A comparison of satellite cloud motion vectors techniques to forecast intra-day hourly solar global horizontal irradiation

D. Aicardi[a], P. Musé[b], R. Alonso-Suárez[a,*]

[a] Laboratorio de Energía Solar, Universidad de la República. Av. L. Batlle Berres, km 508, CP 50000, Salto, Uruguay
[b] Instituto de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de la República. J. Herrera y Reissig 565, CP 11300, Montevideo, Uruguay

Abstract

Solar forecasting provides valuable information for grid management. Satellite-based forecasting tools account for the short-term intra-day time horizons, typically outperforming numerical weather predictions up to 4-5 hours ahead. The method consist of three separated stages, namely, cloud motion estimation, motion extrapolation and satellite-to-irradiation conversion. In this work we compare different satellite-based proposals for hourly irradiation forecast up to 5 hours ahead using a 2-years data set. The widely-used Lorenz's block matching technique and four optical flow (OF) algorithms are assessed, both at image and irradiation levels. All the methods are locally optimized to obtain their peak performance. It is found that the OF algorithm which combines an L^1 data penalty term on the optical flow equation with total variation regularization (TVL1) outperforms the rest. Different image extrapolation approaches and spatial smoothing are also tested. It is found that changing the extrapolation technique does not have much impact in the overall performance and that important gains can be obtained by optimally smoothing the predicted images previous to solar irradiation conversion. By doing this, all methods outperform the exigent convex persistence benchmark, achieving positive forecasting skills. The tests are performed using GOES-East satellite images of south-east South America, and the methods' optimal parameters are given.

Keywords: Solar irradiation forecast, CMV, GHI, GOES satellite, optical flow.

1. Introduction

Solar forecasting is a key requirement to optimally manage solar power resources and, therefore, to increase the solar energy share in the electricity mix. The predictions provide an informed decision-making framework for an efficient energy dispatch, reducing costs associated with solar energy variability, supply/demand balance and back-up generation. The intra-day hourly forecast horizons are covered by Numerical Weather Predictions (NWP) and satellite-based predictions, being the latter a lower uncertainty option for the first time steps, typically up to 4-5 hours ahead (Kühnert et al., 2013; Perez and Hoff, 2013). The

*Corresponding author: R. Alonso-Suarez, r.alonso.suarez@gmail.com

predictions in these time horizons are especially useful for load following applications and electricity market transactions (Diagne et al., 2013; Antonanzas et al., 2016).

Satellite-based solar forecasting uses visible channel geostationary satellite images due to two main reasons. The first one is that cloudiness is easily distinguishable in these images, as clouds reflect more solar irradiance to the outer space than the ground, thus appearing brighter in the visible channel. This is not valid for areas with high albedo terrain, like snow areas or salt flats, where a special treatment is required, typically by also using infrared images (Perez et al., 2013; Dise et al., 2013). The second one is that the images' high time rate (between 30 and 10 minutes) and moderate spatial resolution (between 500 m and 1 km) allow to track the clouds' movement in time by means of estimating their Cloud Motion Field (CMF) or Cloud Motion Vectors (CMV). The CMV is then used to extrapolate the clouds' motion into the near future (from 1 to 5-6 hours ahead), predicting the next images, i.e. the future position of clouds. Finally, to obtain the ground level solar irradiation prediction, a satellite-based assessment model (Perez et al., 2002; Rigollier et al., 2004; Alonso-Suárez et al., 2012; Qu et al., 2017; Laguarda et al., 2020) is applied to each predicted image in each time horizon. Satellite forecasting is a challenging task as clouds not only move, but also form, change their shape and vanish, due to complex atmosphere dynamics that the method does not attempt to model. Also, as the CMV is a two-dimensional (x, y) vector field, the z axis is neglected, implicitly assuming that, locally in space, the cloudiness is at the same altitude plane. This makes the baseline techniques incapable of capturing any phenomenon involving vertical motions of clouds, in particular, the complex convection processes. Further assumptions may be added by the CMV estimation technique to deal with the motion estimation in the images' sequence. Despite these simplifications, satellite-based forecasting techniques have proved to be a solution for the 1-5 hours ahead time horizons, outperforming NWP and reference persistence procedures, and are usually included in solar forecasting products of specialized companies (Perez and Hoff, 2013).

The key part of solar satellite forecasting is the estimation of the CMV. The main technique used for this is the Lorenz et al. (2004) method, which is a block matching algorithm similar to the Particle Image Velocimetry (PIV) technique (Adrian, 1991), widely-used to estimate motion in fluids dynamics. This method uses two consecutive images and estimates the cloud displacement at an image pixel by comparing its surrounding rectangular area (in the current image) with the neighbouring rectangular areas in the previous image, inside a bigger search area. The area that presents the highest similarity with respect to the original one is selected, and the motion vector corresponding to the displacement is assigned to the pixel. A similarity metric for this can be the Root Mean Squared Deviation (RMSD), among others. The procedure is repeated for a set of pixels in the image to avoid the high computational cost of this search. This method (and variations) have been evaluated in several parts of the world accounting for different climates (Perez et al., 2010; Kühnert et al., 2013; Cros et al., 2020; Giacosa and Alonso-Suárez, 2020; Yang et al., 2020b; Pereira et al., 2020). Other CMV estimation techniques have been analyzed (Peng et al., 2013; Cros et al., 2014;

Nonnenmacher and Coimbra, 2014; Urbich et al., 2019; Kallio-Myers et al., 2020), including phase correlation and Optical Flow (OF) algorithms. Among the OF variations, the Lucas and Kanade (1981), Horn and Schunck (1981) and Farnebäck (2003) methods have been used and, recently, also the TVL1 algorithm (Zach et al., 2007; Urbich et al., 2019). Some of these methods directly obtain a dense motion estimation (i.e. at each image pixel) while others rely on a spatial neighbourhood around each pixel. This second category typically obtains a sparse motion estimation to avoid high computational costs, and requires an a posteriori interpolation of the motion field, in the same way as the block-matching algorithm. Comparisons of these algorithm's performance to forecast satellite albedo (or cloud index) and solar irradiation are rarely found in the literature, especially those with extended data sets. For instance, Cros et al. (2014) found that the Lucas and Kanade OF method outperforms the block-matching and phase correlation algorithms, tested with a 6-days data set of cloud index MSG images of western France and northern Spain. There also exist hybrid methods that blend either the satellite images or the CMV with NWP. This may be done by advecting the satellite images using NWP procedures or NWP wind fields information (Arbizu-Barrena et al., 2017; Miller et al., 2018; Wang et al., 2019), by combining the CMV with the NWP wind fields previous to advect (Harty et al., 2019) or by merging the solar prediction from both techniques (Perez et al., 2014), among others strategies. Either as part of a hybrid system or as a standalone methodology, satellite CMV estimation has become an integral part of intra-day cloudiness or solar irradiation forecast.

In this work we provide a detailed performance comparison between the block-matching method and four OF techniques, namely, Lucas and Kanade (1981), Farnebäck (2003), Horn and Schunck (1981) and TVL1 algorithms, to forecast satellite reflectance (satellite albedo) and hourly solar global irradiation at a ground-level horizontal surface (GHI). These methods are selected by their relevance, either for historical, practical or scientific reasons. There are no previous works comparing the cloudiness and solar forecasting performance using such many satellite-based techniques. It represents a two level comparison that addresses cloudiness at image level (using all image pixels except for a small frame around it, to avoid artifacts) and at solar irradiation level at specific measuring sites. The evaluation is done against a 1-year data set of 30-minutes GOES-East satellite images and controlled-quality GHI measurements, accounting for a large time-span in comparison to previous works. This allows for a full characterization of these methods performance and a fair comparison between them. The region under study is the south-east of South America, which includes Uruguay, most of the Argentinian territory, southern Brazil and Paraguay, and is characterized by challenging mesoscale convective systems (Salio et al., 2007; Rasmussen et al., 2014; Pal et al., 2021) and intermediate solar resource short-term variability (Alonso-Suárez et al., 2020). The findings of this article shall also apply to similar climates, especially intermediate solar variability areas where partly cloudy conditions are frequent. In addition, the five methods are optimized for the region by using an independent 1-year satellite data set. This article also provides the optimal parameters for their utilization in this region, which may be extrapolated, as said before, to other regions with similar climate conditions.

We also inspect different mechanisms to perform the image extrapolation and the role of spatial smoothing in the predicted images to better account for hourly irradiation forecast. Both procedures have been rarely addressed in the literature. In particular, image extrapolation to large forecast horizons (i.e. 10 time steps in the future) is not a common computer vision problem and it has several limitations when using a static CMV, as is the case of state-of-art satellite cloudiness and solar irradiation forecasting systems.

This article is organized as follows. Section 2 describes the satellite and solar irradiation data sets. Section 3 presents the methods being used for the CMV estimation, the characteristics of these methods' training and the solar irradiation model. Section 4 briefly presents the forecast performance metrics to be used in the evaluation. Results are presented in Section 5, including the methods' optimal parameters, the performance assessment and comparison, and the findings in relation to image extrapolation and spatial smoothing. Finally, Section 6 summarizes the main conclusions of this work.

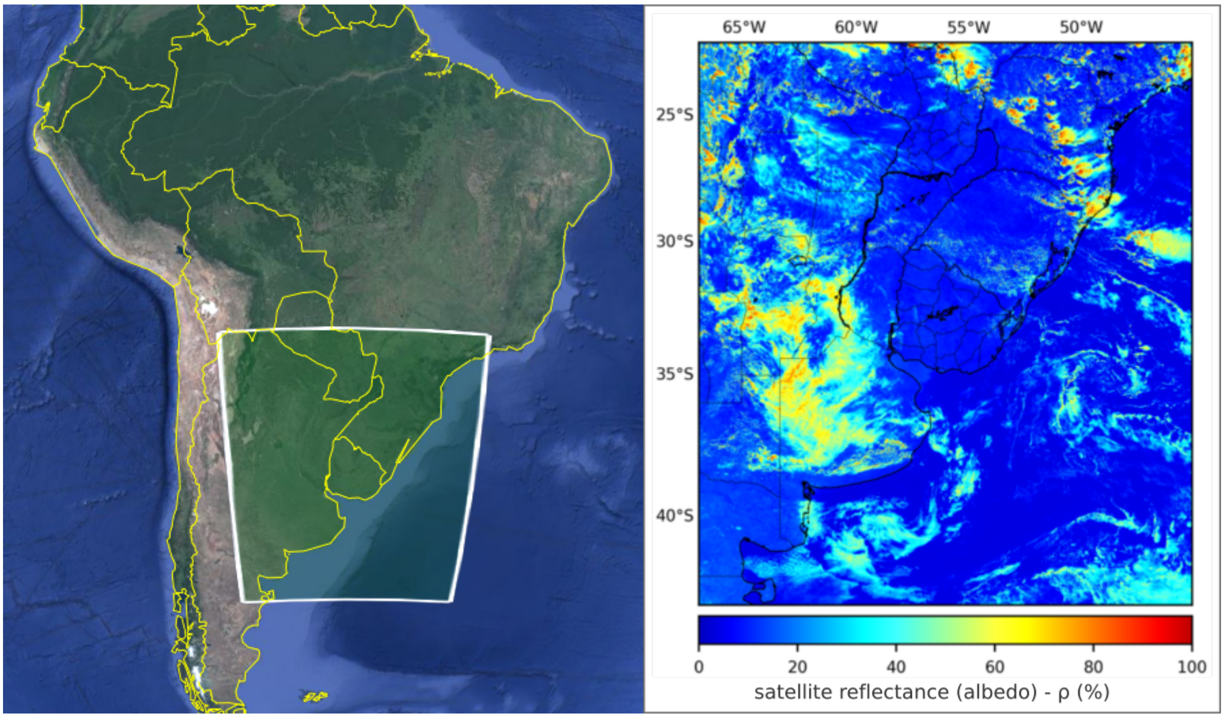
2. Data

This work is based on two data sets: GOES-East geostationary satellite images and GHI ground measurements. A 2-years period of GOES-13 satellite images are used, which correspond to years 2016 and 2017. The data from year 2017 is reserved for algorithm training while the year 2016 is used for evaluation. In the same line, hourly GHI measurements of 2016 are considered, as ground measurements are used here for evaluation purposes (i.e. none of the methods' parameters training rely on ground data, but only on the satellite images). The following two subsections describe the satellite images and GHI measurements.

