ldots, F_T^{(l)} \right\} \tag{19}$$

However, during decoding, only the most recent temporal representation is concatenated with the decoder features:

$$D^{(l)} = \mathrm{Concat}\left(U^{(l)}, F_T^{(l)} \right) \tag{20}$$

where $U^{(l)}$ denotes the upsampled decoder features and $F_T^{(l)}$ is the encoder representation from the final input timestep. This hybrid strategy combines the temporal abstraction learned by the ConvLSTM bottleneck with the high-resolution spatial detail preserved from the most recent observation. This design is particularly suitable for cloud nowcasting, where recent spatial cloud patterns strongly influence short-term future evolution. The final predicted image is generated using a 1×1 convolution followed by a sigmoid activation function:

$$\hat{Y} = \sigma\left(W * F + b \right) \tag{21}$$

where F denotes the final decoder feature representation, W and b are trainable parameters, and $\hat{Y}$ is the normalized predicted COT field. The sigmoid activation constrains predictions to the normalized range [0, 1].

Several robust loss functions have been introduced for image restoration and forecasting tasks to mitigate some of the limitations of conventional pixel-wise losses. One widely adopted approach is the Charbonnier loss (Charbonnier et al., 1994), which can be viewed as a variant of the MAE loss. The Charbonnier loss is defined as:

$$L_{\mathrm{Char}} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\left(y_i - \hat{y}_i \right)^2 + \epsilon^2} \tag{22}$$

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

where y_i denotes the target pixel, $\hat{y}_i$ denotes the predicted pixel, and ϵ is a small positive constant (here equal to 10^{-6}) introduced for numerical stability. Compared to MSE, the Charbonnier loss is less sensitive to large outliers while still maintaining stable gradients during optimization. At the same time, it preserves image sharpness and localized structures more effectively than quadratic loss formulations. The Charbonnier loss has been widely used in image restoration, super-resolution and denoising tasks, as well as video prediction applications (e.g. Zhang et al., 2025). In spatiotemporal satellite forecasting tasks, the Charbonnier loss provides a favorable balance between numerical accuracy and preservation of localized fine-scale spatial details, making it particularly suitable for modeling evolving atmospheric patterns.

Although the Charbonnier loss improves robustness and preserves sharper structures compared to conventional reconstruction losses, it still primarily operates at pixel level and does not explicitly enforce perceptual or structural consistency. To address this limitation and improve the quality of the predicted satellite images while reducing excessive smoothing effects, a combined loss function integrating Charbonnier loss with SSIM was employed in this study:

$$L = \alpha L_{\text{Char}} + (1-\alpha)\,(1-\text{SSIM}) \tag{23}$$

where $a = 0.7$ controls the relative contribution of the two components and was experimentally determined. The Charbonnier term promotes accurate pixel-wise reconstruction and stable optimization, while the SSIM component helps preserving structural information, textures, and perceptual image quality. By optimizing both objectives, the proposed model is able to generate predictions that are both quantitatively accurate and visually realistic (Zubair et al., 2025). This combined loss function is particularly beneficial for cloud forecasting applications, where maintaining sharp cloud boundaries, evolving structures, and mesoscale spatial patterns is essential. In contrast to conventional MSE-based optimization, which often produces overly smooth and blurry predictions, the integration of SSIM helps preserve fine-scale cloud patterns and improves the perceptual consistency of the predicted frames.

2.6.2 Implementation

Model development and training were implemented using TensorFlow with the Keras API (Abadi et al., 2016). The ConvLSTM layers were built using the TensorFlow ConvLSTM2D module (https://www.tensorflow.org/api_docs/python/tf/keras/layers/ConvLSTM2D, last access: 22 September 2026). The U-Net component was implemented from scratch. Optimization was performed using the Adam optimizer with an initial learning rate of 10^{-4} and L2 weight regularization with a coefficient of 10^{-4} applied to convolutional kernels. The model was trained using mini-batch gradient descent and a batch size of 4. After extensive experimentation with different sequence lengths of consecutive COT images given as model input, a temporal sequence length equal to 8 was selected.

The hyperparameters were adjusted based on the model performance on the validation set using the minimization of validation MSE as the optimization objective, ensuring pixel-wise reconstruction accuracy. Initial hyperparameter tuning was performed manually, followed by automated optimization using the Optuna optimization framework (Akiba et al., 2019) to enable efficient exploration of the hyperparameter space. A total of 40 optimization trials were executed with the Optuna tool. The search space for both the ad hoc and Optuna experiments included learning rate in the range between 5×10^{-5} and

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

5×10^{-4}, the batch size values $\{2, 4, 8\}$, and temporal input sequence length $T \in [2, 8]$. Nevertheless, the impact of the explored hyperparameters on model performance was limited as indicated by the variations in the derived MSE, with the learning rate being the most important parameter according to the Optuna output.

The MSE metric was also used for checkpoint selection and early stopping. All experiments were conducted on Google Cloud infrastructure using NVIDIA A100 GPUs. The high-memory GPU environment enabled efficient training of deep spatiotemporal models with multiple ConvLSTM layers and high-resolution satellite imagery.

3 Results and discussion

The performance evaluation of the proposed model was conducted on the test dataset. The predicted cloud fields in 15-minute intervals were compared to the corresponding actual MSG/SEVIRI observed frames. The model performance was evaluated using both numerical metrics, such as MAE and RMSE, and perceptual quality metrics, such as PSNR and SSIM. Numerical metrics quantify pixel-wise reconstruction accuracy, PSNR measures reconstruction fidelity in relation to prediction noise, while SSIM evaluates perceptual structural similarity between predicted and target cloud fields, accounting for luminance, contrast, and structural consistency. The combination of the above metrics provides complementary assessment of numerical forecast accuracy, visual realism, and structural preservation of cloud formations.

The model performance was evaluated using both the directly-predicted COT images and the COT-derived CMF frames. CMF is used to assess how the cloud prediction uncertainty would propagate on a subsequent solar energy forecasting pipeline. The model is benchmarked against CMV-based predictions and persistence forecasts, with the latter being used as the minimum baseline to establish the practical value of our framework. We additionally investigate how different parameters, such as latitude, cloud regimes, and cloud variability, are linked to cloud forecasting performance. Finally, we further examine the model behavior when used to forecast cloud fields over time horizons larger than the next 15 minutes. An overview of the evaluation workflow is included in Fig. 1.

3.1 Cloud optical thickness forecasting performance

First, we compare the predicted COT images for the next 15 minutes with the corresponding actual MSG/SEVIRI observations. The overall performance on COT predictions is summarized in Table 1. The proposed ConvLSTM-UNet framework outperforms CMV-based and persistence forecasting in both numerical accuracy and perceptual image quality. Specifically, our model achieves a performance gain of about 15.5% on MAE and 11.2% on RMSE compared to CMV. The corresponding gains compared to the persistence approach are about 34.9% and 28.7% for MAE and RMSE, respectively. The average PSNR and SSIM values for ConvLSTM-UNet were found to be 29.71 and 0.95, respectively, showing an improvement of about 3.5% and 1.1% compared to CMV and about 10.9% and 2.2% relative to persistence results.

Our CMV results (PSNR $\approx$ 28.7 dB and SSIM $\approx$ 0.94) compare favorably with those reported by Su et al. (2020), who achieved a PSNR of 22.98 dB and an SSIM of 0.41 using an optical flow-based approach for cloud motion prediction on their test dataset. Their proposed Multi-GRU-RCN model, however, achieved a PSNR of 29.62 dB, which is comparable to

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

Table 1. Overall performance statistics for CMV-based and ConvLSTM-UNet COT forecasts 15 min ahead. The arrows indicate whether lower or higher values are better for each metric. The best results per metric are in bold.

	MAE↓	RMSE↓	PSNR [dB]↑	SSIM↑
Persistence	1.92	13.36	26.78	0.93
CMV	1.48	10.72	28.70	0.94
ConvLSTM-UNet	**1.25**	**9.52**	**29.71**	**0.95**

