



## Data Article

# Girasol, a sky imaging and global solar irradiance dataset



Guillermo Terrén-Serrano<sup>a,\*</sup>, Adnan Bashir<sup>b</sup>, Trilce Estrada<sup>b</sup>,  
Manel Martínez-Ramón<sup>a</sup>

<sup>a</sup> Department of Electrical and Computer Engineering, The University of New Mexico, Albuquerque, NM 87131, United States

<sup>b</sup> Department of Computer Science, The University of New Mexico, Albuquerque, NM 87131, United States

## ARTICLE INFO

## Article history:

Received 31 December 2020

Revised 21 February 2021

Accepted 23 February 2021

Available online 4 March 2021

## Keywords:

Sky Imaging  
Global Solar Irradiance  
Fisheye Lens Camera  
Long-wave Infrared Camera  
Data Acquisition System  
Solar Forecasting  
Smart Grids  
Sun-Tracking

## ABSTRACT

The energy available in a microgrid that is powered by solar energy is tightly related to the weather conditions at the moment of generation. A very short-term forecast of solar irradiance provides the microgrid with the capability of automatically controlling the dispatch of energy. We propose a dataset to forecast Global Solar Irradiance (GSI) using a data acquisition system (DAQ) that simultaneously records sky imaging and GSI measurements, with the objective of extracting features from clouds and use them to forecast the power produced by a Photovoltaic (PV) system. The DAQ system is nicknamed the *Girasol Machine* (*Girasol* means Sunflower in Spanish). The sky imaging system consists of a longwave infrared (IR) camera and a visible (VI) light camera with a fisheye lens attached to it. The cameras are installed inside a weatherproof enclosure that it is mounted on a solar tracker. The tracker updates its pan and tilt every second using a solar position algorithm to maintain the Sun in the center of the IR and VI images. A pyranometer is situated on a horizontal mount next to the DAQ system to measure GSI. The dataset, composed of IR images, VI images, GSI measurements, and the Sun's positions, has been tagged with timestamps.

DOI of original article: [10.1016/j.apenergy.2021.116656](https://doi.org/10.1016/j.apenergy.2021.116656)

\* Corresponding author: Tel.: +1-505-450-6215

E-mail address: [guillermoterreren@unm.com](mailto:guillermoterreren@unm.com) (G. Terrén-Serrano).

Social media: (M. Martínez-Ramón)

<https://doi.org/10.1016/j.dib.2021.106914>

2352-3409/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## Specifications Table

Subject	Energy Engineering and Power
Specific subject area	Artificial intelligence system to forecast solar energy in Micro Grids and Smart Grids powered by photovoltaic technology [1].
Type of data	Image (.png), Table (.csv)
How data were acquired	Instruments: Infrared Camera, Visible Camera, Solar Tracker, Pyranometer sensor, Raspberry Pi, Motherboard Make and model and of the instruments used: FLIR <sup>©</sup> Lepton 2.5 infrared camera mounted on Pure Thermal 1 board manufactured by Group Gets. ELP <sup>©</sup> visible camera model 8541707515. Tracker FLIR <sup>©</sup> model PTU-E46-70. LI-COR <sup>©</sup> Pyranometer sensor model LI-200. Raspberry Pi 2B manufactured by SONY <sup>©</sup> . ASRock <sup>©</sup> model Syper Alloy J3455-ITX Quad-core.
Data format	Raw
Parameters for data collection	The sampling interval of the cameras is 15 seconds when the Sun's elevation angles is > 15°. There are approximately 1,200 to 2,400 captures per day from each camera depending of the year day. The GSI signal has a sampling rate ranging from 4 to 6 samples per second. The Sun position files are generate out of the time recorded in the pyranometer files. The weather station data is updated every 10 minutes.
Description of data collection	The tracking system and DAQ software is currently operative. The software was programmed in Python 2.7. The system is placed on the top roof area of the Mechanical Engineering building in UNM central campus. The DAQ session can be visually monitored through a webpage. All devices are interconnected via a LAN network built by DHCP server. The weather station is located at the University of New Mexico Hospital, and both its real-time and historical data are publicly accessible <sup>1</sup> .
Data source location	Institution: The University of New Mexico City/Town/Region: Albuquerque, New Mexico Country: United States of America Latitude and longitude for collected samples: 35.0821, -106.6259
Data accessibility	Repository name: Girasol, a Sky Imaging and Global Solar Irradiance Dataset Data identification number: <a href="https://doi.org/10.5061/dryad.zcrjdfn9m">https://doi.org/10.5061/dryad.zcrjdfn9m</a> Direct URL to data and instructions for accessing these data: <a href="https://doi.org/10.5061/dryad.zcrjdfn9m">https://doi.org/10.5061/dryad.zcrjdfn9m</a>
Related research article	G. Terrén-Serrano, M. Martínez-Ramón, Multi-Layer Wind Velocity Field Visualization in Infrared Images of Clouds for Solar Irradiance Forecasting, Applied Energy, 288 (2021) 116656. <a href="https://doi.org/10.1016/j.apenergy.2021.116656">https://doi.org/10.1016/j.apenergy.2021.116656</a>