2.1. Satellite images

Figure 1 shows the spatial domain of the GOES-East visible channel satellite images being used. The region includes most parts of central and north Argentina, Uruguay and south Paraguay and Brazil (south-east of South America). The majority of this area is classified as Cfa (warm, temperate and humid, with hot summers) in the updated Köppen-Geiger climate classification (Peel et al., 2007), region known as Pampa Húmeda. As mentioned in the introduction, the region is known for frequent mesoscale convective systems which, being the region subtropical, are larger in size and lifespan than the tropical ones. These convective systems tend to peak during daytime over Uruguay and south Brazil (Salio et al., 2007), hence affecting ground level solar irradiation and producing complex daylight cloud patterns. Figure 1 also illustrates a typical GOES-East albedo image (satellite Earth's reflectance, ρ) in which clouds appear brighter that the background (soil, ocean, rivers, lakes, etc.) as they reflect more Sun's radiation to outer space.

For this work we use the GOES-13 albedo images with the calibration procedures recommended by NOAA (Wu and Sun, 2005). During the period 2016-2017 this satellite acquired images with a typical rate of two per hour and a nominal space resolution of 1 km. The satellite's location in geostationary orbit, 75°W, results in a variable pixel size of about $\simeq$ 1-2 km over the region. The original image in irregular satellite projection is converted to a $0.015° \times 0.015°$ latitude-longitude regular grid by a simple pixel averaging procedure. This is then the effective spatial resolution of the images being used, which have a final resolution of 1397×1467 px. Only sector images separated by $\Delta t = 30$ minutes are considered, discarding for instance full-disk images which have a $\Delta t = 7$ minutes with their previous sector image. By doing this, the temporal support of the satellite images is fixed in equally spaced 30 minutes time-steps, with some daylight gaps that result from the operational regime of the satellite for South America. These gaps are not considered for algorithms' training or evaluation. Also, only images in which all pixels have at least a solar altitude of $\simeq 7°$ are considered (cosine of the solar zenith angle higher than 0.1) and some images are discarded due to missing pixels (image acquisition/transmission problems). This procedure results in a daylight data set of 7903 images for the year 2016 and 7794 images for the year 2017, which are going to be used for algorithms' testing and training respectively.



(a) Location of the satellite's footprint.

(b) Example of GOES-East visible channel image.

Figure 1: Location and example of the GOES-East satellite images being used, illustrating the spatial domain of this work.

2.2. GHI ground data

Six GHI measuring sites located at the center of the images' footprint are used. These sites correspond to the Solar Energy Laboratory (LES, http://les.edu.uy/) measurement's network in Uruguay; they are located in rural or semi rural areas and are representative of the broader Pampa Húmeda region. The solar irradiance 10-minutes variability in these sites is of $\sigma \simeq 0.15$ (Marchesoni-Acland and Alonso-Suárez, 2020), a dimensionless variability metric calculated as the standard deviation of the 10 minutes clear-sky index (k_c) changes (Perez et al., 2016) using the McClear clear-sky model for the k_c calculation (Lefèvre et al., 2013). This corresponds to an intermediate short-term solar variability, typical of climates with frequent partly cloudy conditions.

The measuring sites are presented in Table 1. One of these sites is the LES experimental facility (station code LE) in north-western Uruguay. The GHI measurement in this site is acquired by a Kipp & Zonen SOLYS2 ground measurement station with spectrally flat Class A pyranometers according to the ISO 9060:2018 standard. As this site is located in a specialized solar assessment lab with dedicated technical support, routine maintenance of the station is performed on roughly a daily basis, including dome cleaning and horizontal plane check. The other five sites correspond to a field measurement network administrated by LES, where spectrally flat Class A or B pyranometers are used. These stations are located in measurement fields of the National Agronomic Research Institute (INIA, Uruguay) or the National Weather Institute (INUMET, Uruguay), and maintenance is done approximately on a monthly basis by the local operators. All measurements are registered with a 1 minute time rate as an average of 15 seconds samples. The pyranometers are calibrated every two years by the LES following the ISO-9847:1992 standard (Abal et al., 2018) using as reference a Secondary Standard Kipp & Zonen CMP22 pyranometer, which is kept with traceability to the World Radiometric Reference (WRR). Based on the equipments' quality, maintenance routine and calibration, we assign a P95 uncertainty of 3% of the average for the GHI daily measurements at the LE site and of 5% in the others. These uncertainties are significantly lower than the uncertainty of the forecast being evaluated.

The GHI 1-minute data are quality inspected by eliminating the samples tagged as erroneous (i.e. malfunction periods or maintenance days) and by using the BSRN limits for atypical and physically impossible values (McArthur, 2005). After this procedure, the data is hourly integrated using the satellite timestamps as temporal support. That is, for each satellite timestamp, the hourly values for 1 to 5 hours ahead are computed. The choice of the hourly forecast horizons is based on the fact that we have a detailed uncertainty assessment of the hourly solar satellite estimates in the region that use images from the previous GOES-East generation (Laguarda et al., 2020; Alonso-Suárez, 2017), providing a reference uncertainty level for this work and also adjusted tools for satellite-to-irradiation conversion at an hourly time basis. We shall recall here that during 2016-2017 period the GOES-13 provided 30-minutes images for South America with timestamps typically at 8 and 38 minutes, with an irregular acquisition. Only data with solar altitude greater than 7° are considered. Table 1 shows the amount of hourly samples used for evaluation at each site (for 1-hour ahead, as example) and their average, which will be used for performance metrics' normalization.

Table 1: Location of the GHI measuring sites. The last two columns show, respectively, the amount of quality-inspected hourly samples that compose the final data set and the measurements average in each site (for performance metrics normalization).

station name	station code	latitude (deg)	longitude (deg)	altitude (m)	samples (hours)	$\overline{G_h}$ (Wh/m^2)
LES Experimental Facility	LE	−31.28	−57.92	60	3065	411.7
LES Rocha	RC	−34.49	−54.32	30	3324	394.6
INIA Las Brujas	LB	−34.67	−56.34	37	3320	398.2
INIA La Estazuela	ZU	−34.34	−57.69	70	3349	405.6
INIA Tacuarembó	TA	−31.71	−55.83	142	3355	402.7
INUMET Artigas	AR	−30.40	−56.51	121	3378	423.3

3. Methods

In this section we describe the five methods considered for CMV estimation, their local optimization strategy and the satellite-to-irradiation model to convert the predicted images to hourly solar irradiation predictions at the given measuring sites.

3.1. CMV estimation

The CMV is defined by a (u, v) vector field that represents the motion of clouds at each pixel (x, y) in a sequence of two consecutive images. The objective of this estimation is to find the two scalar fields $u(x, y)$ and $v(x, y)$ that describe each component of the vector field across the x and y directions, respectively. Clouds are the only Earth's atmosphere moving object in geostationary satellite images, so motion estimation algorithms can be applied directly to these images to derive the CMV. Due to the anisotropic reflection of solar radiation in the Earth-Atmosphere system, the Sun's apparent movement is also observed in the albedo images. This second order but noticeable spatial change in the albedo may introduce artifacts in the motion estimation, and is considered here as part of the methods' uncertainty. Block-matching and OF techniques are the main two types of CMV estimation methods reported in literature. To estimate objects' motion, both approaches assume that their brightness in the observed frames is preserved. Typically, given a rectangular window or block in an sequence frame, the block-matching procedure consists in identifying its most similar block in a neighbouring position in the next frame. Then, assuming that these two blocks correspond to the same object acquired at consecutive frames, the projected object's motion is computed as the displacement vector between both blocks. Different block similarity metrics can be considered, the most popular being the L^2 or L^1 norm between its pixel values.

In computer vision, motion field estimation methods based on OF first appeared in the early 80s and it is still a fundamental problem. In OF methods, the objects' brightness preservation in time is considered at the pixel level, and the time between consecutive frames is assumed to be short enough to consider that the objects perform infinitesimal motions. More precisely, if $I(x, y, t)$ represents an image sequence, and (dx, dy) and dt represent differentials in space and time respectively, the brightness constancy assumption $I(x, y, t) = I(x + dx, y + dy, t + dt)$ leads to the so called optical flow equation (Horn and Schunck, 1981),

$$I_x \cdot u + I_y \cdot v + I_t = 0, \tag{1}$$

where I_x and I_y denote the spatial derivatives of $I(x, y, t)$, and I_t its time derivative. The use of this equation for each image pixel is an underdetermined linear problem, so the different approaches to solve Eq. (1) constraint the motion field to be smooth or regular in some sense. These constrains can be local in space (Lucas and Kanade, 1981; Farnebäck, 2003) or impose regularity by penalizing non-smoothness via a cost function across the image (Horn and Schunck, 1981; Zach et al., 2007). These second kind of approaches are formulated as variational problems and are solved using a discretized version of the Euler-Lagrange equation. The choice of one or another formulation is problem dependent, as the added constrains and/or the selected penalization functions should be in accordance to the motion scene characteristics and the error probability distribution that result from the residuals of Eq. (1) application (Sun et al., 2008), called data error. The requirement of dense or sparse motion estimation, or the uncertainty introduced by interpolation techniques to convert the latter into the former, also affects the choice.

The classical Lucas and Kanade and Horn and Schunck formulations impose the OF constraint using the L^2 norm as data penalty term. This choice leads to a convex and differentiable optimization problem, but make these methods non robust to outliers (which are inherent to the problem), allowing the larger ones to have an important effect in the motion estimation even being few. Furthermore, the L^2 norm is more suitable for Gaussian data errors (Sun et al., 2008). Black and Anandan (1996) introduced the robust estimation of OF, by proposing the utilization of differentiable but non convex norms, associated with different type of data error distribution (i.e. Lorentzian distribution), in which outliers weight less in the optimization. The non convexity requires sophisticated and time consuming iterative methods to obtain the motion estimation.

The previous discussion applies mostly to the data term, namely, how well Eq. (1) is fulfilled, both in local and variational approaches. The addition of constrains can also be done, as said before, by imposing regularization. This can be included by adopting a regression model for each pixel neighbourhood (Black and Anandan, 1996) or by penalizing large gradients in the (u, v) field, leading to smoother solutions. This is indeed done by the Horn and Schunck method using an L^2 norm on the motion field gradient. An attractive option to penalize (u, v) gradients across the image is the Total Variation (TV) regularization (Rudin et al., 1992), which better preserves discontinuities in the motion field. The popularization of efficient computational techniques to solve non-differentiable convex optimization problems (Chambolle, 2004) has led to robust OF variational approaches that use the TV regularization for the motion field gradients and the L^1 norm for the data term, leading to the TVL1 model (Zach et al., 2007; Sánchez et al., 2013). This up-to-date technique provides a computationally fast framework for motion estimation, which is less sensitive to outliers and can recover piece-wise smooth motion vector fields preserving its discontinuities. Finally, it is worth noting that OF algorithms offer the utilization of down-scaling levels (M) to solve the dense motion estimation from lower to higher resolution images, a multi-level pyramid strategy that has proved to be best performing for large displacement fields (Wedel and Cremers, 2011; Sánchez Pérez et al., 2013).

In this context, we shall consider some of the OF methods to test, based on the recent advances in this field, their previous utilization for solar forecast and practical relevance, for instance, the availability of easily accessible open source algorithms. With all these in consideration, we selected two locally-constrained approaches, the Lucas and Kanade and Farnebäck methods, and two-variational approaches, the Horn and Schunck method and the Sánchez et al. implementation of the TVL1 algorithm (Zach et al., 2007). We will call them, respectively, LK, FRB, HS and TVL1. With this choice, we address the classical methods (LK and HS), a robust approach (TVL1) that is expected to perform similar or better to other more complex robust methods and practical methods previously assessed in the solar forecasting context (which includes the FRB). All of them have freely available open source implementations in python and/or C languages. The implementations of this work result in dense motion estimations composed of non-integer displacements, either due to the algorithms' nature or due to a posterior interpolation that converts a sparse motion estimation into a dense one. In the following subsections each method is briefly described.

All these algorithms offer a set of parameters that can be optimized. Some of these parameters refer to each method's formulation and other to their computational implementation. In this work we address the optimization of two parameters for each methodology, based on their a priori relevance. This decision is made by considering the previous knowledge of each method. We favour this approach in opposition to a full black-box optimization approach, as a way to understand the impact of each selected parameter in the forecasting performance (Subsection 5.1). For instance, the down-scaling levels M will be one of the optimized parameters for the OF algorithms, as it represents a common ground of analysis for them. The second parameter will be the most important method-specific value to tune, as explained in each subsection below. The rest of the parameters will be set as the default or recommended values for each methodology.