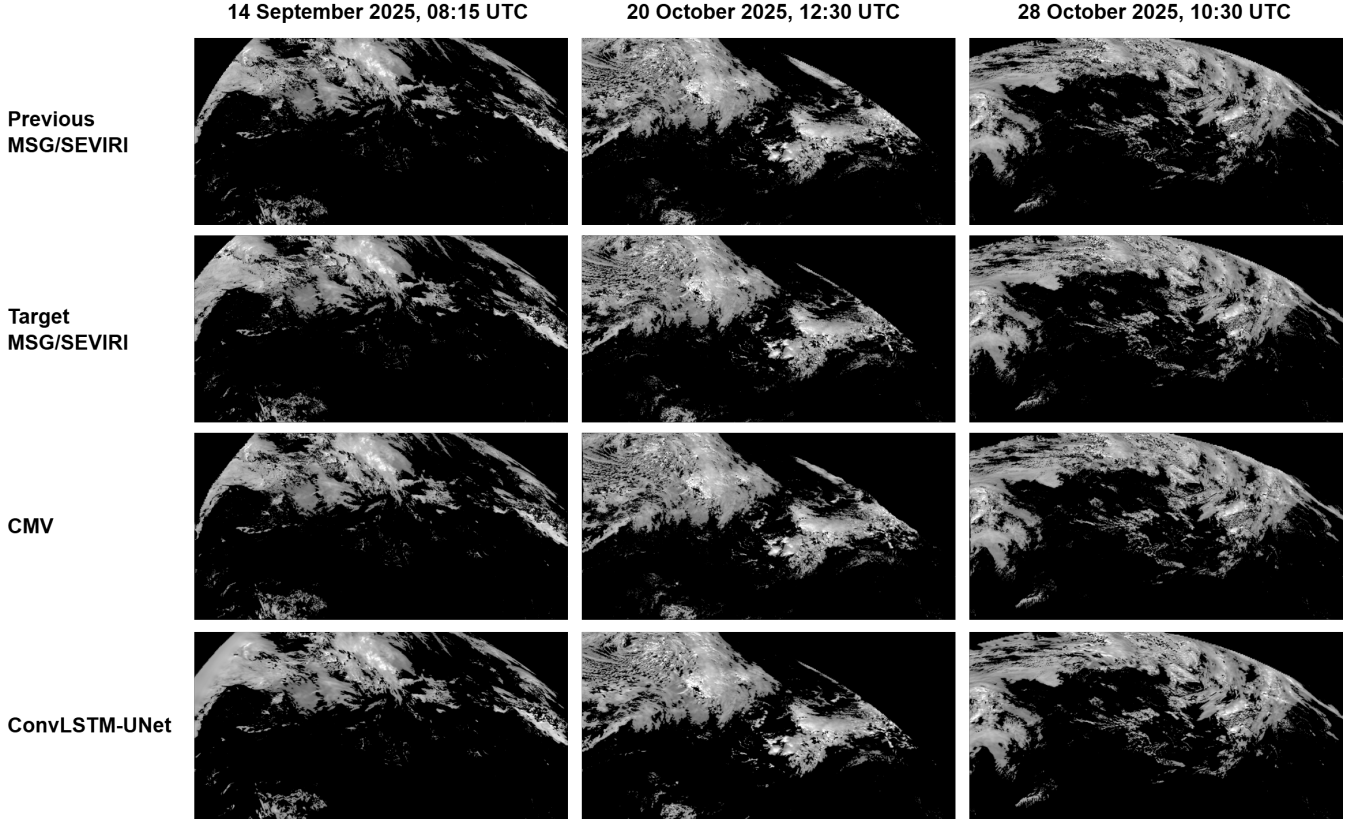


Figure 3. COT image predictions 15 min ahead, performed using CMV (third row) and ConvLSTM-UNet (bottom row) for three distinct COT fields from the test set, acquired on 14 September 2025 at 08:15 UTC (left panels), 20 October 2025 at 12:30 UTC (middle panels), and 28 October 2025 at 10:30 UTC (right panels). The target observed COT fields (second row) and those corresponding to the previous time step (top row) are also presented.

the performance of our proposed DL architecture. Nevertheless, the corresponding SSIM of 0.79 remains lower than the value obtained in the present study. A similar SSIM (0.80) was reported by Wang et al. (2022), whose model achieved a corresponding PSNR of 29.43 dB. A higher average SSIM of 0.98 and a comparable PSNR of about 29.7 dB was reported by Tan et al. (2019), who evaluated a convolutional LSTM-based model on a geostationary satellite cloud dataset.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

The ConvLSTM-UNet RMSE obtained in this study is reasonably close to the results reported by Peng et al. (2026), who evaluated five DL-based spatiotemporal forecasting models for both COT and GHI prediction. They reported overall COT RMSE values ranging from approximately 5 to 10, depending on the forecast lead time. For lead times comparable to the 15-minute interval considered in this work, however, they reported an RMSE of approximately 6, which is substantially lower than the RMSE of 9.52 obtained by our ConvLSTM-UNet model.

To illustrate some results, we randomly picked three distinct COT time frames from the test set, acquired on 14 September 2025 at 08:15 UTC, 20 October 2025 at 12:30 UTC, and 28 October 2025 at 10:30 UTC, and we present them in Fig. 3 along with the corresponding CMV and ConvLSTM-UNet predicted images. The COT fields corresponding to the previous time step (15 minutes before the predicted COT images) are also included. ConvLSTM-UNet seems to reconstruct the predicted images quite well, which is also evident by the overall high PSNR and SSIM values presented in Table 1.

The nonlinear, convolutional design of ConvLSTM-UNet enables it to capture complex spatiotemporal patterns more effectively compared to the purely advective forecasting of the CMV technique. For example, when clouds appear at the edge of the prediction domain, ConvLSTM-UNet can predict them more effectively compared to CMV. Therefore, the proposed model exhibits improved performance during day-night transition periods when clouds appear or disappear due to changes in illumination. Unlike CMV, ConvLSTM-UNet predicts clouds emerging in regions that were previously obscured by darkness. Similarly, it can successfully predict the apparent disappearance of clouds after sunset. These effects are particularly evident near the upper edge of the images (Fig. 3), where day-night transitions occur. Nevertheless, these transition regions are of limited importance for solar energy forecasting because they correspond to very low solar irradiance and therefore contribute negligibly to photovoltaic power generation. Consequently, all pixels affected by day-night transitions are excluded from the following analysis by filtering out pixels with SZA greater than 70°.

3.2 Forecasting the cloud impact on surface solar irradiance

As discussed previously, CMF is used to quantify the impact of clouds on the downwelling SSI. Figure 4 presents the observed and forecasted COT and CMF fields corresponding to the MSG/SEVIRI image acquired on 28 October 2025 at 10:30 UTC (see also Fig. 3), together with the CMF bias for both the CMV and ConvLSTM-UNet forecasts relative to the MSG/SEVIRI observations. Overall, the CMF derived by ConvLSTM-UNet exhibits smaller absolute biases that are more closely centered around zero than those obtained with the CMV approach. Both methods overestimate CMF (and underestimate COT), with larger biases being observed at higher latitudes. Nevertheless, the CMV approach performs noticeably worse at higher latitudes, mainly due to the higher cloud variations observed over those regions, and near the boundaries of the study domain, probably due to the lack of information from neighboring pixels which is important for CMV estimations.

Figure 4 also highlights the nonlinear relationship between COT and CMF. Although COT can take very high values (>200), overcast conditions are already reached at much lower COT values. For COT values of approximately 50-60, CMF approaches zero, indicating that only a small fraction of the solar irradiance reaches the surface compared to clear-sky conditions. Consequently, further increases in COT beyond this range have only a negligible effect on the downwelling solar irradiance. In contrast, relatively small changes in COT under optically thinner or scattered cloud conditions can lead to significant changes

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

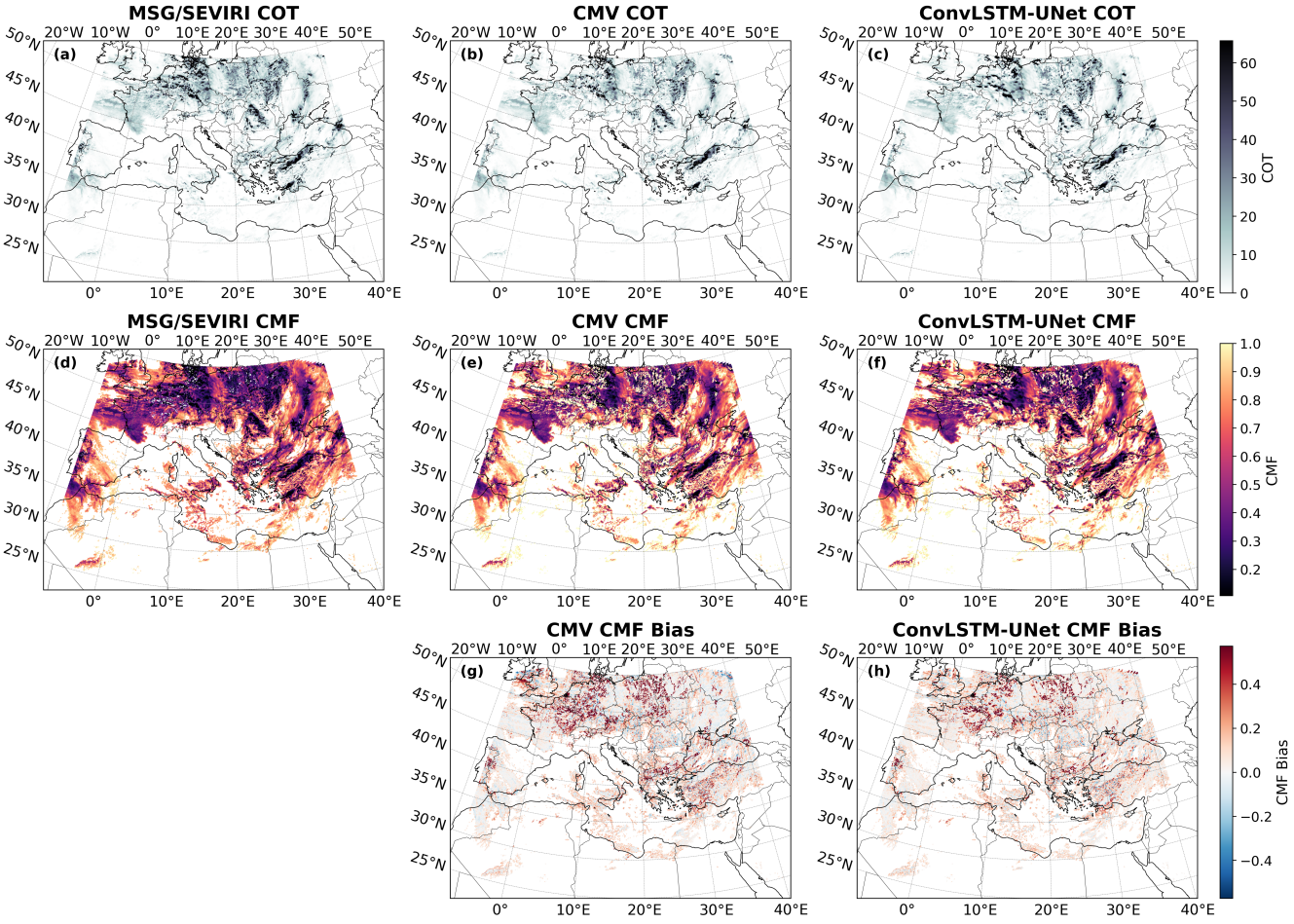


Figure 4. COT and CMF spatial fields on 28 October 2025 at 10:30 UTC **(a, d)** as acquired by MSG/SEVIRI and forecasted with **(b, e)** CMV approach and **(c, f)** ConvLSTM-UNet, together with **(g, h)** the corresponding CMF bias for both the CMV and ConvLSTM-UNet forecasts.

in SSI, e.g., a change of COT from 0 to 1 corresponds to a GHI reduction in the order of 120 W/m^2 and a change in COT from 10 to 11 leads to a change of about 20 W/m^2.