## Value of the Data

- This dataset was generated using a sky imaging DAQ system for solar forecasting [2]. This system is innovative in that it is mounted on a solar tracker so the Sun stays in the center of the images, and that it combines ground-based IR and VI light fisheye images [3]. IR cameras have been previously used but not mounted on a solar tracker [4].
- This dataset was originally conceived to address the problems related to very short-term forecasting of solar irradiance [5]. In very short-term forecasting of solar irradiance the cloud information in the images is useful to accurately predict when the Sun may be occluded by a cloud. In these events, a shadow is projected over a PV system producing a drop in the energy supply [6,7]. A GSI forecasting algorithm will provide the grid with the capability of managing the energy resources [8].

<sup>1</sup> <https://www.wunderground.com/dashboard/pws/KNMALBUQ473>.

- This dataset includes images captured using two different light sensors equipped with lenses. The information provided by the sensors may be combined to increase the diversity of features extracted from the clouds [9]. Image processing is an important factor in the performance of a solar forecasting algorithm [10].
- In the context of machine learning and image processing, the feature extraction algorithm may be adapted to different applications in solar problems [11,12]. For instance, the solar forecasting horizon can be changed depending on the application. The dataset can be used to forecast the effect of the clouds in thermosolar energy generation systems such as concentrated solar power [13,14].

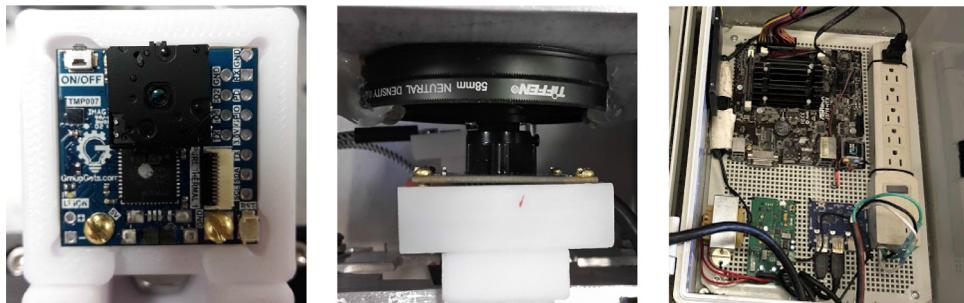
## 1. Data Description

The repository contains recordings of the solar cycle from 242 days of 3 years. The total amount of data is 119GB. The sampling interval of the cameras is 15 seconds and the observation period is when the Sun's elevation angle is higher than 15°. There are approximately 1,200 to 2,400 captures per day from each camera depending on the day of the year.