3.1.1. Lucas-Kanade method

The LK method determines the (u, v) values at each pixel by considering a constrain inside a rectangular window $\mathcal{W}$ centered on it. In the standard LK method this constraint forces the values of (u, v) to remain constant within the window, yielding an over-determined problem with no direct solution, but whose least mean square solution can be found by solving the following minimization problem for each pixel:

$$\underset{u,v}{\arg\min} \left\{ \int_{\mathcal{W}} \left(I_x \cdot u + I_y \cdot v + I_t \right)^2 \right\}. \tag{2}$$

If the texture of the neighbourhood $\mathcal{W}$ allows to solve the indetermination (satellite images comply with this), the solution to this problem for each window is unique and allows to estimate a dense motion field. We call this method LK-avg and implement it with the CalcOpticalFlowLK function of the OpenCV 2.x python libraries. This function does not have embedded the multi-level pyramid strategy, so it was implemented by us using the expand and reduce functions of the OpenCV library.

Another way to impose a motion field constrain under the Lucas and Kanade framework is to use an affine transformation, in which (u, v) must approximate a parametric function within the region (not just a constant value) in the form $(u, v) = f(x, y, p)$, where p represents the parameters. In such case, Eq. (2) is adapted to find the p optimal parameters in each neighbourhood:

$$\arg\min_{p} \left\{ \int_{\mathcal{W}} (\nabla I \cdot f(p) + I_t)^2 \right\}, \tag{3}$$

where $\nabla I = (I_x, I_y)$ denotes the image gradient. We call this method LK-afn and for its implementation we use the calcOpticalFlowPyrLK function of the OpenCV 3.x python libraries (Bouguet, 2000).

The two main parameters to tune locally for these two methods are the window length w (in pixels) and the down-scaling levels M, which will be optimized for the region. As square regions are considered, it follows that the amount of pixels in W is w^2. The other function's parameters, such as stopping criteria, iterations, etc., are set as default. Although these two variations of the LK method are based on a pixel neighbourhood, their computational implementation provide a dense motion estimation. This is done by performing the estimation at each pixel in the image with an overlapping one-pixel displaced spatial window.

3.1.2. Farnebäck method (FRB)

Farnebäck (2003) proposed an OF method based on a second order polynomial expansion of the neighbourhood of each pixel. This proposal is intended to deal with noisy sequences, for instance, sequences with high frequency variations in the CMV. Satellite images are prone to noise in the signal processing sense (Peng et al., 2013), hence this technique is an attractive option. The displacement of the polynomial expansion leads to a Lucas-Kanade-like minimization within each neighbourhood, given in Eq. (12) of Farnebäck (2003). This technique has been previously used for satellite solar forecasting in Kallio-Myers et al. (2020) and can provide a dense motion estimation without a high computational cost and without using interpolation techniques. The Farnebäck method is included in the python OpenCV 3.x libraries, calcOpticalFlowFarneback, and was used in this work. The available algorithm also includes a weighted Gaussian minimization in which central pixels in the neighbourhood are assigned higher importance. The algorithm has several parameters which can be locally tuned, from which we selected the window length w and the down-scaling level M, as we think these are the two more problem-specific (local adaptation with GOES-East satellite images) parameters. The other function's parameters are set as recommended in the OpenCV documentation website: pyr_scale=0.5, poly_n=5 and poly_sigma=1.1.

3.1.3. HS method

The Horn and Schunck variational method obtains the (u, v) vector field that minimizes the following convex and differentiable cost function across the whole image I:

$$\arg\min_{u,v} \left\{ \int_I (\nabla u)^2 + (\nabla v)^2 + \lambda \cdot (I_x \cdot u + I_y \cdot v + I_t)^2 \right\}. \tag{4}$$

The cost function consists of two terms: (i) a regularization term on the L^2 norm of gradients ∇u and ∇v, which promotes smoothness on the CMVs, and (ii) a data misfit term that enforces the OF constraint. The parameter λ controls the trade-off between these two terms. For its implementation, the function CalcOpticalFlowHS of the OpenCV 2.x versions was used. The local parameters to be tuned are the trade-off parameter λ and the down-scaling levels M, leaving the other function's parameters as default (in particular, the number of iterations was left in $N_{\text{iter}} = 100$). The multi-level pyramid strategy was implemented in the same way as in the LK-avg method.

3.1.4. TVL1 method

The TVL1 method has a similar formulation to the previous one, but the OF equation is enforced using the L^1 norm, and regularity is imposed by penalizing the total variation of the vector field (u, v):

$$\arg\min_{u,v} \left\{ \int_I |\nabla u| + |\nabla v| + \lambda \cdot |I_x \cdot u + I_y \cdot v + I_t| \right\}. \tag{5}$$

The total variation semi-norm promotes piece-wise smooth solutions, and allows to better preserve strong discontinuities in the vector field (Wedel and Cremers, 2011), for instance, those observed in interfaces between different height clouds moving in different directions. This problem can be solved using convex, non-differentiable optimization techniques Chambolle (2004). We use here the open-source implementation of Sánchez et al. (coded in C language) with its default parameters except for λ and M which are locally optimized. The other parameters are set as recommended in Sánchez et al. (2013): $\tau = 0.25$, $\theta = 0.30$, $\epsilon = 0.01$, $\eta = 0.5$ y $N_{\text{warps}} = 5$.

3.1.5. PIV method

For the block matching algorithm of Lorenz et al., we use the implementation provided by the OpenPIV python library (extended_search_area_piv). The details of this widely-used method have been briefly provided in the introduction and a complete description can be found in Kühnert et al. (2013). Previous to its utilization, two parameters need to be optimized to the region: the size of the neighbouring block (W_n) and the size of the search area (W_s) in the previous image, both assumed square, and so defined by their lengths w_n and w_s, respectively. Due to computational reasons, a grid-search of 30×30 px was used for the motion field space support. The resulting sparse CMV is then interpolated by us to each image pixel with a bi-linear interpolation of u and v. We found no significant performance difference when using a 5×5 px support. The choice of the space step affects the overlap function's parameter, which was set to $w_n - 30$ px. The rest of the function's parameters are fixed to their default values.

3.2. Training

The five algorithms presented before have parameters to tune. This optimization may depend on the region's characteristics, specially its clouds typical regime. Here we have tuned these parameters for the territory specified in Figure 1. Images from year 2017 were used for this purpose, leaving the year 2016 for evaluation. The optimization is done for the first time step, this is, the present time and previous images (30 minutes difference) are used to estimate the next one via each CMV technique, and the optimal parameters are the ones which minimize the average root mean squared deviation (RMSD) of this estimation. The extrapolation technique to estimate the next image from the current image $I(x, y, t)$ and the non-integer (u, v) field is the standard one (backward search), in which $I(x, y, t + 1)$ is constructed pixel-by-pixel with a sub pixel bilinear interpolation of $I(x - u, y - v, t)$. The choice to optimize the parameters for the first time step relies in the iterative construction of the predicted images, in where the prediction $I(x, y, t+h)$ is generated by using the CMV and the previous prediction $I(x, y, t + h - 1)$, initiating with $I(x, y, t)$, so each extrapolation is only for one time step ahead at each time. Subsection 5.1 shows the training of each CMV technique and provides the optimal parameters for each methodology in the region. As will be observed in Subsection 5.2, different extrapolation methods yield to quite similar performance results, hence the choice of the extrapolation procedure for training is not expected to affect significantly the parameters tuning.

3.3. Solar irradiation conversion

The last step of the prediction chain is the conversion of the predicted hours-ahead images to global horizontal solar hourly irradiation. This is done in this work by means of a Cloud Index Model (CIM) that has been specifically adjusted to the target region in Laguarda et al. (2020). The model combines the ESRA clear sky model (Rigollier et al., 2000) with a simple linear cloud attenuation factor, $F(\eta)$, which is based on the cloud index η, calculated from the albedo images ρ. This model is referred in the cited article as CIM-ESRA. The Linke turbidity factors to use in the region are given as seasonal daily trends in Laguarda and Abal (2016). The model has for the region a relative mean bias deviation (rMBD) of about -1% and a relative root mean square deviation (rRMSD) of 12.5%, both expressed as percentage of the measurements average and at hourly scale. The use of these seasonal cycles does not downgrades the optimized model's performance in a significant extent for GHI estimation in the region. For instance, the use of the McClear model (Lefèvre et al., 2013) instead of the ESRA model with the CIM strategy, CIM-McClear, achieves the same bias and a rRMSD of 12.1%, which is slightly lower than the CIM-ESRA. Due to the ease of implementation and simplicity, we prefer here the CIM-ESRA model for image-to-GHI conversion.

4. Performance metrics

The evaluation makes use of common metrics to assess the performance of deterministic forecast, namely, the mean bias deviation (MBD), the root mean square deviation (RMSD) and the Forecasting Skill (FS), as a function of the hourly forecast horizons h, from 1h to 5h. The first two metrics are also given in their relative versions, expressed as a percentage of the observed average (rMBD and rRMSD, respectively). The FS metric quantifies the gain in terms of RMSD of the prediction with respect to a persistence procedure. The performance evaluation is presented at two levels: albedo image and solar hourly irradiation. More details on the evaluation framework for deterministic solar forecast can be found in (Yang et al., 2020a). For the sake of completeness, we describe in the following the image level evaluation (which is rather uncommon in the literature) and the persistence procedures used for both levels, especially for the irradiation level, as its choice critically affects the FS metric. The two-levels evaluation is intended to bridge the gap between them, in particular, to understand at which extent an improvement at image level impacts the performance at irradiation level, which is the ultimate goal of these techniques.

The image level evaluation is carried out by comparing the predicted albedo images with their corresponding ground truth (real albedo image) for each h. All images' pixels are used, except for a small frame of 50 px around them ($\simeq 0.5\%$ of the image). This is intended to avoid image's border artifacts and problems, that are common in this strategy. An MBD and RMSD are obtained for each image comparison, which are then averaged for the same h to obtain the overall performance. Images from the year 2016 are used for this purpose while images from 2017 are used for algorithms' optimization, thus a $\simeq 50/50$ split is implicitly performed for training/test. The evaluation is done for each CMV methodology and for the image persistence, which is implemented here by simply maintaining the dimensionless albedo image constant in the future h time steps. The MBD and RMSD normalization is done by using the average of each image mean value in the testing set. The FS is constructed from the RMSD curve of each method and the persistence.

The irradiation level evaluation is performed by using the predicted and measured time-series of GHI at the sites in Table 1. For this evaluation in particular there is much work done in defining an exigent and universal persistence benchmark. As the guidelines in Yang et al. (2020a) indicate, the persistence should be based on the clear sky index, k_c, the normalization of the GHI by the corresponding estimated clear sky irradiation from a clear sky model. This dimensionless magnitude eliminates the geometrical behavior of the GHI and is a better signal, in the stationary sense, than the clarity index, k_t, in where the normalization is done with the horizontal extraterrestrial irradiation. The persistence used in this work is therefore based on k_c and, as we used the locally adjusted ESRA model for image-to-GHI conversion, we used it also for the clear sky index calculation. There are some ways to use k_c for persistence, from where we shall consider two of them as reference for the the later discussion: (i) the regular persistence ($k_c(t+h) = k_c(t)$), PERS, and (ii) the convex combination of the regular persistence with the climatological value (Yang, 2019b), CC. The first one is the simple naive and widely-known persistence procedure and the second one is the recommended benchmark (Yang et al., 2020a), as it yields to more exigent performance levels than the former. In Alonso-Suárez et al. (2021) an analysis of these benchmarks for the target region is provided, including the sensibility of the second method to the training period, which is low. In this work we use the measurements during the training year (2017) to calculate the covariance values required for the CC.