The southern part of the study domain, particularly over North Africa, is dominated by clear-sky conditions, resulting in COT = 0 (CMF = 1) for most pixels. Approximately 92.4% of the pixel-level samples within the latitude band 25°-35°N correspond to cloud-free conditions during the test period. This percentage decreases toward higher latitudes but still reaches about 72% over the entire domain. Such cloud-free cases are accurately predicted by both the CMV and ConvLSTM-UNet models, as well as the persistence approach. Consequently, including them in the evaluation may bias the performance metrics toward artificially low errors, particularly over regions with limited cloud variability.

To further investigate the spatial distribution of forecast performance and its dependence on cloud variability, Fig. 5 presents the temporally averaged spatial distributions of the CMF MAE for both CMV and ConvLSTM-UNet, together with the mean

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

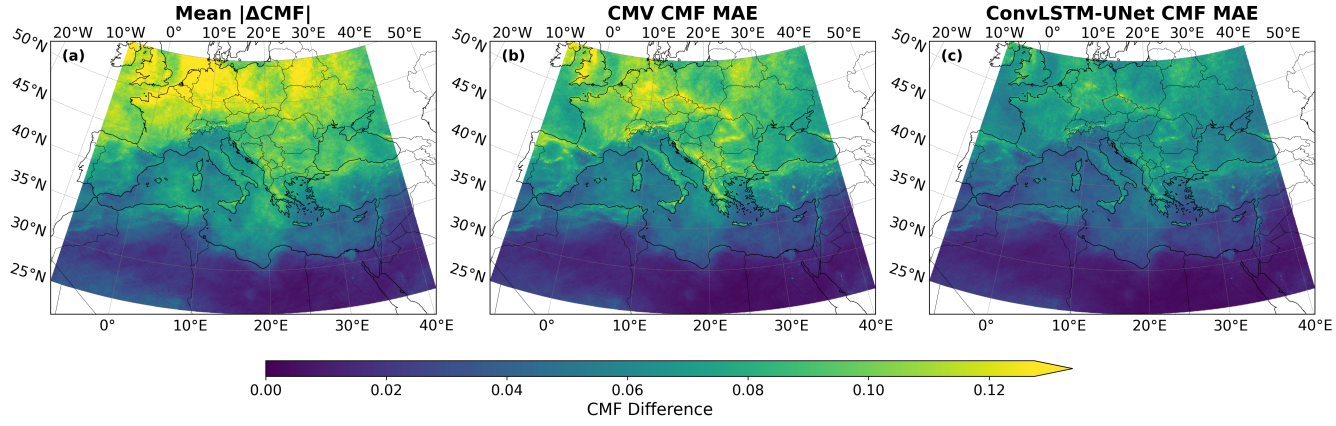


Figure 5. Temporally averaged spatial distributions of **(a)** mean |ΔCMF| and CMF MAE for both **(b)** CMV and **(c)** ConvLSTM-UNet forecasts 15 min ahead.

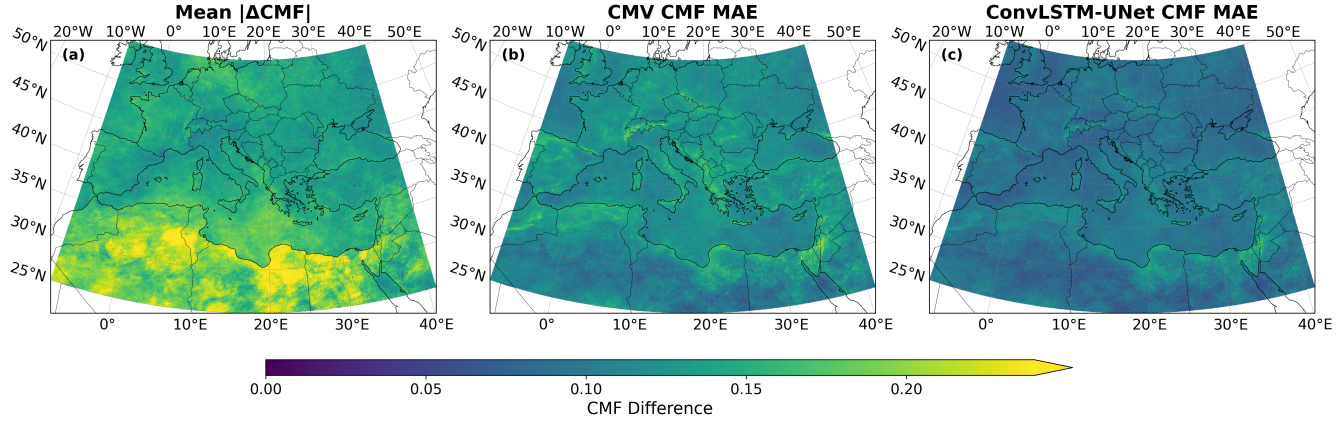


Figure 6. Same as Fig. 5 after removing persistent clear-sky pixels. The color scale is different from Fig. 5.

absolute temporal change in CMF between consecutive observations (|ΔCMF|). The latter represents the error that would be obtained using a persistence forecast and therefore provides a useful baseline against which forecasting models can be evaluated. As expected, |ΔCMF| exhibits increased values at higher latitudes, where cloud variability is enhanced compared to the quite stable and often cloudless conditions prevailing at southern regions. Similarly, Hou et al. (2022) computed mean |ΔCMF| over different regions ranging between approximately 0.05 and 0.07 with larger values corresponding to higher latitudes. Both forecasting approaches exhibit MAE values generally lower than |ΔCMF| at higher latitudes, with the lowest errors being derived by ConvLSTM-UNet. At lower latitudes (approximately below 35°N), however, the forecast MAE becomes comparable to |ΔCMF|.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

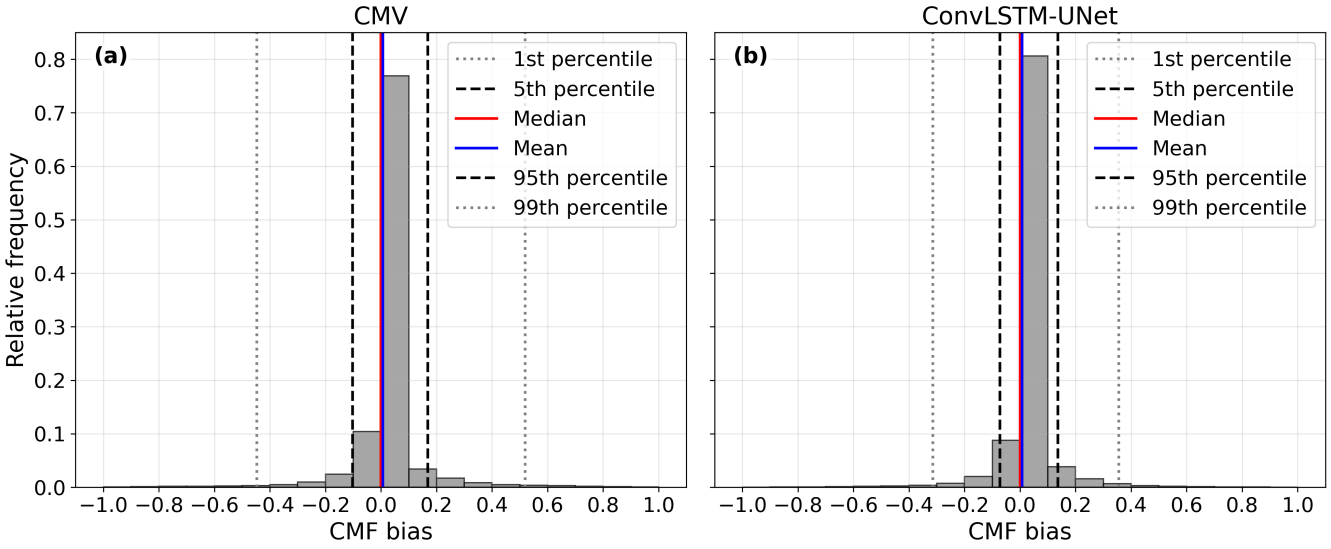


Figure 7. CMF bias distributions derived from all **(a)** CMV and **(b)** ConvLSTM-UNet forecasts 15 min ahead, corresponding to the maps of Fig. 5.