- VI images are 16 bits with a resolution of  $450 \times 450$ , intensity channel only. Approx. 240KB per frame. Between 200 MB to 400 MB per day depending on the amount of images in the directory. The images are saved in a lossless png format in the directory \*/visible. The images are named by the UNIX time in seconds.
- IR images are 16 bits with a resolution of  $60 \times 80$ . Approx. 8KB per frame. 20MB per day roughly constant. Images are saved in lossless png format in the directory \*/infrared. The images are named by the UNIX time in seconds.
- The pyranometer is sampled from 4 to 6 times per second. The measurements are saved in the directory \*/pyranometer in.csv files named with their date (yyyy\_mm\_dd), approx. 4,500KB to 7,500KB per file. The files contain UNIX time in the first column and GSI in  $W/m^2$  in the second column.
- The Sun position files are generated from the time recorded in the pyranometer files. The positions are saved in the directory \*/sun\_position in.csv files named with their date (yyyy\_mm\_dd), approx. 6,500KB to 11,500KB per file. The files contain the UNIX time in the first column, the elevation angle in the second column, and the azimuth angle in the third column, all of them computed with full precision UNIX time.
- The weather station sample interval is 10 minutes. Linear interpolation was applied to match the sampling interval of the pyranometer. The weather station files are saved in the directory \*/weather\_station in.csv files named with their date (yyyy\_mm\_dd), approx. 14.3MB to 25.4MB per file. The files are organized by columns which contain, from left to right: UNIX time, temperature in °C, dew point in °C, atmospheric pressure in mmHg, wind direction in radians, wind velocity in miles/s and relative humidity in %.

## 2. Experimental Design, Materials and Methods

The *Girasol Machine* is composed of data acquisition hardware and software. In this section, we first summarize the specifications of the hardware. The described parts of the hardware are: the IR camera, the VI camera, the solar tracker and the pyranometer. We later describe the formulation of the algorithms in the software. These algorithms are designed to compute the Sun's position, to attenuate the noise in IR and VI images, and to fuse multiple exposure VI images together.



**Fig. 1.** Details of IR (left) and VI (middle) cameras screwed onto their individual 3D printed support, inside the custom made enclosure, and placed in front of their respective apertures. The supports are fixed to the enclosure's structured by means of adjustable clamps. The junction box (right) contains the computer and controllers.



**Fig. 2.** Germanium and neutral density lenses are attached to the IR and VI cameras, respectively, to filter the light beams and protect the cameras from weather hazards (left). The solar tracker's servomotors are mounted on the tripod and screwed to a metallic structure which counterbalances the enclosure's weight (middle). The sky-imaging system (right) installed on the roof of a UNM building.

## 2.1. Data Acquisition Hardware

We built a system composed of VI and IR solar radiation cameras, a tracking system, and a pyranometer, to collect the data. A weatherproof enclosure contains the two USB cameras, and it is mounted on top of two servomotors. The cameras are connected to a motherboard running a solar tracking algorithm in parallel. This motherboard is placed inside of a different weatherproof enclosure together with a router, a servomotor control unit, a data storage system, and the respective power supplies. The enclosure's degree of protection is IP66. This is an international standard utilized for electronic equipment that provides protection against dust and water (see Figs 1–2).

### 2.1.1. Infrared Sensor

The IR sensor used to capture the images is a FLIR<sup>©</sup> Lepton 2.5 camera<sup>2</sup> which is mounted on a Pure Thermal 1 board<sup>3</sup> manufactured and distributed by Group Gets. The Lepton 2.5 sensor produces thermal images by measuring long-wave infrared (see Fig. 1). It captures IR radiation with a nominal response wavelength band from 8 to 14  $\mu\text{m}$ .

<sup>2</sup> <http://www.flir.com>.

<sup>3</sup> <https://groupgets.com>.