5. Results

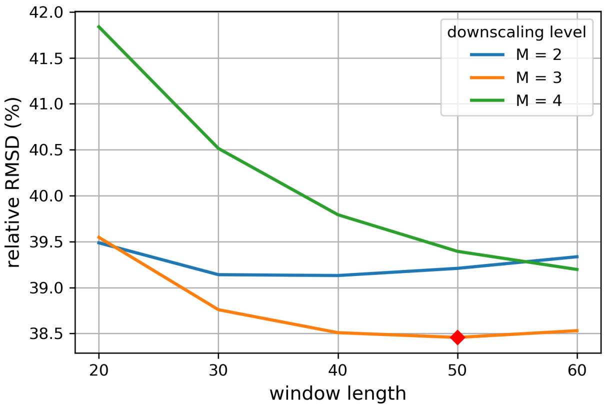
Reproducibility is an important feature in both solar resource assessment and forecasting works (Yang, 2019a). This is a quite difficult task when a large volume of satellite images is involved. In this work we address this issue by using publicly available satellite images and open python libraries or open source algorithms. We also provide a detailed description of the methods and their implementation, so they can be reproduced elsewhere. In this section we present our results regarding the algorithms optimization (Subsection 5.1) and performance evaluation, at image (Subsection 5.2) and irradiation (Subsection 5.3) levels. For the sake of clarity, in this section we focus the discussions mainly on the rRMSD metric and FS score. The rMBD plots along with a brief discussion are provided in AppendixA.

5.1. Optimization

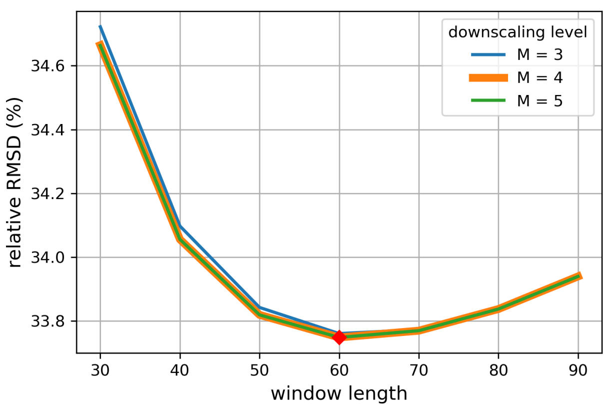
The optimization results for each algorithm are presented in Table 2 and Figure 2. Table 2 shows the optimal parameters obtained for each methodology. The plots of Figure 2 show the rRMSD trend as a function of these parameters, being minimum at the optimal values. As the adjustment is done with the training set (images for the 2017 year), the rRMSD curves of Figure 2 must not be taken as each method real performance. Instead, these plots shall be considered as a help to visualize that these optimum parameters indeed exist and that slight variations from them are not critical. All the training rRMSD curves are smooth around the optimal values of λ, w or w_n. Intermediate values of the down-scaling levels M achieve better training performance and variations of ± 1 from the optimal value typically do not downgrade the training rRMSD in more than 1%. The optimal values of Table 2 will be used in the following for performance assessment of the methods, but over the evaluation set.

Table 2: Optimal parameters for each CMV technique.

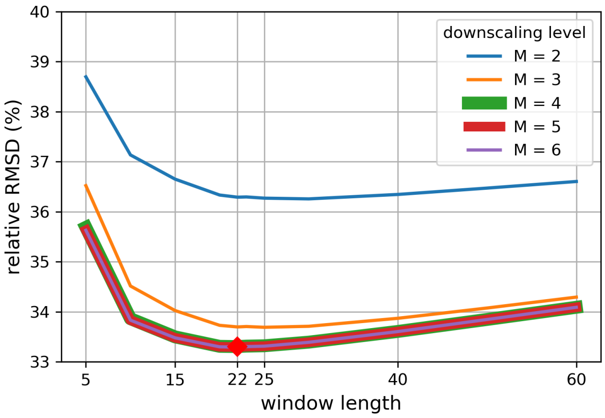
LK-avg		LK-afn		FRB		HS		TVL1		PIV	
w (px)	50	w (px)	60	w (px)	22	λ	0.200	λ	0.055	w_n (px)	120
M	3	M	4	M	5	M	5	M	6	w_s (px)	144



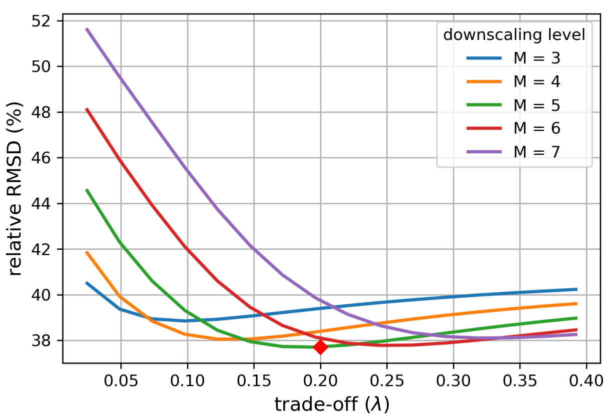
(a) LK-avg algorithm.



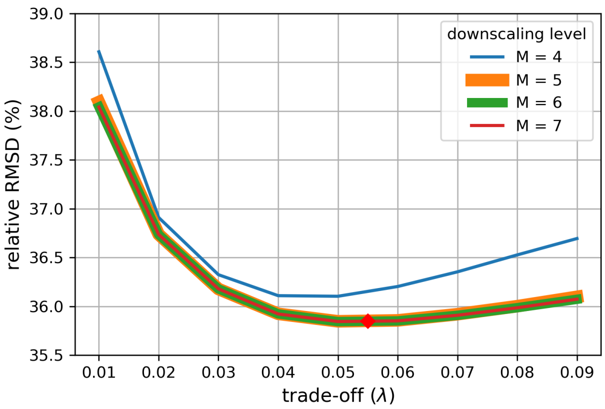
(b) LK-afn algorithm.



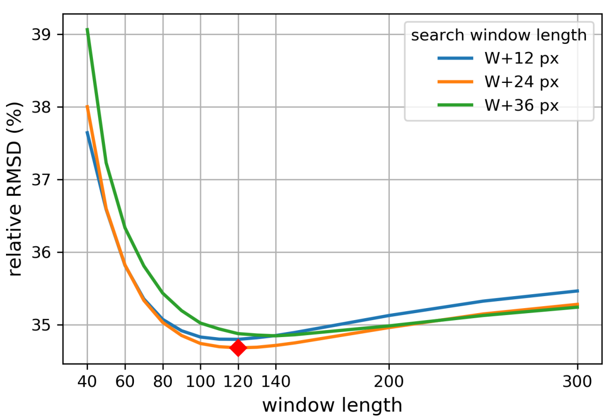
(c) FRB algorithm.



(d) HS algorithm.



(e) TVL1 algorithm.



(f) PIV algorithm.

Figure 2: Optimization of the CMV methodologies at image level. The curves (an their optimum) are not representative of the real prediction performance, as the parameters adjustment is done over the training set (year 2017) and for $\Delta t = 30$ minutes.

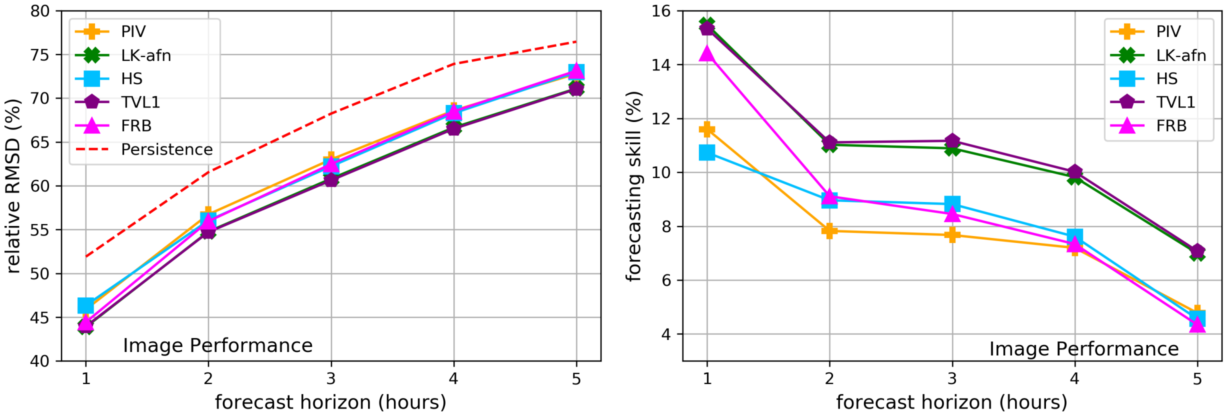
5.2. Image forecasting

Performance evaluation at the image level is carried out by comparing, pixel-to-pixel, the predicted albedo images at each time horizon with their corresponding real image. Image prediction is a very challenging task, as it requires to predict the irregular cloudiness at each pixel. Table 3 presents the performance metrics as a function of the forecast horizon for each optimized method and the image persistence. Figure 3 shows graphically its rRMSD and FS information. The image extrapolation method used here is the baseline pixel's backward search with a sub-pixel bilinear interpolation. All CMV forecast strategies beat the image persistence with positive FS and have rather similar rRMSD. It should be noted that all methods are locally optimized, so the performance differences tights between each other and are better observed with the FS score (Figure 3b). At image level, the TVL1 and LK-afn methods outperform the rest for all forecast horizons, having similar FS scores. For the first hour ahead the FRB methods also provides competitive performance, but it downgrades for higher forecast horizons ($h \geq 2$). Overall, the best image forecast methods (TVL1 and LK-afn) provide a gain in FS of $\simeq 4\%$ for $h = 1$ and of 2-3% for $h \geq 2$ over the worst performing methods (HS and PIV). Table 3 also provides the performance comparison between the LK-avg and LK-afn techniques. It is observed that the LK-afn has an unequivocal better performance, so for the sake of clarity we will only consider this LK method in the following discussion. It shall be pointed out that the LK-avg method was also assessed in our tests at solar irradiation level and the same conclusion holds: it did not improve the performance of the LK-afn method.

Table 3: Image forecasting performance as a function of the forecast horizon. The reference average is $\bar{\rho} = 0.28$.

h (hour)	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
method	relative MBD (%)					relative RMSD (%)					forecasting skill (%)				
LK-avg	+1.3	+1.6	−0.9	−4.9	−5.6	46.8	56.5	62.8	68.6	73.2	+9.7	+8.2	+7.9	+7.1	+4.3
LK-afn	+1.6	+2.0	−0.5	−4.6	−5.7	43.9	54.8	60.8	66.7	71.1	+15.5	+11.0	+10.9	+9.8	+7.0
FRB	+1.6	+1.7	−1.1	−5.7	−6.3	44.4	55.9	62.5	68.5	73.1	+14.4	+9.1	+8.5	+7.3	+4.3
HS	+0.4	+0.5	−2.9	−7.3	−9.0	47.5	55.5	61.6	67.5	72.0	+10.7	+9.0	+8.8	+7.6	+4.6
TVL1	+0.8	+0.9	−2.1	−6.4	−7.7	43.9	54.7	60.6	66.5	71.0	+15.3	+11.1	+11.2	+10.0	+7.1
PIV	+1.5	+1.7	−0.2	−3.7	−3.8	45.9	56.7	63.0	68.6	72.8	+11.6	+7.8	+7.7	+7.2	+4.8
persistence	+2.1	+1.1	+0.3	−1.7	−5.0	51.1	61.2	67.5	72.8	76.7	–	–	–	–	–

A word shall be said about the extrapolation strategy for image forecasting, the second step of the forecasting methodology. We tested three techniques that use the non-integer CMV: (i) the usual backward pixel's search in which the opposite (u, v) field is used to find in the previous image the value to translate via nearest neighbour (DN) and with a sub pixel bilinear interpolation (DL), (ii) an optimized (block size) block moving technique from the previous to the predicted image (OW). In this last method, if two or more values get to the same pixel, then the average is calculated. Figure 4 shows the rRMSD achieved over the



(a) Relative RMSD. (b) Forecasting Skill.

Figure 3: Performance comparison at image level for all the forecasting algorithms.

testing set of these extrapolation methodologies when using the CMV estimated by the TVL1 algorithm, as a representative example. The results when using the other CMV algorithms are similar, as shown in Figure B.11 of AppendixB. Two things can be observed. First, the performance when using the different extrapolation strategies is similar, and second, the common backward search with bilinear interpolation provides the best prediction performance from these three, hence is the one used in this work (previously and afterwards).