Since low-latitude regions are dominated by cloud-free conditions, Fig. 5 was recomputed after excluding persistent clear-sky pixels, defined as pixels for which both the previous observation and the target observation correspond to CMF = 1 (Fig. 6). Following this filtering, the remaining samples represent only approximately 7.6% of the data for latitudes between 25° and 35°N, and 5.2% for the 25°-30°N latitude band. It should be noted that different color scales are used in Figs. 5 and 6 for visualization purposes. After removing persistent clear-sky pixels, both |ΔCMF| and forecast MAE increase across the entire domain. Particularly, |ΔCMF| increases substantially over southern regions, indicating that although cloud occurrence is infrequent there, cloud events are associated with considerable temporal variability. Under these conditions, both forecasting approaches exhibit larger errors, which, however, are quite lower relative to |ΔCMF|. The ConvLSTM-UNet MAE remains lower compared to both the CMV approach and the persistence benchmark across most of the domain.

Table 2. Performance statistics for CMF derived from CMV and ConvLSTM-UNet COT forecasts 15 min ahead using all available data and excluding persistent clear-sky pixels in parenthesis.

	R	Slope	Intercept	MB	MAE	RMSE	FS [%]
CMV	0.85 (0.71)	0.835 (0.736)	0.152 (0.194)	0.007 (0.030)	0.044 (0.130)	0.126 (0.217)	10.37 (13.63)
ConvLSTM-UNet	0.91 (0.83)	0.887 (0.850)	0.107 (0.123)	0.007 (0.029)	0.034 (0.100)	0.099 (0.168)	29.95 (33.02)

The CMF bias distributions for the CMV and ConvLSTM-UNet forecasts, computed using the unfiltered dataset, are presented in Fig. 7. As expected, both distributions are strongly centered around zero because approximately 72% of the pixels

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

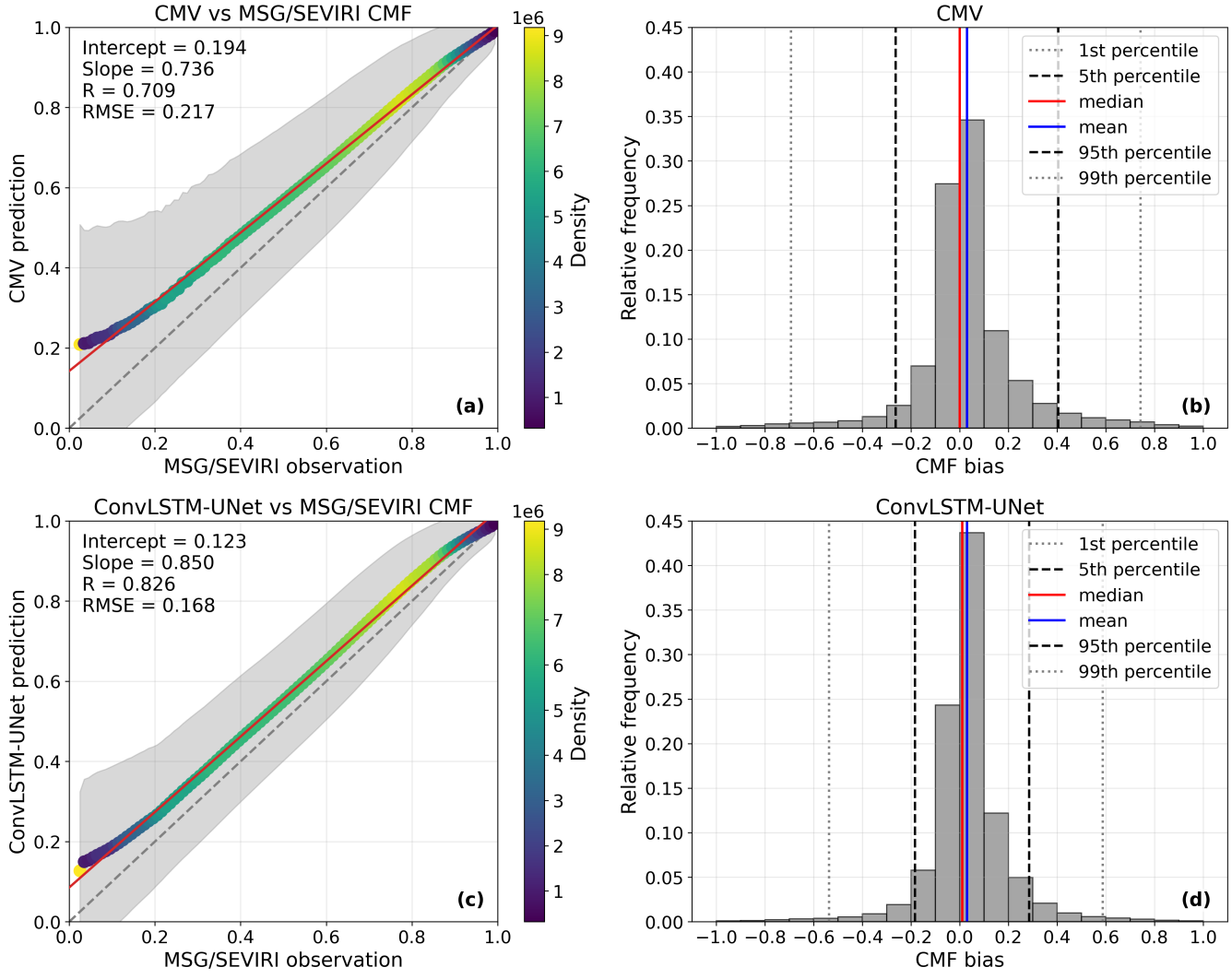


Figure 8. (a, c) Bin-averaged CMF scatter density plots and **(b, d)** CMF bias distributions derived from **(a, b)** CMV and **(c, d)** ConvLSTM-UNet forecasts 15 min ahead, excluding persistent pixels. The shaded area in the scatter plots corresponds to the standard deviation within each bin.

correspond to stable cloud-free conditions (CMF = 1). The mean CMF bias is 0.007 for both CMV and ConvLSTM-UNet models. The corresponding performance metrics are summarized in Table 2. Overall, both forecasting approaches exhibit strong agreement with the MSG/SEVIRI observations, yielding high correlation coefficients and low errors. Nevertheless, the proposed ConvLSTM-UNet consistently outperforms the CMV approach across all evaluation metrics.

435 ConvLSTM-UNet demonstrates a substantial improvement over the persistence forecast, achieving a FS of about 29.95%. The corresponding average FS for CMV is quite lower (10.37%). Papachristopoulou et al. (2024) evaluated the FS of CMV-

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

derived GHI at several BSRN stations and reported similar positive values of up to approximately 14%, depending on location and forecast lead time. The only exception was Tamanrasset, the southernmost station in their study, located in North Africa, for which a negative FS was reported (-1.3%). They further concluded that persistence outperforms the CMV approach under cloud-free conditions and regions characterized by limited cloud variability (negative FS), while CMV performed best (FS > 20%) under cloudy conditions (CMF < 0.4) and shorter forecast time horizons.

To obtain a more meaningful assessment of model performance under cloudy conditions, all pixels corresponding to persistent cloud-free conditions (CMF = 1 at both the previous and target time steps), were excluded from the subsequent analysis. The resulting performance metrics are included in parenthesis in Table 2, while the corresponding bias distributions and bin-averaged scatter density plots are illustrated in Fig. 8. Due to the very large number of data pairs, individual samples are not displayed in the scatter plots. Instead, CMF values were averaged within bins of fixed width (0.01) to facilitate visualization.

As anticipated, the performance metrics worsen once persistent clear-sky pixels are removed, because the assessment is no longer skewed toward unrealistically low errors. The CMV approach is affected more strongly by the filtering, exhibiting poorer agreement with the actual observations and considerably larger errors. In particular, the correlation coefficient drops from 0.85 to 0.71, while the regression slope and intercept change by approximately -11.9%, and 27.6%, respectively. The MB and MAE increase by a factor of three and two, respectively, whereas the RMSE increases by about 72%.

The ConvLSTM-UNet model likewise shows reduced performance after filtering; however, it still clearly outperforms the CMV method. The correlation coefficient decreases from 0.91 to 0.83, the regression slope is reduced by only about 4.2%, and the intercept increases by approximately 15%. Although the magnitude of the impact on ConvLSTM-UNet forecast errors is similar with that on CMV errors, the derived MAE and RMSE values are considerably lower, with MB being comparable. These results indicate that the proposed ConvLSTM-UNet is considerably more robust than the CMV approach under cloudy and dynamically evolving conditions.

It is worth noting that the FS is the only performance metric that improves after excluding persistent clear-sky pixels for both forecasting methods. This behavior is expected because persistence performs worse under variable cloud conditions than under cloud-free conditions, thereby increasing the relative skill of both forecasting approaches.

3.3 Dependence of forecast performance on cloud regime and latitude

According to the spatial distributions of the CMF error (Figs. 5 and 6), forecast performance exhibits a clear dependence on latitude. In addition, the scatter density plots in Fig. 8 suggest that the prediction accuracy is also influenced by cloudiness, with improved agreement under cloud-free conditions. To better understand these two factors, we further investigated the CMF bias as a function of geographical latitude and cloud conditions (Fig. 9).