The dimensions of a Lepton 2.5 are  $8.5 \times 11.7 \times 5.6$  mm. It has  $51^\circ$  horizontal Field Of View (FOV) and  $63.5^\circ$  diagonal FOV with type f1.1 silicon doublet lens. The resolution is 80 (horizontal)  $\times$  60 (vertical) active pixels. The thermal sensitivity is  $< 50$  mK. The camera integrates digital thermal image processing functions that include automatic thermal environment compensation, noise filtering, non-uniformity correction and gain control. This sensor can produce or stream an image in  $< 0.5$  seconds. It operates at 150 mW nominal power, and has a low power standby mode. The Pure Thermal 1 board adds functionalities to the IR sensor such as thermal video transmission over USB, which works with a USB Video Class (UVC) library on Windows, Linux, Mac and Android. It also includes on-board image processing without the need of an external system. The microprocessor is an STM32F411CEU6 ARM. It also includes a contactless thermopile temperature sensor to manually calibrate a Lepton module. The operational temperatures for Flat Field Correction (FFC) function range from  $-10^\circ\text{C}$  to  $65^\circ\text{C}$ . The captured images are digitized in Y16 format. It is a single channel format that only quantifies intensity levels.

### 2.1.2. Visible Sensor

The VI camera is a 5 megapixel color sensor with USB 2.0 manufactured by ELP<sup>©</sup>. The sensor is an OV5640 with maximum resolution of 2592 (horizontal)  $\times$  1944 (vertical) pixels and  $170^\circ$  FOV with a fisheye type lens that is adjustable within the range of 2.1 to 6 mm (see Fig. 1). The pixel size is  $1.4 \times 1.4 \mu\text{m}$ , and the image area is  $3673.6 \times 2738.4 \mu\text{m}$ . The image stream rate is 30 frames per second at  $640 \times 480$  pixel resolution. The communication protocol is UVC and its interface is a USB 2.0 high speed. The dynamic range is 68 dB, and it has a mount-in shutter that can control the frame exposure time. The camera-board has built-in functions that can be enabled for Automatic Gain Control (AGC), Automatic Exposure Control (AEC), Automatic White Balance (AWB) and Automatic Black Focus (ABF). The sensor's brightness, contrast, hue, saturation, sharpness, gamma, white balance, exposure and focus are software adjustable. The power consumption in VGA resolution is 150mW. The dimensions of the camera board are 38 x 38 mm. The recommended operational temperatures for stable images range from  $0^\circ\text{C}$  to  $60^\circ\text{C}$ . The output format of the camera is YUYV. YUYV is a 3 channel format in the YUV color space, where Y represents brightness, U blue color projection and V red color projection. UVC drivers can operate in Windows, Linux, MAC and Android.