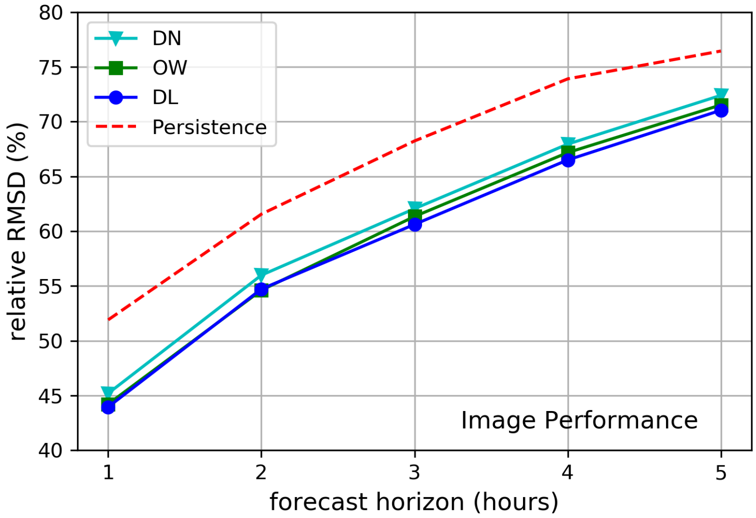
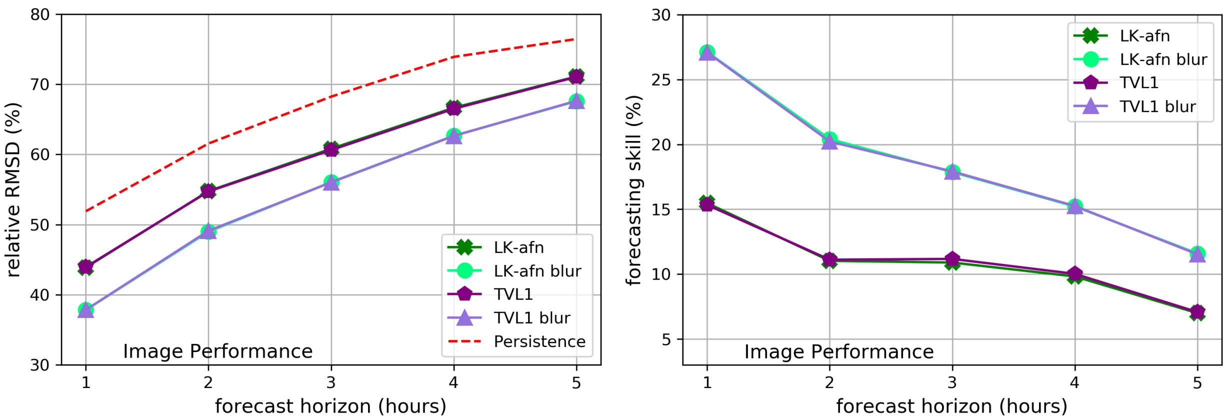


Figure 4: rRMSD of the TVL1 algorithm for different image extrapolation strategies (performance over the training data set).

To finish the image level forecasting analysis, two further studies are presented: (i) the performance gain obtained by using a spatial smoothing (image blurring) and (ii) the relative RMSD dependence with the amount of cloudiness in each image, approximated here by the mean albedo of the ground truth image.

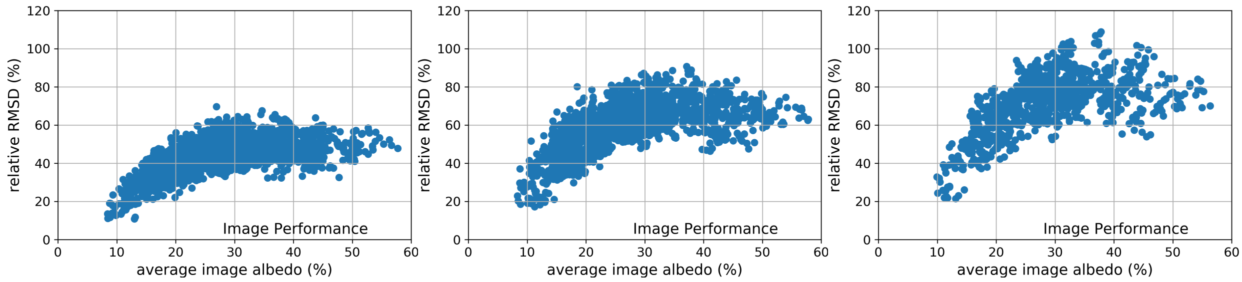
The first analysis is reported in Figure 5 for the two best performing techniques, the TVL1 and LK-afn algorithms. The blur inspected here is the 20×20 px fixed spatial average, as an exploratory analysis that will be complemented in the next subsection. It is clear that the blurring improves the forecasting accuracy at image level, increasing the FS in $\simeq 12\%$ for $h = 1$ and in $\simeq 5\%$ for $h = 5$ with respect to the single-pixel approach without spatial smoothing. The improvement of this procedure is similar for both techniques, as the curves in each plot are almost the same. The second analysis is shown in Figure 6 for three time horizons (1, 3 and 5 hours ahead) and the TVL1 algorithm (without blurring), as the plots are quite similar for all techniques. The figure shows the relative RMSD trend as a function of the mean albedo. Each dot is a comparison between a predicted image and the corresponding ground truth for the given forecast horizon. As expected, the prediction downgrades with the time horizon. The mean rRMSD value of each plot coincides with the quantitative values provided in Table 3. It is observed that images that have a lower mean albedo, i.e. less presence of clouds, are better predicted, achieving, for instance, rRMSD values of 15-20% for 1-hour ahead (at $\bar{\rho} \simeq 10\%$). The rRMSD increases with the cloudiness amount in the range of $\bar{\rho} =$ 10-35%, with a flat peak around mid values ($\bar{\rho} \simeq$ 35-40%). Then, for images with high average albedo, the performance tend to slightly improve, with a soft decrease in rRMSD for $\bar{\rho}$ values above 40%. This overall behavior becomes more evident as the forecast horizon increases, as can be seen in Figure 6. The location of the flat peak also tend to increase with the forecast horizon, being located at $\bar{\rho} \simeq 30\%$ for $h = 1$ and at $\bar{\rho} \simeq$ 35-40% for $h = 5$. The plots of Figure 6 have the same x and y axis to facilitate the inter-comparison.



(a) Relative RMSD.

(b) Forecasting Skill.

Figure 5: Performance at image level with and without spatial smoothing for the best two methods (TVL1 and LK-afn).



(a) 1 hour forecast horizon. (b) 3 hour forecast horizon. (c) 5 hour forecast horizon.

Figure 6: Relative RMSD dependence with the mean albedo of the ground truth image for the TVL1 forecasting algorithm.

5.3. GHI forecasting

The GHI forecast is evaluated on a hourly basis, so, hourly irradiation values should be compared with hourly predicted irradiation. A pixel in an instantaneous image fails to accurately represent the hourly average behavior of both cloudiness and solar irradiation at given sites. To account for this, assessment models usually make use of spatial smoothing, so that the space average represents the time average via an ergodic assumption (Laguarda et al., 2020), achieving a fair comparison and reducing the uncertainty of the hourly solar estimation. The third step of the forecasting methodology is to convert instantaneous predicted images into hourly irradiation values, so the same principle applies. This can be done in two ways. The first one is by using the optimal spatial smoothing of the time t assessment (fixed for all forecast horizons), thus considering this third step as independent from the forecast itself, which is then exclusively left to the image prediction. The second one is by obtaining the optimal spatial smoothing for each forecast horizon, hence exploiting the procedure to also filter inaccuracies in the image prediction step (as it is equivalent to an image blurring). In the following, both approaches are analyzed. Although the single pixel utilization does not hold for hourly values, it is also included in the discussion, as this is not usually provided in the literature. However, detailed results for single pixel are left to AppendixC.

The spatial smoothing effect is illustrated in Figure 7 using the TVL1 algorithm as example. The plot shows the rRMSD curve as a function of the smoothing window length (in px) for each forecast horizon. The optimum window length is shown in red squares and the analysis includes the satellite assessment optimum ($h = 0$). The assessment optimal averaging window is 20×20 px and higher values are obtained for hourly prediction, increasing with the forecast horizon, as larger h imply higher inaccuracies. The curve's minimum is better marked for assessment than for prediction, and it flattens out with the forecast horizon. A non negligible rRMSD improvement is observed when using the optimal window in comparison to the fixed one. This improvement increases with the lead time. The same analysis, but for the other algorithms, is provided in AppendixD, where the same trends are observed with small changes in the optimal window lengths depending on the methodology.

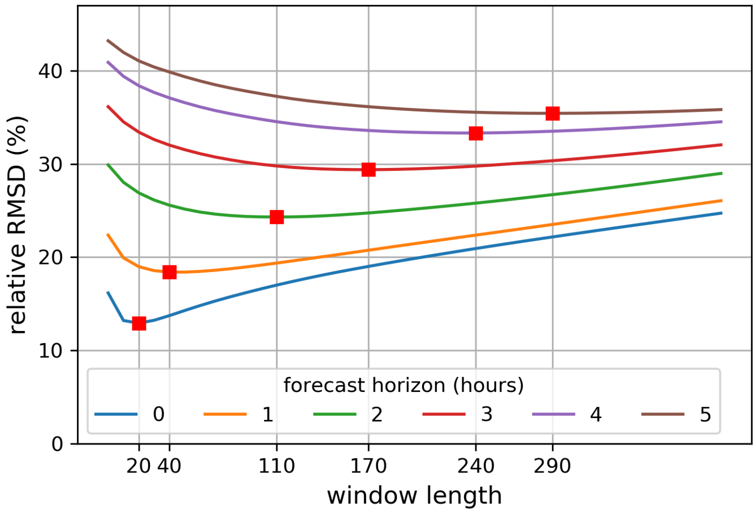


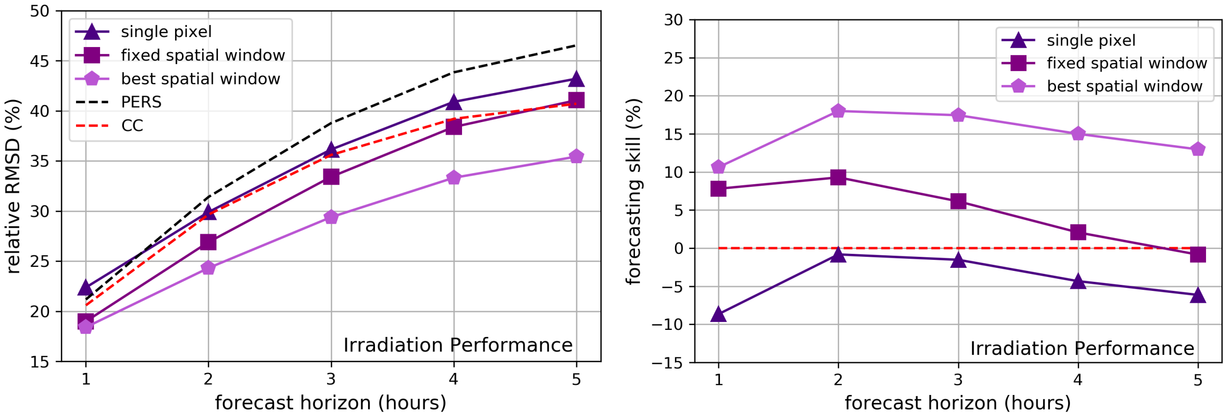
Figure 7: Effect of the image spatial smoothing on the relative RMSD of the best performing algorithm (TVL1)

.

To conclude the smoothing analysis, Figure 8 shows the rRMSD and FS curves of the TVL1 method when using a single pixel, a fixed and an optimal spatial smoothing. The rRMSD plot (Figure 8a) includes the CC and PERS procedures but the FS is calculated with respect to the CC. Single pixel shows the worst performance of the three approaches, as expected. It overcomes the PERS performance but not the CC. Both spatially smoothed versions outperform CC at almost all lead times, with the exception of the 5-hours ahead forecast with fixed window, in which the CC provides a slightly lower rRMSD. The optimal spatial smoothing provides the best performance, with a significant gain over the fixed window approach, especially for the larger forecast horizons, and showing FS scores between 11-18%. As it will be shown next, the FS curve of Figure 8b of the spatially smoothed TVL1 algorithm represents also the best solar irradiation performance among all the methods analyzed in this work. For further reference, Table 4 provides the relative RMSD of the hourly irradiation assessment ($h = 0$). This is the baseline uncertainty of the satellite-to-irradiation model for each spatial smoothing. As the fixed smoothing is defined here as the optimal assessment spatial window, both the fixed and optimal smoothing coincide at $h = 0$, being of rRMSD $\simeq 13\%$. The findings in Table 4 are in agreement with Laguarda et al. (2020), where an analysis of the satellite assessment uncertainty for the region is provided as a function of the spatial average.