To quantify the latitudinal dependence, the dataset was first divided into five latitude bands of width ±2.5° centered at 30°, 35°, 40°, 45°, and 50°N (Fig. 9a). The corresponding performance statistics are summarized in Table 3. All forecasting approaches exhibit a clear improvement with increasing latitude, particularly between 30° and 40°N, after which the performance remains nearly constant. The ConvLSTM-UNet consistently outperforms the CMV approach across all latitude bands, maintaining higher correlation coefficients and lower errors. Though ConvLSTM-UNet exhibits improved correlation and MAE

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

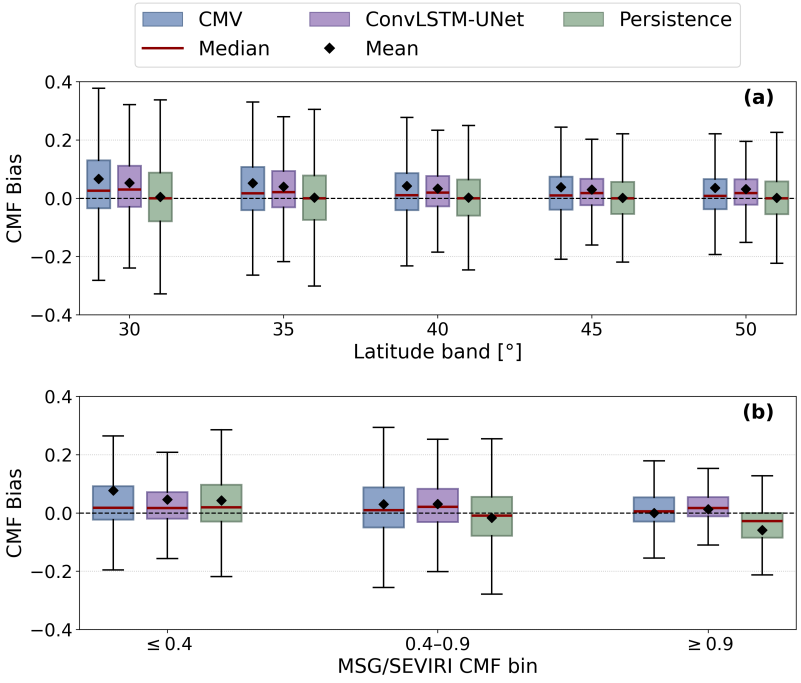


Figure 9. Box and whisker plots of CMF bias derived from CMV, ConvLSTM-UNet, and persistence models over **(a)** five latitude bands of width ±2.5° centered at latitudes 30°, 35°, 40°, 45°, and 50°N and **(b)** three different CMF bins.

compared to the persistence method, its RMSE between 40° and 50°N is slightly higher. In addition, persistence shows quite lower MB values, as persistence forecast over a long time series can sometimes balance out close to zero, due to alternating overestimations and underestimations rather than structural accuracy. In particular, persistence overestimates the SSI during clear-to-cloudy transitions, while it underestimates it during cloudy-to-clear transitions.

The apparent dependence on latitude is primarily linked to the regional cloud climatology, as discussed earlier, rather than to latitude per se. As shown in Figs. 5 and 6, southern regions are dominated by persistent clear-sky conditions, whereas the relatively infrequent cloud events are characterized by enhanced temporal variability. Consequently, cloud occurrences at lower latitudes represent more challenging forecasting situations, leading to larger errors for both forecasting approaches. In contrast, cloud systems at higher latitudes are generally more persistent, resulting in improved forecast performance.

To further investigate the influence of cloud conditions independently of latitude, the data were classified into three cloud regimes using the observed CMF derived from the MSG/SEVIRI COT retrievals: overcast (CMF $\leq$ 0.4), partially cloudy (0.4 < CMF < 0.9), and clear-sky (CMF $\geq$ 0.9) conditions (Fig. 9b). For this analysis, pixels with CMF = 1 at either the target or previous time step were excluded to prevent biasing the clear-sky category (CMF $\geq$ 0.9) toward negative values. Such bias would occur if only pixels with persistent clear sky (CMF = 1 at both times) were removed, leaving cases where the previous step has CMF < 0.9 and the current step has CMF = 1. The same filtering procedure was applied for the latitude band

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

Table 3. Statistics for CMF derived from CMV, ConvLSTM-UNet, and persistence COT forecasts over five latitude bands. The best results per metric are in bold.

	Latitude band (±2.5°N)	R	MB	MAE	RMSE
CMV	30	0.745	0.066	0.125	0.192
	35	0.773	0.052	0.115	0.180
	40	0.789	0.042	0.106	0.172
	45	0.790	0.038	0.099	0.166
	50	0.786	0.036	0.094	0.163
ConvLSTM-UNet	30	0.827	0.052	0.105	0.155
	35	0.859	0.040	0.095	0.141
	40	0.880	0.033	0.083	0.129
	45	**0.891**	0.030	0.074	0.119
	50	0.880	0.031	**0.074**	0.122
Persistence	30	0.747	0.004	0.116	0.160
	35	0.797	0.002	0.107	0.150
	40	0.861	0.002	0.090	0.127
	45	0.877	**0.001**	0.081	**0.116**
	50	0.856	0.001	0.084	0.122

Table 4. Statistics for CMF derived from CMV, ConvLSTM-UNet, and persistence COT forecasts over three different CMF bins. The best results per metric are in bold.

	CMF bin	R	MB	MAE	RMSE
CMV	≤ 0.4	0.383	0.077	0.116	0.215
	0.4–0.9	0.649	0.029	0.102	0.150
	≥ 0.9	0.236	**-0.0005**	0.053	0.075
ConvLSTM-UNet	≤ 0.4	0.552	0.046	0.084	0.150
	0.4–0.9	**0.752**	0.030	0.085	0.122
	≥ 0.9	0.253	0.013	**0.046**	**0.063**
Persistence	≤ 0.4	0.525	0.043	0.094	0.139
	0.4–0.9	0.673	-0.017	0.093	0.130
	≥ 0.9	0.203	-0.058	0.068	0.114

analysis to ensure that the results can be directly comparable; however, the latitudinal classified results were barely affected. The performance statistics for different cloud regimes are presented in Table 4.

All forecasting approaches exhibit the best accuracy under clear-sky conditions, where cloud evolution is limited and persistence becomes a strong predictor. Forecast errors increase substantially as cloudiness increases, with the largest errors occurring under overcast conditions. This behavior is particularly evident for the CMV approach, which exhibits the highest positive bias and the largest MAE and RMSE for CMF ≤ 0.4, indicating a systematic tendency to overestimate CMF (i.e., underestimate COT) for optically thick cloud fields. Although the performance of ConvLSTM-UNet also deteriorates under overcast condi-

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

tions, it generally outperforms both the CMV and persistence approaches across all cloud regimes. The largest improvement over CMV is achieved under overcast conditions, where the MAE is reduced from 0.116 to 0.084 (by approximately 27.6%), while also remaining lower than the persistence MAE of 0.094. However, within the same cloud regime, persistence forecast exhibits comparable MB and slightly lower RMSE. Under partially cloudy conditions, the ConvLSTM-UNet achieves the highest correlation overall and lower MAE and RMSE compared to the other two methods. Moreover, it achieves the lowest error metrics under clear-sky conditions indicating predictability gains even when cloud evolution is weak. Interestingly, CMV shows a very low negative bias (-0.0005) under cloudless skies.

The persistence MB exhibits a systematic dependence on the initial cloud regime (Fig 9b and Table 4), being predominantly positive for low CMF values (CMF $\leq$ 0.4) and negative for high CMF values (CMF $\geq$ 0.9), while remaining closer to zero for intermediate CMF values (0.4 < CMF < 0.9). Since the persistence forecast assumes no change in CMF, the sign of the MB reflects the mean evolution of CMF relative to its initial value. Specifically, the positive MB at low CMF values indicates a tendency for CMF to decrease over the forecast period, whereas the negative MB at high CMF values indicates a tendency for CMF to increase. These opposing regime-dependent biases partially compensate when aggregated over the full CMF range, resulting in the substantially smaller overall MB (e.g., Fig 9a and Table 3).

Overall, these results indicate that the apparent dependence of forecast performance on latitude primarily reflects differences in cloud climatology and temporal cloud variability rather than a direct geographical effect. Regions characterized by frequent cloud-free conditions and rapidly evolving cloud systems represent the most challenging forecasting environments, whereas the ConvLSTM-UNet demonstrates consistently greater robustness than both the CMV and persistence approaches under all cloud regimes and latitude bands.

3.4 Forecast performance at extended time horizons

In the following, we investigate the performance of each forecasting method over different prediction horizons. Since the ConvLSTM-UNet model was trained to predict only the next 15-minute cloud field, the results presented here correspond to recursive predictions up to 3 h ahead. Other approaches, such as training a model that makes predictions for multiple lead times (sequence-to-sequence model) or training different models for different time horizons, would reduce error propagation but are computationally more demanding. The CMV and persistence methodologies were also recursively applied. Figure 10 illustrates the evolution of the CMF forecast performance as a function of prediction lead time up to 3 h ahead for each approach. The corresponding FS of CMV and ConvLSTM-UNet relative to persistence is also included. The persistent clear-sky pixels were excluded also from the analysis over the extended time horizons.