### 2.1.3. Solar Tracker

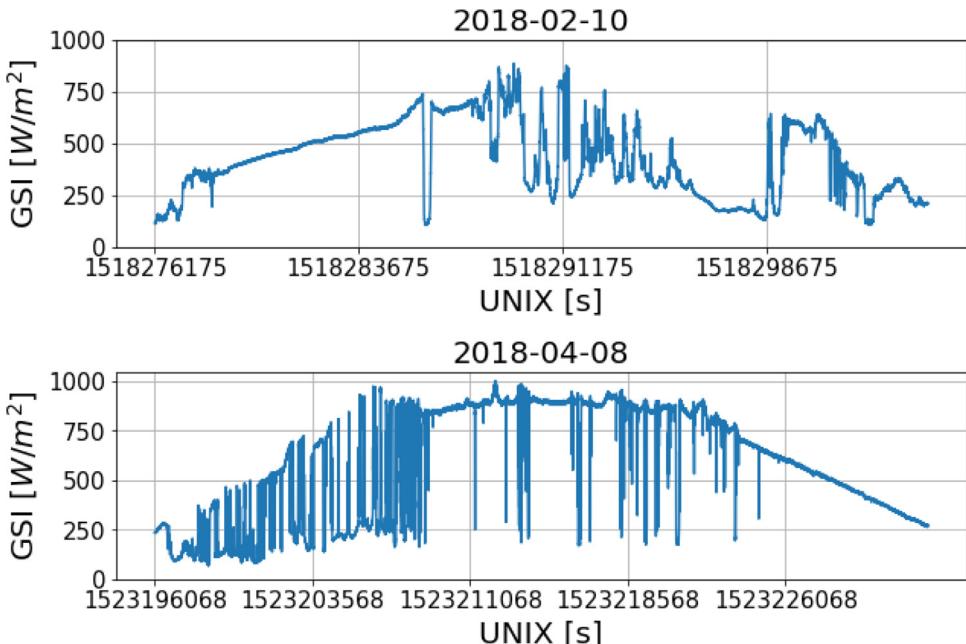
The device used for solar tracking is a commercial model manufacture by FLIR<sup>©</sup> and distributed by moviTHERM (Fig. 2). The model is a PTU-E46-70. The tracking system is driven by two servomotors mounted to allow rotation on tilt and pan axes. The tracker is rated for a payload of 4.08 kg, (the maximum weight of the system). The rotational speed of the tracker is  $60^\circ$  per second, and it has a resolution of  $0.003^\circ$ .

The degree of freedom of the tilt servomotor allows vertical maneuvers of  $78^\circ$  with a vertical limit ranging from  $-47^\circ$  to  $+31^\circ$  with respect to the horizontal axis. The pan servomotor has more degrees of freedom. This allows the Sun to be tracked throughout the entire day. Its range is  $\pm 159^\circ$ . The motor speeds are adjustable up to  $0.003^\circ$  per second. The tracker requires an input voltage ranging from 12 to 30 VDC. The corresponding power is 13 W in maximum power mode, 6 W low-power mode and 1 W in holding power off mode. The system is controlled via Ethernet from a host computer. The control unit is connected to the servomotors via a DB-9 female connector.

Regarding the mechanical description of the tracker, the system weight is 1.36 kg, its dimensions are 7.6 cm height  $\times$  13 cm width  $\times$  10.17 cm depth. The control unit weights 0.227 kg, and its dimensions are 3.18 cm height  $\times$  8.26 cm width  $\times$  11.43 cm depth.

### 2.1.4. Pyranometer

As part of the DAQ, a pyranometer sensor has been designed and tested to provide the device with high portability and accuracy. Here, the circuit design for signal conditioning of the pyranometer sensor (model LI-COR LI-200) is summarized. The prototype was designed to



**Fig. 3.** This figure shows two days of samples from the pyranometer. Day 2018-02-10 has 101510 samples and day 2018-04-08 has 131412 samples. The x axis shows the UNIX time in seconds. The y axis shows the GSI in  $W/m^2$ .

send data to a laptop, Raspberry Pi or any other computer through USB communication, to which a simple Python script needs to be added and executed. For portability purposes, the board we designed is provided with a 5 V power supply, so that a Raspberry Pi may also be connected to it.

The pyranometer design objective was to condition and measure the signal of the sensor by converting it to a 12-bit digital signal and saving the data by timestamp in a.csv file. The system is portable, easy to setup, and reliable. The pyranometer output voltage ranges from 0 to 20 mV, where 10 mV represents a radiation power density of 1,000  $W/m^2$  (see Fig. 3). A 120 V power outlet is used to power the device. The system is connected to the Internet and has user log-in capabilities.

Our design uses an Analog Devices AD629 instrumentation amplifier with 13 dB gain that conditions the signal to an ADC. An Atmel SAMD21G18 microcontroller is used to measure the signal output from the ADC and to transmit it by a USB port. The board is connected to a Raspberry Pi to save the signal's data in files. This Raspberry Pi is connected to the internet, so it is also able to transmit the data.