Table 4: Relative RMSD of the irradiation assessment ($h = 0$) for different spatial smoothing.

	single pixel	**fixed smoothing**	**optimal smoothing**
relative RMSD (%)	16.2	12.9	12.9



(a) Relative RMSD metric. (b) Forecasting Skill score.

Figure 8: Comparison of the GHI prediction performance of the TVL1 algorithm with the single pixel approach and with a fixed and optimal spatial windows. PERS and CC benchmarks are given as reference, and the FS is calculated with the latter.

Tables 5 and 6 present the performance metrics when using the fixed and optimal window's length, respectively, and for all the algorithms. It includes both persistence procedures, but the FS is calculated by using the CC, as it is the most exigent benchmark. The rRMSD and FS are illustrated in Figure 9, where the fixed window utilization is at the left panels and the optimal window utilization is at the right panels.

When using a fixed spatial window, the only CMV algorithm that outperforms the CC for almost all lead times is TVL1 (except for the last one, in which is slightly beaten by the CC). LK-afn, FRB and HS algorithms only provide positive FS in the first three lead times, and PIV only in the first one. This means that, under the fixed window strategy, only TVL1 can be considered a clear improvement over the most exigent benchmark. On the other hand, it shall be noted that all methods succeed to outperform the regular PERS procedure, with which satellite-based forecasting strategies has been evaluated in the past. For the best knowledge of the authors, this is the first work in which the CC persistence is used as benchmark for solar satellite forecast up to 5 hours ahead.

When using the optimal spatial smoothing, all methods succeed to outperform the CC benchmark, providing positive FS. The methods performance tighten, as the spatial smoothing effect provides different gains for each algorithm with respect to the fixed window. However, the general conclusions are the same: the TVL1 method is the best overall method and the PIV is the weakest one. Under the optimal smoothing framework the other three methods (LK-afn, FRB and HS) provide a competitive performance to that of the TVL1, and for $h = 1$ the HS method succeed to slightly outperform it. The second best performing methods are the HS and LK-afn algorithms, but they downgrade for the last two lead times. The FRB method, although promising at image level, especially for the first lead time, does not stand out at any forecast horizon for GHI prediction.

Table 5: Irradiation forecasting performance metrics with the fixed spatial smoothing.

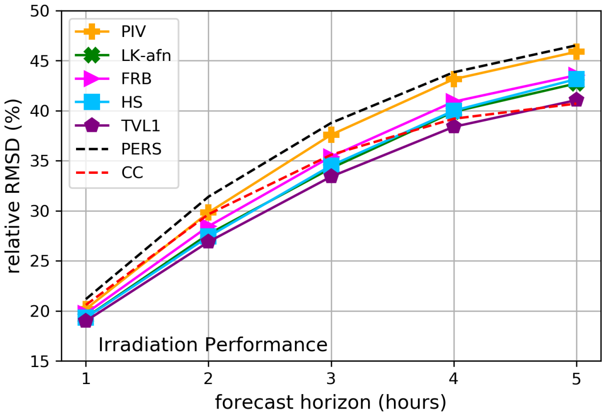
h (hour)	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
method	relative MBD (%)					relative RMSD (%)					forecasting skill (%)				
LK-afn	−1.3	−1.9	−2.3	−3.0	−3.0	19.3	27.7	34.3	39.9	42.7	+6.1	+6.7	+3.8	−1.8	−5.0
FRB	−1.3	−2.1	−2.9	−4.0	−4.5	19.7	28.5	35.4	40.9	43.5	+4.1	+4.0	+0.5	−4.3	−7.0
HS	−1.3	−2.0	−3.0	−4.0	−4.5	19.3	27.4	34.5	40.0	43.2	+6.2	+7.5	+3.0	−2.0	−6.0
TVL1	−1.0	−1.2	−1.4	−1.9	−2.0	19.0	26.9	33.4	38.4	41.1	+7.8	+9.3	+6.1	+2.1	−0.9
PIV	−0.6	−0.1	+0.1	−0.4	−0.3	20.2	29.8	37.6	43.2	45.9	+2.0	−0.6	−5.6	−10.1	−12.7
CC	−0.8	−1.3	−1.9	−2.5	−2.6	20.6	29.7	35.6	39.2	40.7	–	–	–	–	–
PERS	−0.8	−1.4	−2.1	−2.8	−3.0	21.2	31.4	38.8	43.8	46.5	–	–	–	–	–

Table 6: Irradiation forecasting performance metrics with the optimal spatial smoothing.

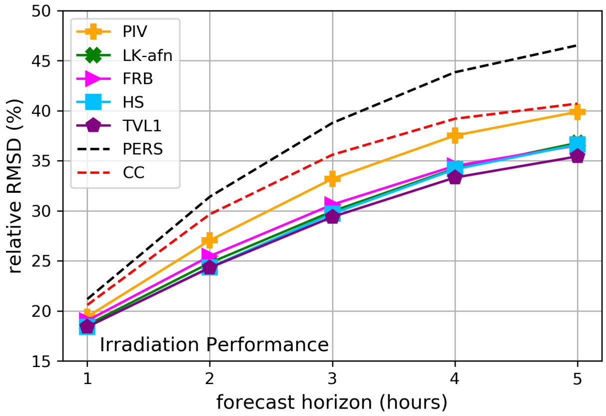
h (hour)	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
method	relative MBD (%)					relative RMSD (%)					forecasting skill (%)				
LK-afn	−0.9	−0.9	−1.2	−1.5	−1.2	18.6	24.8	30.0	34.2	36.8	+9.6	+16.4	+15.9	+12.7	+9.6
FRB	−1.0	−1.0	−1.6	−2.0	−1.7	19.0	25.5	30.6	34.5	36.5	+7.6	+14.1	+14.0	+12.0	+10.3
HS	−0.8	−0.9	−1.6	−2.3	−2.4	18.4	24.4	29.7	34.2	36.6	+10.7	+17.9	+16.5	+12.9	+10.1
TVL1	−0.7	−0.3	−0.3	−0.4	−0.2	18.4	24.3	29.4	33.3	35.4	+10.6	+18.0	+17.4	+15.0	+13.0
PIV	−0.2	+0.7	+1.0	+1.0	+1.5	19.4	27.1	33.2	37.5	39.9	+5.7	+8.8	+6.7	+4.3	+2.0
CC	−0.8	−1.3	−1.9	−2.5	−2.6	20.6	29.7	35.6	39.2	40.7	–	–	–	–	–
PERS	−0.8	−1.4	−2.1	−2.8	−3.0	21.2	31.4	38.8	43.8	46.5	–	–	–	–	–

6. Conclusions

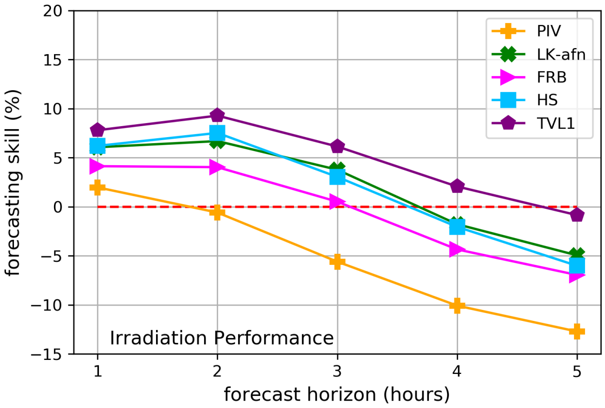
Satellite-based solar forecasting methods still present a large space for improvements. In this work, a comprehensive assessment of up-to-date techniques based on static two-dimensional cloud motion vectors (CMV) is provided for a region with intermediate short-term solar irradiance variability. The performance analysis is done for hourly forecast up to 5 hours ahead at image and irradiation level, and consist of five methods: the popular block-matching algorithm and four optical flow methodologies. A detailed analysis is presented, discussing also the image extrapolation and the spatial smoothing strategy, issues which are not commonly addressed. This work compares the five CMV methods, being the larger solar satellite forecast comparison to date, and uses the exigent clear-sky index convex combination (CC) of persistence and climatology as performance benchmark, that has not been previously reported in this satellite framework. The work also provides the optimization of the two main parameters for each method to be used in the Pampa Húmeda region of South America, which may be extrapolated to other regions with similar climate characteristics or solar resource variability. There may be further room for improvements by optimizing the



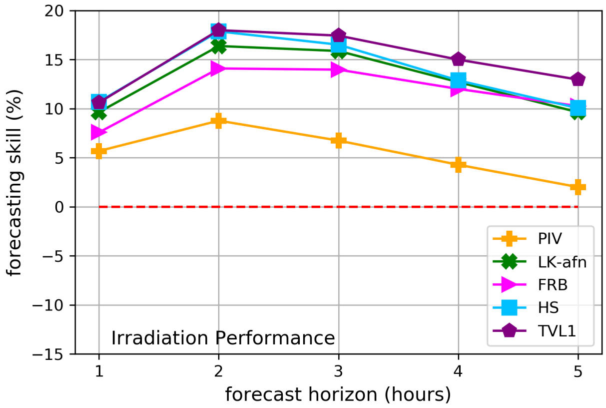
(a) Relative RMSD metric (fixed smoothing).



(b) Relative RMSD metric (optimal smoothing).



(c) Forecasting Skill score (fixed smoothing).



(d) Forecasting Skill score (optimal smoothing).

Figure 9: Performance comparison at irradiation level for all the forecasting algorithms.

other algorithms' parameters, which here are set as default or recommended, and/or by considering different parameters' values depending on the cloudiness' amount, type or regime that is observed in the sequence. Both detailed studies are left for future work.

We found that the TVL1 method is the best satellite CMV strategy for this region. It provides the best performance for image and hourly irradiation forecasting. For instance, it is the only method that outperforms the CC benchmark for most lead times when using a fixed spatial smoothing window adjusted for hourly solar assessment and it provides almost the best performance for all forecast horizons when the optimal spatial smoothing is used. On the other hand, the classical block-matching algorithm, implemented here as the PIV algorithm, resulted in the weakest option for all forecast horizons at image and irradiation levels, using fixed or optimal smoothing. We also found that the LK method with an affine transformation constrain performs better than the simplest LK approach in which the motion field is forced to be constant within a neighborhood. Same qualitative results are observed at image and irradiation levels, but the quantitative relative RMSD values differ in scale, as a consequence of different data ranges, geometrical behavior and reference average values (satellite albedo and hourly irradiation). Optimal spatial smoothing, performed ad-hoc for each method, changes some performance curves as a function of the lead times, as it tights the forecasting skill curves, but does not modify the general conclusion. This optimal smoothing procedure also allows all the methods to outperform the exigent CC solar benchmark.

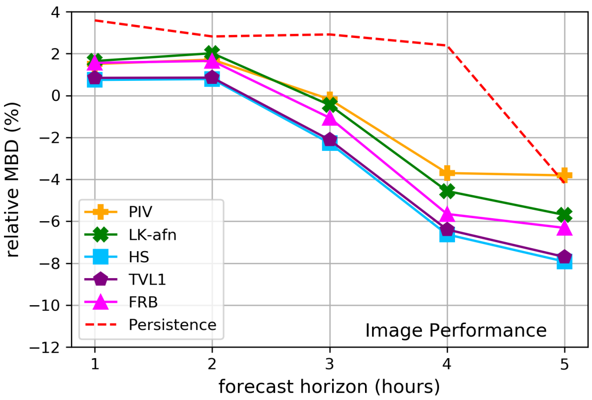
The findings of this comparison and the parameters' tuning may be location-dependent, so further similar studies are needed in other areas of the globe. Especially, the relation between the regional cloudiness regime (cloud type, cloud's typical development and movement, intermittency, etc.) with the CMV performance and with each method's parameters should be further inspected. The image acquisition time rate is another potential source of discrepancy, being here of 30 minutes, with non-regular acquisition during daylight periods. The impact of the acquisition time rate is an interesting study that can be performed for this region with the current 10-minutes GOES-16 satellite images, and is part of our current work.

Acknowledgments

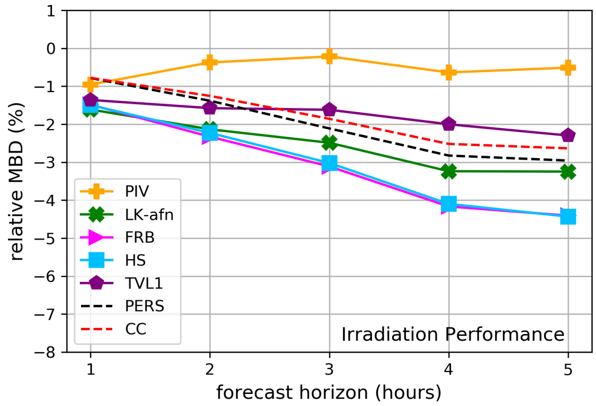
The authors gratefully acknowledge financial support from Uruguay's National Research and Innovation Agency (ANII) under the FSE-ANII-2015-109937 and FSE-ANII-2016-131799 grants.