The accuracy of all forecasting approaches progressively declines with increasing lead time due to the accumulation of forecast uncertainty and the increasingly nonlinear evolution of cloud fields. This degradation is reflected by decreasing correlation coefficients and increasing error metrics. Despite this expected deterioration, the proposed ConvLSTM-UNet consistently outperforms the CMV approach over almost the entire forecasting horizon. Although both methods exhibit increasing errors with lead time, the deterioration of the ConvLSTM-UNet is more gradual, indicating that the learned spatiotemporal representations

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

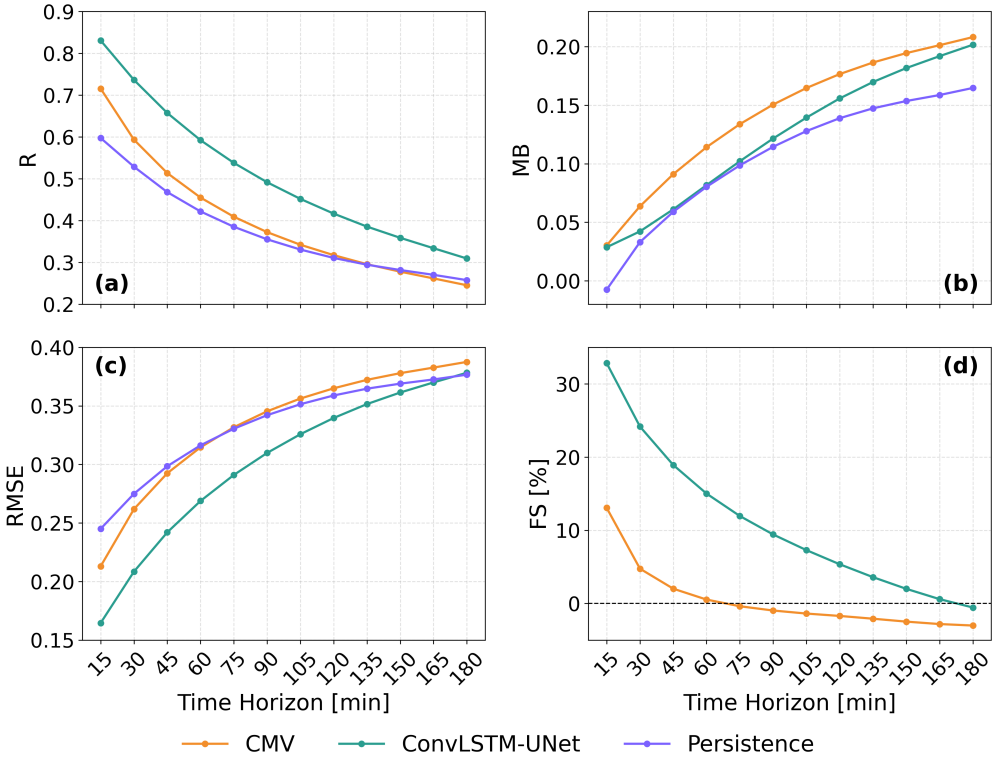


Figure 10. **(a)** Correlation coefficient, **(b)** mean bias, **(c)** root mean square error, and **(d)** forecast skill of CMF derived from CMV (orange line), ConvLSTM-UNet (green line), and persistence (purple line) forecasts at each 15 min time step up to 3 h ahead.

remain effective well beyond the 15 min training horizon. In contrast, the CMV forecasts exhibit a more rapid increase in MB, reflecting the inherent limitation of CMV in representing cloud formation, dissipation, and structural evolution.

The FS analysis (lower right panel in Fig. 10) further highlights these differences. The ConvLSTM-UNet maintains positive FS relative to persistence for lead times up to approximately 165 min, demonstrating consistent added predictive value over a persistence forecast. At the longest investigated lead time (180 min), however, the FS is slightly negative, indicating that persistence becomes more accurate. The CMV approach exhibits predominantly negative FS after the first 1 h, indicating that simple persistence provides a more reliable estimate than CMV for the investigated domain and longer forecast horizons.

Improved persistence forecast correlation compared to CMV is also evident for longer lead times (> 120 min). The consistently lower persistence biases reflect the alternating overestimations and underestimations of the persistence forecast rather than improved accuracy, as discussed earlier. The increasing positive persistence MB suggests that, over the collection of forecast cases, future CMF tends to be lower than its initial value, increasingly so as the forecast horizon increases. Persistence keeps the initial CMF unchanged, and therefore increasingly overestimates CMF at longer lead times. It is worth noting that,

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

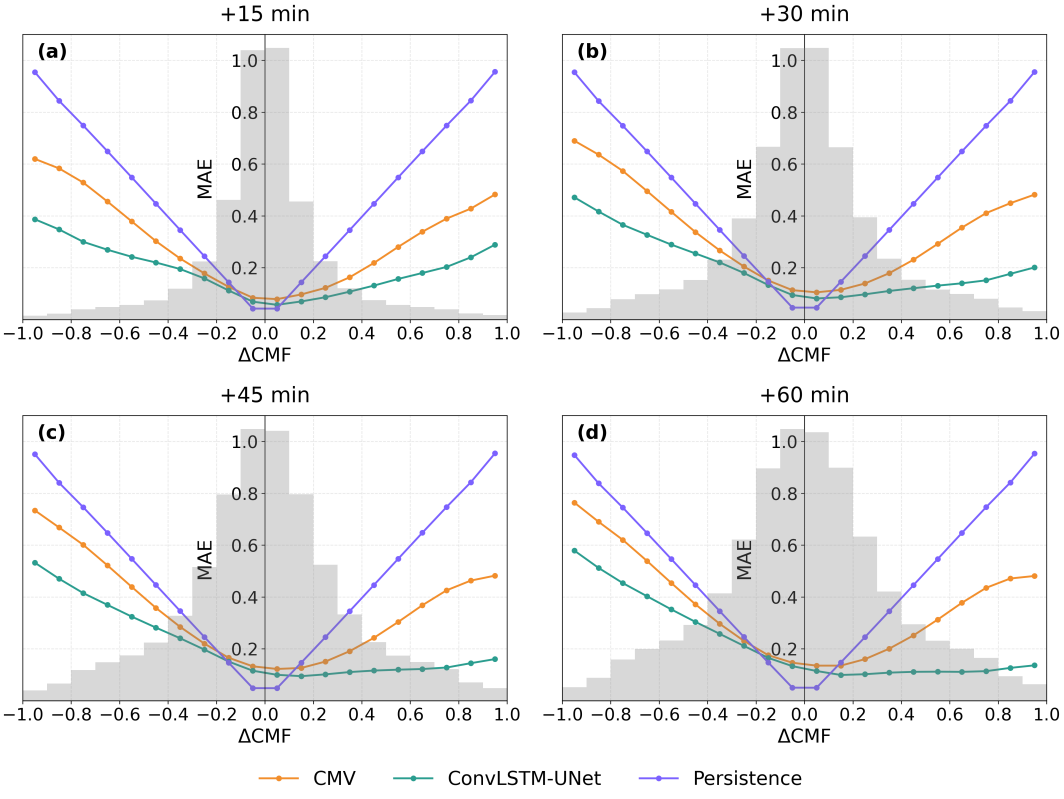


Figure 11. Distribution of the temporal change in the cloud modification factor (ΔCMF) computed from consecutive MSG/SEVIRI observations (grey bars), together with the corresponding mean absolute forecast error (MAE) of the CMV and ConvLSTM-UNet models as well as the persistence approach as a function of ΔCMF for forecast lead times of **(a)** 15 min, **(b)** 30 min, **(c)** 45 min, and **(d)** 60 min.

although the proposed ConvLSTM-UNet was trained to predict only the next 15-minute cloud field, its recursive application up to 3 h ahead demonstrates that the learned cloud evolution patterns remain informative for quite longer forecasting horizons.

To further investigate the relationship between forecast performance and cloud evolution, Fig. 11 presents the MAE as a function of the observed temporal change in the CMF (ΔCMF) for lead times of 15, 30, 45, and 60 min. The corresponding distribution of ΔCMF is also included to illustrate the frequency of occurrence of different cloud variability regimes. The analysis was conducted by dividing the ΔCMF into classes with a fixed width of 0.1.

The ΔCMF distribution is strongly clustered at small values, indicating that small temporal changes in cloud conditions occur much more frequently than rapidly evolving cloud fields. For both CMV and ConvLSTM-UNet models, the MAE increases with increasing |ΔCMF|, demonstrating that the forecast error is primarily controlled by the magnitude of cloud evolution rather than by the forecast horizon alone. Stable cloud conditions (small |ΔCMF|) are predicted with relatively high accuracy, whereas rapidly evolving cloud fields represent the most challenging forecasting situations. The CMV and ConvLSTM MAE over cases of small cloud variability become higher also as the forecast horizon increases. As expected,

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

there is a 1:1 relationship between the persistence MAE and |ΔCMF|, as the latter represents the error that would be obtained using the persistence forecast.

ConvLSTM-UNet consistently maintains lower MAE than the CMV approach across the entire ΔCMF range and for all investigated lead times. The advantage of the proposed model becomes more pronounced for moderate and large |ΔCMF| values, where cloud formation, dissipation, and structural evolution cannot be sufficiently represented by the CMV approach. In contrast, under quite stable cloud conditions (small |ΔCMF|), the performance difference between the two approaches becomes smaller because simple cloud advection dominates the cloud evolution. These findings further support the conclusion that the improved performance of the ConvLSTM-UNet primarily depends on its ability to learn complex spatiotemporal cloud dynamics beyond simple cloud motion. Consequently, the largest improvements over the CMV approach were achieved under the cloud conditions that are most difficult to predict and that contribute most to the overall forecast uncertainty.

4 Summary and conclusions

This study presents a deep learning framework based on a hybrid ConvLSTM-UNet architecture for short-term forecasting of cloud optical thickness (COT) from geostationary MSG/SEVIRI observations. The proposed model combines multiscale spatial feature extraction with recurrent spatiotemporal learning to capture both the spatial structure and temporal evolution of cloud fields. In addition to evaluating the accuracy of the predicted COT fields, the impact of cloud forecasting uncertainties on surface solar irradiance (SSI) was assessed through the cloud modification factor (CMF), providing a direct link between cloud prediction performance and solar energy applications.

The proposed ConvLSTM-UNet consistently outperformed both the optical flow cloud motion vector (CMV) approach and the persistence forecast. For 15-min-ahead COT predictions, the model achieved lower reconstruction errors (MAE $\approx$ 1.25, RMSE $\approx$ 9.52) and higher image quality (PSNR = 29.71, SSIM = 0.95), demonstrating its ability to preserve realistic cloud structures while reducing the smoothing commonly observed in deep learning forecasts. The improved performance is primarily attributed to the model's capability to learn complex spatiotemporal cloud dynamics, including cloud formation, dissipation, and structural evolution, which cannot be adequately represented by conventional motion-based extrapolation techniques.