## 2.2. Data Acquisition Software

The tracking system and DAQ software is currently operative. The software was programmed in a single Python 2.7 script. The system is placed on the roof of the Electrical and Computer Engineering building at UNM central campus (35.0821, -106.6259). The DAQ sessions can be visually monitored through a webpage. All devices were interconnected via a LAN network built by a DHCP server.

### 2.2.1. Sun Position

The Sun position is computed to update the servomotors' position every second. The following set of equations was used to calculate azimuth and elevation angles<sup>4</sup>. These angles correspond to the servomotors' pan and tilt axes. The tracking system must be aligned to the true North or South of the local geographic coordinates defined as  $(\lambda, \phi)$ , which are respectively longitude and latitude. The system must be leveled with respect to the horizontal and vertical axis. The first step is to calculate the Local Standard Time Meridian (LSTM), which is the time zone meridian with respect to the Greenwich meridian. The LSTM is computed as,

$$\text{LSTM} = 15^\circ \cdot \Delta T_{GMT}, \quad (1)$$

where  $\Delta T_{GMT}$  is the difference between the Local Time (LT) and Greenwich Mean Time (GMT) and it is measured in hours. The equation of time (EoT) is an empirical equation that considers the eccentricity of earth's orbit and its axial tilt. It is expressed in minutes and is described by

$$\text{EoT} = 9.87 \cdot \sin(2B) - 7.53 \cdot \cos(G) - 1.5 \cdot \sin(B), \quad (2)$$

where

$$B = \frac{360}{365} \cdot (d - 81). \quad (3)$$

The units of B are degrees and  $d$  corresponds to the number of days since the beginning of the year. Time Correction factor (TC) accounts for local effects that produce variations in the Local Solar Time(LST).These effects are caused by the eccentricity of the Earth's orbit. TC is quantified in minutes, and its equation is

$$\text{TC} = 4 \cdot (\lambda - \text{LSTM}) + \text{EoT}, \quad (4)$$

Constant 4 is related to the rotation of the earth, which rotates  $1^\circ$  every 4 minutes, and  $\lambda$  is the longitude. The LST is calculated by using the previous correction and Local Time (LT). The following equation is expressed in hours

$$\text{LST} = \text{LT} + \frac{\text{TC}}{60}. \quad (5)$$

The Earth rotates  $15^\circ$  per hour, so each hour away from the solar noon corresponds to an angular position. The Hour Angle (HRA) converts LST to the Sun's angle. Before noon, angles are negative and in the afternoon the angles obtained are positive. The expression for this is

$$\text{HRA} = 15^\circ \cdot (\text{LST} - 12). \quad (6)$$

The Sun's declination angle  $\delta$  varies seasonally, and it accounts for the angular difference between the Equator and the Earth's center.  $\delta$  is obtained from formula

$$\delta = 23.45^\circ \cdot \sin \left[ \frac{360}{365} \cdot (d - 8) \right], \quad (7)$$

where  $d$  represents the day since the beginning of the year. The rotations on the tracker's pan and tilt axes are provided by the Elevation  $\epsilon$  and Azimuth  $\alpha$  angles. The Elevation angle quantifies the Sun's angular height, and the azimuth  $\alpha$  is the Sun's angular position on the horizon. The angles are calculated from the following equations,

$$\begin{aligned} \epsilon &= \sin^{-1} [\sin \delta \cdot \sin \phi + \delta \cdot \cos \phi \cdot \cos(\text{HRA})], \\ \alpha &= \cos^{-1} \left[ \frac{\sin \delta \cdot \cos \phi - \cos \delta \cdot \sin \phi \cdot \cos(\text{HRA})}{\cos \xi} \right], \end{aligned} \quad (8)$$

where HRA is the hour angle,  $\phi$  is latitude, and

$$\xi = \sin^{-1} [\sin \delta \cdot \sin \phi + \delta \cdot \cos \phi \cdot \cos(\text{HRA})]. \quad (9)$$

<sup>4</sup> <http://www.pveducation.org/>.