AppendixA. Mean bias deviation plots

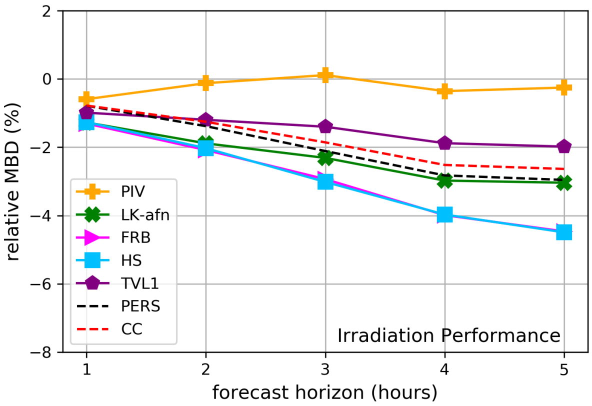
Figure A.10 shows the relative MBD curves for all the settings: at image level and irradiation level with a single pixel approach and fixed and optimal spatial smoothing. This is the same information that was shown in Tables 3, 5 and 6 and is shown in Table C.7. There is not a clear relationship between biases at image and irradiation levels. The TVL1 algorithm is also the best choice from the MBD point of view, as seen in Figure A.10d. It shall be noticed that the PIV algorithm is the only one which yields positive MBD at irradiation level, being always upper from the rest.



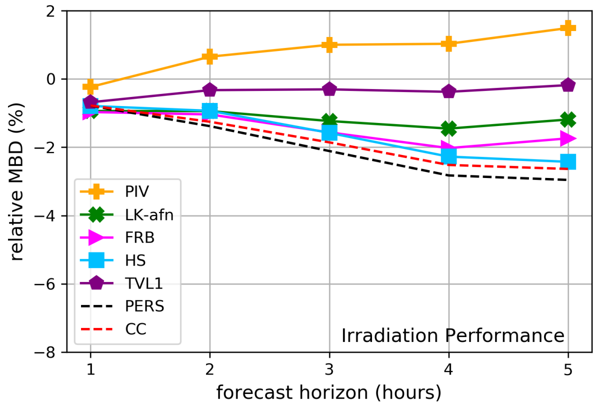
(a) Albedo image level.



(b) Irradiation level (single pixel).



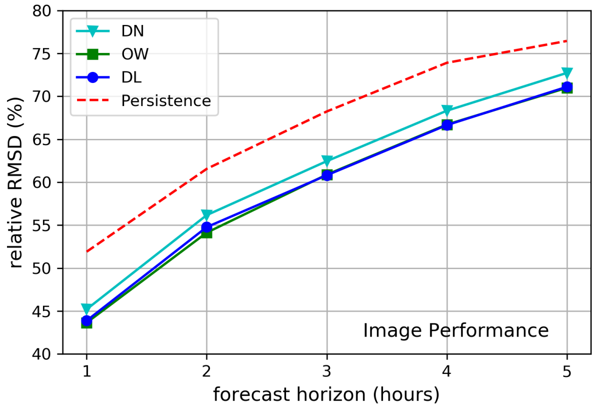
(c) Irradiation level (fixed spatial smoothing).



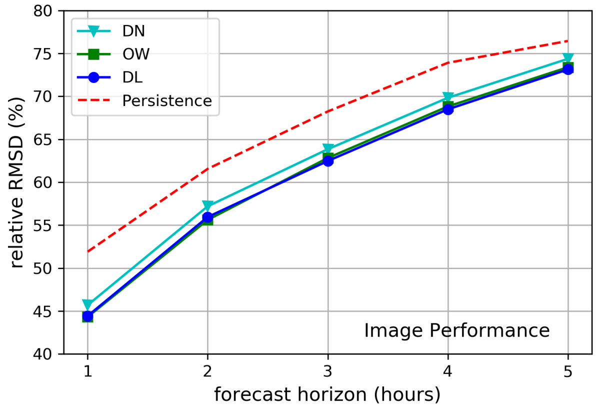
(d) Irradiation level (optimal spatial smoothing).

Figure A.10: Relative MBD as a function of the forecast horizon at image and irradiation levels.

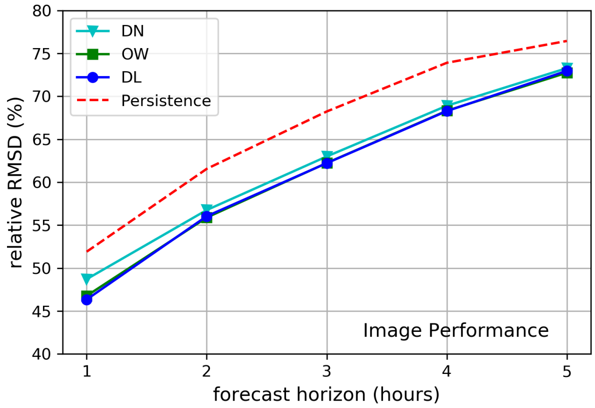
AppendixB. Extrapolation strategies plots for each method



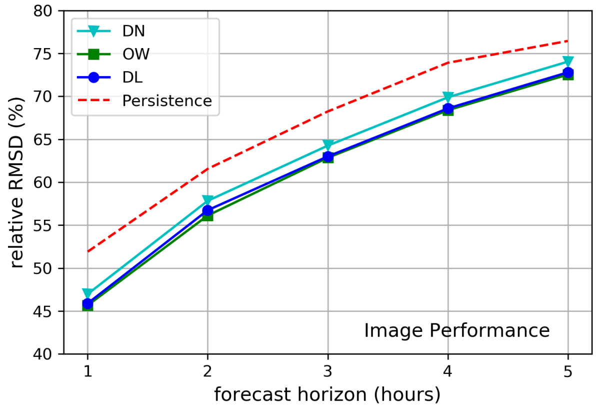
(a) LK-afn algorithm.



(b) FRB algorithm.



(c) HS algorithm.



(d) PIV algorithm.

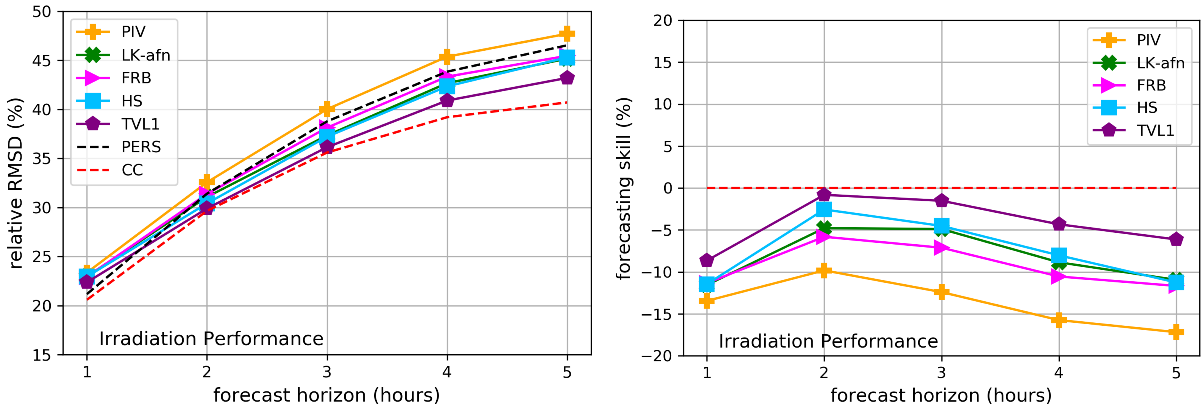
Figure B.11: Relative RMSD for different image extrapolation strategies in each method (results over the test data set).

AppendixC. Single pixel GHI forecasting

For the sake of completeness, Table C.7 and Figure C.12 provide, respectively, the performance metrics and plots regarding the single pixel approach for GHI conversion. The rMBD plot was given in AppendixA, jointly with the others.

Table C.7: Irradiation forecasting performance metrics with the single pixel approach.

h (hour)	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
method	relative MBD (%)					relative RMSD (%)					forecasting skill (%)				
LK-afn	−1.6	−2.1	−2.5	−3.2	−3.2	23.0	31.1	37.3	42.7	45.2	−11.6	−4.8	−4.9	−8.9	−11.0
FRB	−1.5	−2.3	−3.1	−4.2	−4.4	22.9	31.4	38.1	43.3	45.5	−11.4	−5.8	−7.1	−10.6	−11.7
HS	−1.5	−2.2	−3.0	−4.1	−4.4	23.0	30.4	37.2	42.3	45.3	−11.5	−2.6	−4.5	−8.0	−11.3
TVL1	−1.4	−1.6	−1.6	−2.0	−2.3	22.4	29.9	36.1	40.9	43.2	−8.7	−0.8	−1.5	−4.3	−6.1
PIV	−1.0	−0.4	−0.2	−0.6	−0.5	23.4	32.6	40.0	45.4	47.7	−13.5	−9.8	−12.4	−15.7	−17.2
PERS	−0.8	−1.4	−2.1	−2.8	−3.0	21.2	31.4	38.8	43.8	46.5	–	–	–	–	–
CC	−0.8	−1.3	−1.9	−2.5	−2.6	20.6	29.7	35.6	39.2	40.7	–	–	–	–	–

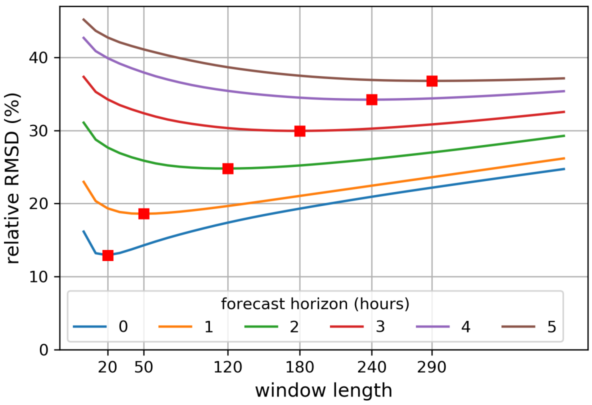


(a) Relative RMSD metric.

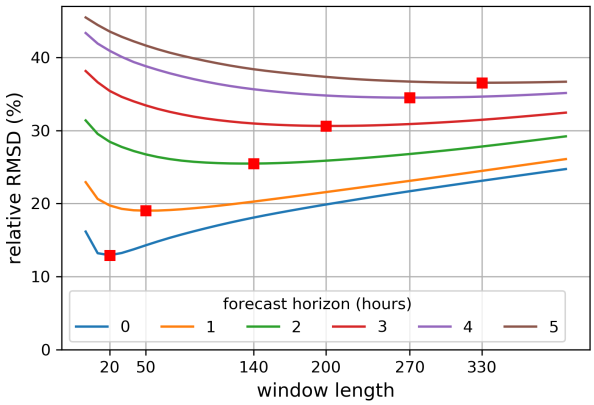
(b) Forecasting Skill metric with convex persistence.

Figure C.12: Performance comparison at irradiation level for all the forecasting algorithms (single pixel approach).

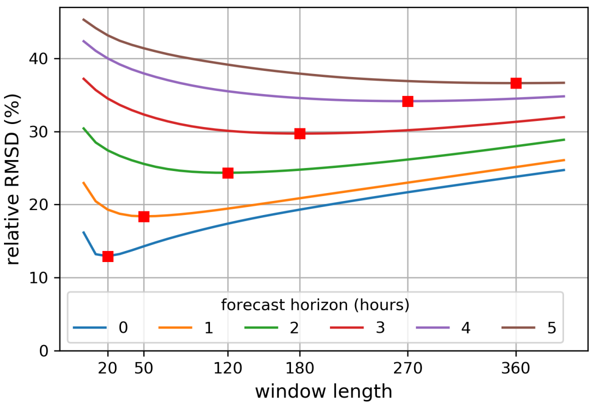
AppendixD. Spatial smoothing plots for each method



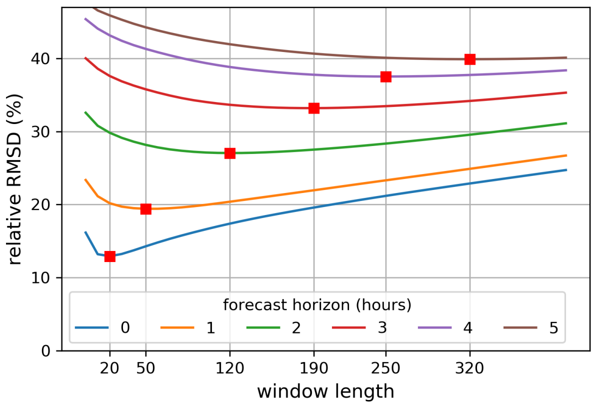
(a) LK-afn algorithm.