The improved performance of the proposed framework is also reflected in the derived CMF forecasts. The ConvLSTM-UNet produced higher correlations and considerably lower errors than the CMV approach across the study domain and achieved a positive forecast skill of approximately 30% relative to persistence, while the corresponding CMV FS was about 10.4% under all-sky conditions. The advantage of the proposed model became even more pronounced after excluding persistent clear-sky conditions, thereby demonstrating its robustness under dynamically evolving cloud fields, which are of primary relevance for operational solar radiation forecasting.

The analysis further shows that forecast performance depends strongly on cloud variability and cloud regime rather than geographical location alone. Although all forecasting approaches perform better under stable clear-sky conditions, the proposed ConvLSTM-UNet generally maintains the highest accuracy across most cloud regimes and latitude bands. The largest improvements over the CMV approach were obtained under optically thick cloud fields (27.6% reduction in MAE for CMF

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

≤ 0.4) and rapidly evolving cloudiness (much lower errors for moderate and large $|\Delta CMF|$), where conventional optical flow methods are inherently limited.

Finally, the recursive application of the model demonstrated that the learned spatiotemporal representations remain informative well beyond the 15-minute training horizon. Despite the expected decline in accuracy for longer lead times, the ConvLSTM-UNet model maintained positive FS relative to persistence for forecast horizons up to about 165 minutes, highlighting its potential for operational short-term cloud and solar radiation forecasting.

Overall, the results demonstrate that deep spatiotemporal learning provides a robust and accurate alternative to traditional cloud motion forecasting techniques. The improvement of both cloud prediction accuracy and its propagation on SSI estimates highlight the proposed framework potential to enhance operational satellite-based cloud and solar radiation forecasting systems.

Future efforts will aim to extend the framework by exploiting observations from the Meteosat Third Generation (MTG) mission, whose enhanced spatial and temporal resolution is expected to further improve forecasting performance. In addition, efficient patch-based training strategies should be investigated to reduce computational requirements while preserving large-scale cloud structures. Future research is also needed to explore transformer-based spatiotemporal architectures (Dosovitskiy et al., 2020), which have recently shown considerable promise for EO forecasting tasks (e.g. Bi et al., 2023), and to extend the proposed framework to directly predict multiple future time steps rather than relying on recursive single-step forecasting. These developments are expected to improve both forecast accuracy and operational applicability for cloud and solar energy forecasting.

Data availability. The MSG/SEVIRI data are available from the EUMETSAT Data Store at https://data.eumetsat.int/search. Any other data used in this paper are available from the authors upon request.

Author contributions. Conceptualization: TD, SK, VA; data curation: TD, AN, AG; formal analysis, visualization and writing original draft: TD; funding acquisition: VA; investigation and software: TD, AN, DD; methodology: TD, AN, DD, KP, SK; resources: AG, GT, VA; supervision: SK, GT, VA; validation: TD, DD, KP; writing (review and editing): TD, AN, DD, KP, SK, AG, KK, GT. All authors gave final approval for publication.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques.

Acknowledgements. This research was funded from the European Union's HORIZON EUROPE Programme under the Grant Agreement No 101182795 - PANORAMA.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: A system for large-scale machine learning, https://arxiv.org/abs/1605.08695, 2016.

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M.: Optuna: A next-generation hyperparameter optimization framework, in: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, pp. 2623–2631, 2019.

Ballas, N., Yao, L., Pal, C., and Courville, A.: Delving Deeper into Convolutional Networks for Learning Video Representations, https://arxiv.org/abs/1511.06432, 2016.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., and Tian, Q.: Accurate medium-range global weather forecasting with 3D neural networks, Nature, 619, 533–538, https://doi.org/10.1038/s41586-023-06185-3, 2023.

Bi, L., Hu, G., Raza, M. M., Kandel, Y., Leandro, L., and Mueller, D.: A Gated Recurrent Units (GRU)-Based Model for Early Detection of Soybean Sudden Death Syndrome through Time-Series Satellite Imagery, Remote Sensing, 12, https://doi.org/10.3390/rs12213621, 2020.

Bouguet, J.-Y.: Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm, http://robots.stanford.edu/cs223b04/algo_tracking.pdf, 2000.

Bradski, G.: The OpenCV Library, Dr. Dobb's Journal of Software Tools, 2000.

Calbó, J., Pagès, D., and González, J.-A.: Empirical studies of cloud effects on UV radiation: A review, Reviews of Geophysics, 43, https://doi.org/https://doi.org/10.1029/2004RG000155, 2005.

Charbonnier, P., Blanc-Feraud, L., Aubert, G., and Barlaud, M.: Two deterministic half-quadratic regularization algorithms for computed imaging, in: Proceedings of 1st International Conference on Image Processing, vol. 2, pp. 168–172 vol.2, https://doi.org/10.1109/ICIP.1994.413553, 1994.

Cros, S., Sébastien, N., Liandrat, O., and 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, edited by Comern, A., Stein, K., Kassianov, E. I., Gonglewski, J. D., and Schfer, K., vol. 9242 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 924202, https://doi.org/10.1117/12.2066853, 2014.

Dai, K., Li, X., Ma, C., Lu, S., Ye, Y., Xian, D., Tian, L., and Qin, D.: Learning Spatial–Temporal Consistency for Satellite Image Sequence Prediction, IEEE Transactions on Geoscience and Remote Sensing, 61, 1–17, https://doi.org/10.1109/TGRS.2023.3303947, 2023.

Derrien, M. and Gléau, H. L.: MSG/SEVIRI cloud mask and type from SAFNWC, International Journal of Remote Sensing, 26, 4707–4732, https://doi.org/10.1080/01431160500166128, 2005.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N.: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, https://arxiv.org/abs/2010.11929, 2020.

Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Seyboth, K., Matschoss, P., Kadner, S., Zwickel, T., Eickemeier, P., Hansen, G., Schlömer, S., and von Stechow, C., eds.: IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation, Cambridge University Press, Cambridge, UK and New York, NY, USA, ISBN 978-1-107-60710-1, prepared by Working Group III of the Intergovernmental Panel on Climate Change, 2011.

Farnebäck, G.: Two-Frame Motion Estimation Based on Polynomial Expansion, in: Image Analysis, edited by Bigun, J. and Gustavsson, T., pp. 363–370, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-540-45103-7, 2003.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.

Fountoulakis, I., Kosmopoulos, P., Papachristopoulou, K., Raptis, I.-P., Mamouri, R.-E., Nisantzi, A., Gkikas, A., Witthuhn, J., Bley, S., Moustaka, A., Buehl, J., Seifert, P., Hadjimitsis, D. G., Kontoes, C., and Kazadzis, S.: Effects of Aerosols and Clouds on the Levels of Surface Solar Radiation and Solar Energy in Cyprus, Remote Sensing, 13, https://doi.org/10.3390/rs13122319, 2021.

Hou, X., Papachristopoulou, K., Saint-Drenan, Y.-M., and Kazadzis, S.: Solar Radiation Nowcasting Using a Markov Chain Multi-Model Approach, Energies, 15, https://doi.org/10.3390/en15092996, 2022.

IEA: Renewables 2022, https://www.iea.org/reports/renewables-2022, licence: CC BY 4.0, 2022.

Jiang, X., Chen, J., Chen, X., kin Wong, W., Wang, M., and Wang, S.: Comparative study of cloud evolution for rainfall nowcasting using AI-based deep learning algorithms, Journal of Hydrology, 639, 131 593, https://doi.org/https://doi.org/10.1016/j.jhydrol.2024.131593, 2024.

Kallio-Myers, V., Riihelä, A., Lahtinen, P., and Lindfors, A.: Global horizontal irradiance forecast for Finland based on geostationary weather satellite data, Solar Energy, 198, 68–80, https://doi.org/https://doi.org/10.1016/j.solener.2020.01.008, 2020.

Kamdem De Teyou, G., Tarabalka, Y., Manighetti, I., Almar, R., and Tripodi, S.: DEEP NEURAL NETWORKS FOR AUTOMATIC EXTRACTION OF FEATURES IN TIME SERIES OPTICAL SATELLITE IMAGES, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2020, 1529–1535, https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-1529-2020, 2020.

Kosmopoulos, P., Kouroutsidis, D., Papachristopoulou, K., Raptis, P. I., Masoom, A., Saint-Drenan, Y.-M., Blanc, P., Kontoes, C., and Kazadzis, S.: Short-Term Forecasting of Large-Scale Clouds Impact on Downwelling Surface Solar Irradiation, Energies, 13, https://doi.org/10.3390/en13246555, 2020.

Kosmopoulos, P., Dhake, H., Melita, N., Tagarakis, K., Georgakis, A., Stefas, A., Vaggelis, O., Korre, V., and Kashyap, Y.: Multi-Layer Cloud Motion Vector Forecasting for Solar Energy Applications, Applied Energy, 353, 122 144, https://doi.org/https://doi.org/10.1016/j.apenergy.2023.122144, 2024.

Lammers, R., Li, A., Nag, S., and Ravindra, V.: Prediction models for urban flood evolution for satellite remote sensing, Journal of Hydrology, 603, 127 175, https://doi.org/https://doi.org/10.1016/j.jhydrol.2021.127175, 2021.