Alternatively,  $\xi$  can be also calculated as  $\xi = 90^\circ + \phi - \delta$ . The zenith angle  $\zeta$  is the difference between the elevation angle  $\epsilon$  and the vertical axis, so that

$$\zeta = 90^\circ - \epsilon. \quad (10)$$

The tracking system initializes one hour after the sunrise and stops one hour before sunset, so that it is inactive during the hours when the Sun's elevation is too low to generate energy by PV systems. Inverters require a minimum amount of power generation, which is reached when Sun's position is above and below these time limits. The sunrise (SR) and sunset (SS) are expressed in decimal hours and they are calculated in the following equations

$$\begin{aligned} \text{SR} &= 12 - \frac{1}{15^\circ} \cdot \cos^{-1} \left( \frac{-\sin \phi \cdot \sin \delta}{\cos \phi \cdot \cos \delta} \right) - \frac{\text{TC}}{60}, \\ \text{SS} &= 12 + \frac{1}{15^\circ} \cdot \cos^{-1} \left( \frac{-\sin \phi \cdot \sin \delta}{\cos \phi \cdot \cos \delta} \right) - \frac{\text{TC}}{60}. \end{aligned} \quad (11)$$

### 2.2.2. Noise Attenuation

The captures from the IR cameras are noisy. Therefore, we propose to attenuate the noise by averaging  $N = 10$  frames taken at capture  $k$ , with a frame rate of 9 images per second. The camera frame rates are lower than the computer socket reading speed, so we need to assess whether each acquired frame read from the buffer is new or if it has been previously acquired in our set. Also, it is possible that a capture was defective. To detect these situations, we first acquire  $N$  consecutive images and we compute the sample mean image of the set  $\bar{\mathbf{I}} = \frac{1}{N} \cdot \sum_{i=1}^N \mathbf{I}_i$ . Then, for all images in the set, we compute the Pearson coefficient

$$\rho_{i,j} = \frac{(\mathbf{I}_i - \bar{\mathbf{I}}) \cdot (\mathbf{I}_j - \bar{\mathbf{I}})}{\sqrt{(\mathbf{I}_i - \bar{\mathbf{I}})^2} \cdot \sqrt{(\mathbf{I}_j - \bar{\mathbf{I}})^2}}, \quad \forall i, j = 1, \dots, N. \quad (12)$$

If a pair of frames has a coefficient equal to one, we discard one of them since the frames are the same. If an image has a coefficient less than 0.9 with respect the rest, we assume that the frame is defective, and it is also discarded. The mean is recomputed with the remaining frames and then stored (see Fig. 4).

### 2.2.3. Exposure Switching

The scattering effect produced by solar irradiance in the VI images precludes the detection of the clouds in the circumsolar region (when the exposure time is long). The light in the outer region of the image is faint when the exposure time is short. To resolve this issue, we propose to average the captured images at different exposure times. A frame is taken with four different exposure times equal to 1 ms, 4 ms, 12 ms and 28 ms. The process is repeated 10 times, and the frames with same exposure time are averaged using the procedure in Subsection 4.2.2, which results in 4 different frames corresponding to the four different exposures. This images are further fused using the procedure below.