(b) FRB algorithm.



(c) HS algorithm.



(d) PIV algorithm.

Figure D.13: Effect of the image spatial smoothing on the relative RMSD metric at irradiation level for the different CMV techniques. This plot for the best performing technique is shown in Figure 7.

References

Abal, G., Monetta, A., and Alonso-Suárez, R. (2018). Outdoor solar radiometer calibration under ISO-9847:1992 standard and alternative methods. In *2018 IEEE 9th Power, Instrumentation and Measurement Meeting (EPIM)*, pages 1–6.

Adrian, R. (1991). Particle-imaging techniques for experimental fluid mechanics. *Annual Review of Fluid Mechanics*, 23(1):261–304.

Alonso-Suárez, R. (2017). *Estimación del recurso solar en Uruguay mediante imágenes satelitales.* PhD thesis, Facultad de Ingeniería, Universidad de la República. Available at: https://hdl.handle.net/20.500.12008/20200.

Alonso-Suárez, R., Abal, G., Siri, R., and Musé, P. (2012). Brightness-dependent tarpley model for global solar radiation estimation using GOES satellite images: application to Uruguay. *Solar Energy*, 86(11):3205–3215.

Alonso-Suárez, R., David, M., Branco, V., and Lauret, P. (2020). Intra-day solar probabilistic forecasts including local short-term variability and satellite information. *Renewable Energy*, 158:554–573.

Alonso-Suárez, R., Marchesoni, F., and Aicardi, D. (2021). Analysis of persistence-based solar irradiance forecasting benchmarks. In *Proceedings of the ISES Solar World Congress 2021, extended abstract.*

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., de Pison, F. M., and Antonanzas-Torres, F. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136:78–111.

Arbizu-Barrena, C., Ruiz-Arias, J., Rodríguez-Benítez, F., Pozo-Vázquez, D., and Tovar-Pescador, J. (2017). Short-term solar radiation forecasting by advecting and diffusing MSG cloud index. *Solar Energy*, 155:1092–1103.

Black, M. J. and Anandan, P. (1996). The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104.

Bouguet, J.-Y. (2000). Pyramidal implementation of the lucas kanade feature tracker. *Intel Corporation, Microprocessor Research Labs.*

Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1):89–97.

Cros, S., Badosa, J., Szantaï, A., and Haeffelin, M. (2020). Reliability predictors for solar irradiance satellite-based forecast. *Energies*, 13(21).

Cros, S., Sébastien, N., Liandrat, O., and Schmutz, N. (2014). Cloud pattern prediction from geostationary meteorological satellite images for solar energy forecasting. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 9242.

Diagne, M., David, M., Lauret, P., Boland, J., and Schmutz, N. (2013). Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews*, 27:65–76.

Dise, S., Kankiewicz, A., Schlemmer, J., Hemker, K., Kivalov, S., Hoff, T., and Perez, R. (2013). Operational improvements in the performance of the SUNY satellite-to-solar irradiance model using satellite infrared channels. In *2013 IEEE 39th Photovoltaic Specialists Conference (PVSC)*, pages 0960–0963.

Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. In Bigun, J. and Gustavsson, T., editors, *Image Analysis*, pages 363–370, Berlin, Heidelberg. Springer Berlin Heidelberg.

Giacosa, G. and Alonso-Suárez, R. (2020). Performance assessment of intra-day solar irradiation forecast in Uruguay using satellite Cloud Motion Vectors. In *Proceedings of the ISES Solar World Congress 2019 and IEA SHC International Conference on Solar Heating and Cooling for Buildings and Industry 2019*, pages 2083–2090.

Harty, T., Holmgren, W., Lorenzo, A., and Morzfeld, M. (2019). Intra-hour cloud index forecasting with data assimilation. *Solar Energy*, 185:270–282.

Horn, B. K. P. and Schunck, B. G. (1981). Determining optical flow. *Artif. Intell.*, 17(1-3):185–203.

Kallio-Myers, V., Riihelä, A., Lahtinen, P., and Lindfors, A. (2020). Global horizontal irradiance forecast for Finland based on geostationary weather satellite data. *Solar Energy*, 198:68–80.

Kühnert, J., Lorenz, E., and Heinemann, D. (2013). Chapter 11 - satellite-based irradiance and power forecasting for the german energy market. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 267–297. Academic Press, Boston.

Laguarda, A. and Abal, G. (2016). Índice de turbidez de Linke a partir de irradiación solar global en el Uruguay. *Aceptado para Avances en Energías Renovables y Medio Ambiente*, pages 1–14.

Laguarda, A., Giacosa, G., Alonso-Suárez, R., and Abal, G. (2020). Performance of the site-adapted cams database and locally adjusted cloud index models for estimating global solar horizontal irradiation over the pampa húmeda. *Solar Energy*, 199:295–307.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Qu, Z., Wald, L., Homscheidt, M. S., and Arola, A. (2013). McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques, European Geosciences Union*, 6:2403–2418.

Lorenz, E., Hammer, A., Heinemann, D., et al. (2004). Short term forecasting of solar radiation based on satellite data. In *EUROSUN2004 (ISES Europe Solar Congress)*, pages 841–848.

Lucas, B. D. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'81, pages 674–679, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Marchesoni-Acland, F. and Alonso-Suárez, R. (2020). Intra-day solar irradiation forecast using rls filters and satellite images. *Renewable Energy*, 161:1140–1154.

McArthur, L. (2005). Baseline Surface Radiation Network (BSRN) Operations Manual. Td-no. 1274, wrcp/wmo, World Meteorological Organization (WMO). www.wmo.org.

Miller, S. D., Rogers, M. A., Haynes, J. M., Sengupta, M., and Heidinger, A. K. (2018). Short-term solar irradiance forecasting via satellite/model coupling. *Solar Energy*, 168:102–117. Advances in Solar Resource Assessment and Forecasting.

Nonnenmacher, L. and Coimbra, C. F. (2014). Streamline-based method for intra-day solar forecasting through remote sensing. *Solar Energy*, 108:447–459.

Pal, S., Dominguez, F., Dillon, M. E., Alvarez, J., Garcia, C. M., Nesbitt, S. W., and Gochis, D. (01 Feb. 2021). Hydrometeorological observations and modeling of an extreme rainfall event using wrf and wrf-hydro during the relampago field campaign in argentina. *Journal of Hydrometeorology*, 22(2):331–351.

Peel, M. C., Finlayson, B. L., and Mcmahon, T. A. (2007). Updated world map of the köppen-geiger climate classification. *Hydrology and Earth System Sciences Discussions*, 11:1633–1644.

Peng, Z., Yoo, S., Yu, D., and Huang, D. (2013). Solar irradiance forecast system based on geostationary satellite. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 708–713.

Pereira, E. S., Gonçalves, A. R., Costa, R., and E., P. (2020). Algoritmo de geraçao de imagens de previsão de campos de nebulosidade a partir de imagens de satélite. In *Anais do VIII Congresso Brasileiro de Energia Solar*.

Perez, R., Cebecauer, T., and Šúri, M. (2013). Chapter 2 - semi-empirical satellite models. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 21–48. Academic Press, Boston.

Perez, R., David, M., Hoff, T. E., Jamaly, M., Kivalov, S., Kleissl, J., Lauret, P., and Perez, M. (2016). Spatial and temporal variability of solar energy. *Foundations and Trends® in Renewable Energy*, 1(1):1–44.

Perez, R. and Hoff, T. E. (2013). Chapter 10 - SolarAnywhere forecasting. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 233–265. Academic Press, Boston.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F. (2002). A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, 73(5):307–317.

Perez, R., Kankiewicz, A., Schlemmer, J., Hemker, K., and Kivalov, S. (2014). A new operational solar resource forecast model service for pv fleet simulation. In *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, pages 0069–0074.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., and Hoff, T. E. (2010). Validation of short and medium term operational solar radiation forecasts in the US. *Solar Energy*, 84(12):2161–2172.

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., and Wald, L. (2017). Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. *Meteorologische Zeitschrift*, 26(1):33–57.

Rasmussen, K. L., Zuluaga, M. D., and Houze Jr., R. A. (2014). Severe convection and lightning in subtropical south america. *Geophysical Research Letters*, 41(20):7359–7366.

Rigollier, C., Bauer, O., and Wald, L. (2000). On the clear sky model of the ESRA –European Solar Radiation Atlas– with respect to the Heliosat method. *Solar Energy*, 68(1):33–48.

Rigollier, C., Lefevre, M., and Wald, L. (2004). The method Heliosat-2 for deriving shortwave solar radiation from satellite images. *Solar Energy*, 77(2):159–169.

Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268.

Salio, P., Nicolini, M., and Zipser, E. J. (01 Apr. 2007). Mesoscale convective systems over southeastern south america and their relationship with the south american low-level jet. *Monthly Weather Review*, 135(4):1290–1309.

Sun, D., Roth, S., Lewis, J. P., and Black, M. J. (2008). Learning optical flow. In Forsyth, D., Torr, P., and Zisserman, A., editors, *Computer Vision – ECCV 2008*, pages 83–97, Berlin, Heidelberg. Springer Berlin Heidelberg.

Sánchez, J., Meinhardt-Llopis, E., and Facciolo, G. (2013). TV-L1 optical flow estimation. *Image Processing Online (IPOL).*

Sánchez Pérez, J., Monzón López, N., and Salgado de la Nuez, A. (2013). Robust Optical Flow Estimation. *Image Processing On Line*, 3:252–270.

Urbich, I., Bendix, J., and Müller, R. (2019). The seamless solar radiation (sesora) forecast for solar surface irradiance—method and validation. *Remote Sensing*, 11(21).

Wang, P., van Westrhenen, R., Meirink, J., van der Veen, S., and Knap, W. (2019). Surface solar radiation forecasts by advecting cloud physical properties derived from meteosat second generation observations. *Solar Energy*, 177:47–58.

Wedel, A. and Cremers, D. (2011). *Stereo Scene Flow for 3D Motion Analysis.* Springer Publishing Company, 1st edition.

Wu, X. and Sun, F. (2005). Post-launch calibration of GOES imager visible channel using MODIS. In *Proc. SPIE*, volume 5882.

Yang, D. (2019a). A guideline to solar forecasting research practice: Reproducible, operational, probabilistic or physically-based, ensemble, and skill (ropes). *Journal of Renewable and Sustainable Energy*, 11(2):022701.

Yang, D. (2019b). Making reference solar forecasts with climatology, persistence, and their optimal convex combination. *Solar Energy*, 193:981–985.

Yang, D., Alessandrini, S., Antonanzas, J., Antonanzas-Torres, F., Badescu, V., Beyer, H. G., Blaga, R., Boland, J., Bright, J. M., Coimbra, C. F., David, M., Âzeddine Frimane, Gueymard, C. A., Hong, T., Kay, M. J., Killinger, S., Kleissl, J., Lauret, P., Lorenz, E., van der Meer, D., Paulescu, M., Perez, R., Perpiñán-Lamigueiro, O., Peters, I. M., Reikard, G., Renné, D., Saint-Drenan, Y.-M., Shuai, Y., Urraca, R., Verbois, H., Vignola, F., Voyant, C., and Zhang, J. (2020a). Verification of deterministic solar forecasts. *Solar Energy*, 210:20–37. Special Issue on Grid Integration.

Yang, L., Gao, X., Hua, J., Wu, P., Li, Z., and Jia, D. (2020b). Very short-term surface solar irradiance forecasting based on Fengyun-4 geostationary satellite. *Sensors (Switzerland)*, 20(9).

Zach, C., Pock, T., and Bischof, H. (2007). A duality based approach for realtime TV-L1 optical flow. In *Proceedings of the 29th DAGM Conference on Pattern Recognition*, pages 214–223, Berlin, Heidelberg. Springer-Verlag.