Marchesoni-Acland, F., Herrera, A., Mozo, F., Camiruaga, I., Castro, A., and Alonso-Suárez, R.: Deep learning methods for intra-day cloudiness prediction using geostationary satellite images in a solar forecasting framework, Solar Energy, 262, 111 820, https://doi.org/https://doi.org/10.1016/j.solener.2023.111820, 2023.

Menzel, W. P.: Cloud Tracking with Satellite Imagery: From the Pioneering Work of Ted Fujita to the Present, Bulletin of the American Meteorological Society, 82, 33 – 48, https://doi.org/10.1175/1520-0477(2001)082<0033:CTWSIF>2.3.CO;2, 2001.

Miller, L., Pelletier, C., and Webb, G. I.: Deep Learning for Satellite Image Time-Series Analysis: A review, IEEE Geoscience and Remote Sensing Magazine, 12, 81–124, https://doi.org/10.1109/mgrs.2024.3393010, 2024.

Moskolaï, W. R., Abdou, W., Dipanda, A., and Kolyang: Application of Deep Learning Architectures for Satellite Image Time Series Prediction: A Review, Remote Sensing, 13, https://doi.org/10.3390/rs13234822, 2021.

Météo-France: Algorithm Theoretical Basis Document for the Cloud Product Processors of the NWC/GEO (GEO-CMA-v4.0, GEO-CT-v3.0, GEO-CTTH-v3.0, GEO-CMIC-v1.0), issue 1, Revision 1, 2016.

Natsis, A. N., Bais, A., and Meleti, C.: Analysis of cloud enhancement events in a 30-year record of global solar irradiance at Thessaloniki, Greece, Atmospheric Research, 312, 107 731, https://doi.org/https://doi.org/10.1016/j.atmosres.2024.107731, 2024.

Nguyen, C., Nachamkin, J. E., Sidoti, D., Gull, J., Bienkowski, A., Bankert, R., and Surratt, M.: Machine Learning–Based Cloud Forecast Corrections for Fusions of Numerical Weather Prediction Model and Satellite Data, Artificial Intelligence for the Earth Systems, 2, e220 072, https://doi.org/10.1175/AIES-D-22-0072.1, 2023.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.



Nonnenmacher, L. and Coimbra, C. F.: Streamline-based method for intra-day solar forecasting through remote sensing, Solar Energy, 108, 447–459, https://doi.org/https://doi.org/10.1016/j.solener.2014.07.026, 2014.

Papachristopoulou, K., Fountoulakis, I., Bais, A. F., Psiloglou, B. E., Papadimitriou, N., Raptis, I.-P., Kazantzidis, A., Kontoes, C., Hatzaki, M., and Kazadzis, S.: Effects of clouds and aerosols on downwelling surface solar irradiance nowcasting and short-term forecasting, Atmospheric Measurement Techniques, 17, 1851–1877, https://doi.org/10.5194/amt-17-1851-2024, 2024.

Park, Y.-J., Kim, D., Seo, M., Jeon, H.-G., and Choi, Y.: Data-driven Precipitation Nowcasting Using Satellite Imagery, https://arxiv.org/abs/2412.11480, 2024.

Partio, M., Hieta, L., and Kokkonen, A.: CloudCast—Total Cloud Cover Nowcasting with Machine Learning, Artificial Intelligence for the Earth Systems, 4, e240 104, https://doi.org/10.1175/AIES-D-24-0104.1, 2025.

Pelland, S., Remund, J., Kleissl, J., Oozeki, T., and De Brabandere, K.: Photovoltaic and Solar Forecasting: State of the Art, https://iea-pvps.org/wp-content/uploads/2013/10/Photovoltaic_and_Solar_Forecasting_State_of_the_Art_REPORT_PVPS__T14_01_2013.pdf, 2013.

Peng, W., He, T., Ma, Y., and Zheng, Y.: A physics-informed machine learning method for short-term solar radiation forecast with satellite-derived cloud optical depth, Applied Energy, 406, 127 284, https://doi.org/https://doi.org/10.1016/j.apenergy.2025.127284, 2026.

Perera, M., De Hoog, J., Bandara, K., Weeratunge, H., and Halgamuge, S.: Distributed solar generation forecasting using attention-based deep neural networks for cloud movement prediction, Energy, 361, 141 985, https://doi.org/https://doi.org/10.1016/j.energy.2026.141985, 2026.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., and Hoff, T. E.: Validation of short and medium term operational solar radiation forecasts in the US, Solar Energy, 84, 2161–2172, https://doi.org/https://doi.org/10.1016/j.solener.2010.08.014, 2010.

Pfeifroth, U., Drücke, J., Kothe, S., Trentmann, J., Schröder, M., and Hollmann, R.: SARAH-3 – satellite-based climate data records of surface solar radiation, Earth System Science Data, 16, 5243–5265, https://doi.org/10.5194/essd-16-5243-2024, 2024.

Qiao, M., He, X., Cheng, X., Li, P., Luo, H., Zhang, L., and Tian, Z.: Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks, International Journal of Applied Earth Observation and Geoinformation, 102, 102 436, https://doi.org/https://doi.org/10.1016/j.jag.2021.102436, 2021.

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., and Wald, L.: Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat?4 method, Meteorologische Zeitschrift, 26, 33–57, https://doi.org/10.1127/metz/2016/0781, 2017.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., and Mohamed, S.: Skilful precipitation nowcasting using deep generative models of radar, Nature, 597, 672–677, https://doi.org/10.1038/s41586-021-03854-z, 2021.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, https://arxiv.org/abs/1505.04597, 2015.

Saleem, H., Salim, F., and Purcell, C.: STC-ViT: Spatio Temporal Continuous Vision Transformer for Medium-range Global Weather Forecasting, https://arxiv.org/abs/2402.17966, 2025.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., kin Wong, W., and chun Woo, W.: Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, https://arxiv.org/abs/1506.04214, 2015.

Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D.-Y., kin Wong, W., and chun Woo, W.: Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model, https://arxiv.org/abs/1706.03458, 2017.

https://doi.org/10.5194/egusphere-2026-5862
Preprint. Discussion started: 15 October 2026
© Author(s) 2026. CC BY 4.0 License.



Sirch, T., Bugliaro, L., Zinner, T., Möhrlein, M., and Vazquez-Navarro, M.: Cloud and DNI nowcasting with MSG/SEVIRI for the optimized operation of concentrating solar power plants, Atmospheric Measurement Techniques, 10, 409–429, https://doi.org/10.5194/amt-10-409-2017, 2017.

Su, X., Li, T., An, C., and Wang, G.: Prediction of Short-Time Cloud Motion Using a Deep-Learning Model, Atmosphere, 11, https://doi.org/10.3390/atmos11111151, 2020.

Tahir, Z., Haseeb, M., Mahmood, S. A., Batool, S., Abdullah-Al-Wadud, M., Ullah, S., and Tariq, A.: Predicting land use and land cover changes for sustainable land management using CA-Markov modelling and GIS techniques, Scientific Reports, 15, 3271, https://doi.org/10.1038/s41598-025-87796-w, 2025.

Tan, C., Feng, X., Long, J., and Geng, L.: FORECAST-CLSTM: A New Convolutional LSTM Network for Cloudage Nowcasting, https://arxiv.org/abs/1905.07700, 2019.

Taylor, M., Kosmopoulos, P., Kazadzis, S., Keramitsoglou, I., and Kiranoudis, C.: Neural network radiative transfer solvers for the generation of high resolution solar irradiance spectra parameterized by cloud and aerosol parameters, Journal of Quantitative Spectroscopy and Radiative Transfer, 168, 176–192, https://doi.org/https://doi.org/10.1016/j.jqsrt.2015.08.018, 2016.

Wang, R., Teng, D., Yu, W., Zhang, X., and Zhu, J.: Improvement and Application of a GAN Model for Time Series Image Prediction—A Case Study of Time Series Satellite Cloud Images, Remote Sensing, 14, https://doi.org/10.3390/rs14215518, 2022.

Weyn, J. A., Durran, D. R., and Caruana, R.: Improving Data-Driven Global Weather Prediction Using Deep Convolutional Neural Networks on a Cubed Sphere, Journal of Advances in Modeling Earth Systems, 12, e2020MS002 109, https://doi.org/https://doi.org/10.1029/2020MS002109, e2020MS002109 10.1029/2020MS002109, 2020.

Xia, P., Zhang, L., Min, M., Li, J., Wang, Y., Yu, Y., and Jia, S.: Accurate nowcasting of cloud cover at solar photovoltaic plants using geostationary satellite images, Nature Communications, 15, 510, https://doi.org/10.1038/s41467-023-44666-1, 2024.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., and Chen, Z.: Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach, Remote Sensing of Environment, 233, 111 358, https://doi.org/https://doi.org/10.1016/j.rse.2019.111358, 2019.

Xu, Z., Du, J., Wang, J., Jiang, C., and Ren, Y.: Satellite Image Prediction Relying on GAN and LSTM Neural Networks, ICC 2019 - 2019 IEEE International Conference on Communications (ICC), pp. 1–6, https://api.semanticscholar.org/CorpusID:198169294, 2019.

Zhang, J., Gong, W., Ye, L., Chen, M., Liu, G., and Cheng, Y.: A novel window-based hybrid attention network for low-dose CT image denoising, Complex & Intelligent Systems, 12, 9, https://doi.org/10.1007/s40747-025-02116-z, 2025.

Zubair, B., Hu, W., and Peng, J.: Structural-Perceptual Image Super Resolution Using Charbonnier-SSIM Loss in an Efficient Sub-Pixel Convolutional Network, in: 2025 Photonics & Electromagnetics Research Symposium - Fall (PIERS-Fall), pp. 1–8, https://doi.org/10.23919/PIERS-Fall62445.2025.11394197, 2025.