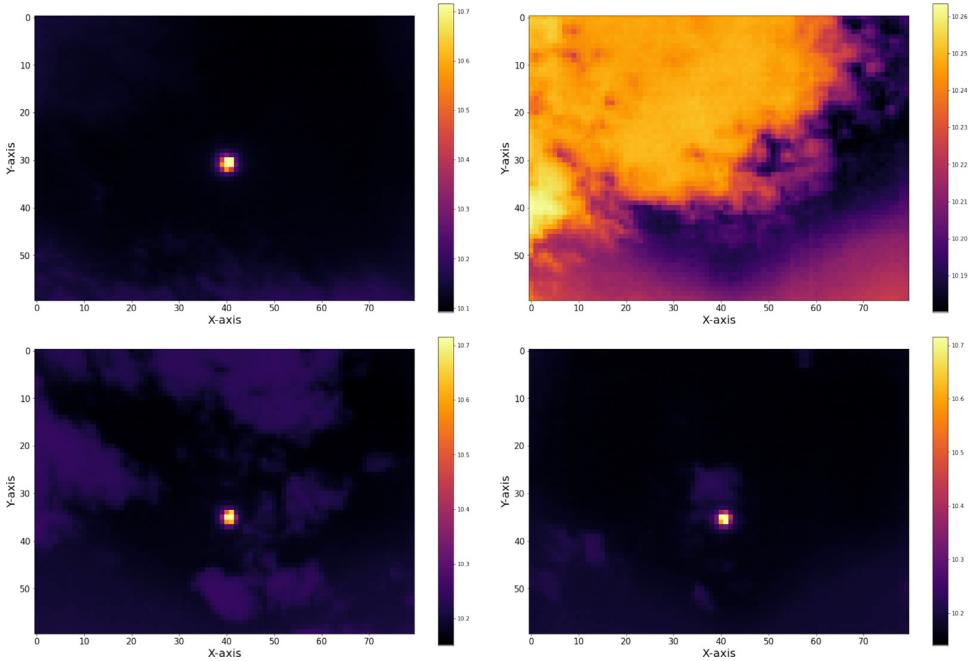
### 2.2.4. Visible Image Fusion

We introduce a method to fuse images with different exposure times, generating a High Dynamic Range (HDR) of 16 bits, and combining cloud information from the different images. The frame is defined as  $\mathbf{I}_{e,c}^k$ ,  $1 \leq e \leq 4$  as each one of the RGB components  $1 \leq c \leq 3$  of image  $\mathbf{I}_e^k$  at instant  $k$  and exposure  $T_e \in \{1, 4, 12, 28\}$  ms. To merge the images we first regularized the RGB components in each exposition to avoid division by zero when light is too dim,

$$\mathbf{I}_{e,c}^k + \lambda, \quad \mathbf{I}_{e,c}^k \in \mathbb{R}^{D \times D}. \quad (13)$$

We convert the images to grayscale by a weighted sum of the image RGB channels with the corresponding *Luma* coding system coefficients  $\beta_c = \{0.299, 0.587, 0.114\}$  [15],

$$\mathbf{I}_e^k = \sum_{c=1}^C \mathbf{I}_{e,c}^k \cdot \beta_c \quad (14)$$



**Fig. 4.** This figure shows samples (stored in the repository) of IR images that were acquired in different days. The intensity of the pixels in the images is displayed in logarithmic scale. The resolution of the images is  $60 \times 80$  and their sample rate is four images per minute.

The proposed method is based on the distance from a pixel to the Sun. We know that the Sun is always centered in the frames. Therefore, we define our fusion mask as a binary valued frame function of the radial distance of the Sun to a pixel, such as

$$\mathcal{M}(r) \triangleq \left\{ \mathbf{M}_{i,j} = \mathbb{I}\left( (x_i - x_{\text{sun}})^2 + (y_j - y_{\text{sun}})^2 \leq r^2 \right) \right\}, \quad (15)$$

where  $x_i$  and  $y_j$  are coordinates of a frame and  $\{x_{\text{sun}}, y_{\text{sun}}\}$  are the coordinates of the Sun. We define a mask  $\mathbf{M}_e$  for each exposition time corresponding to a radius,

$$\mathbf{M}_e = \mathcal{M}(r_e), \quad \forall r_e = \{5, 6.25, 12.5, 25\} \quad (16)$$

the set of fusion masks is then  $\mathbf{M}_e = \{\mathbf{0}, \mathbf{M}_2, \dots, \mathbf{M}_E, \mathbf{1}\}$ . We convolve each one of the masks with a Gaussian kernel,

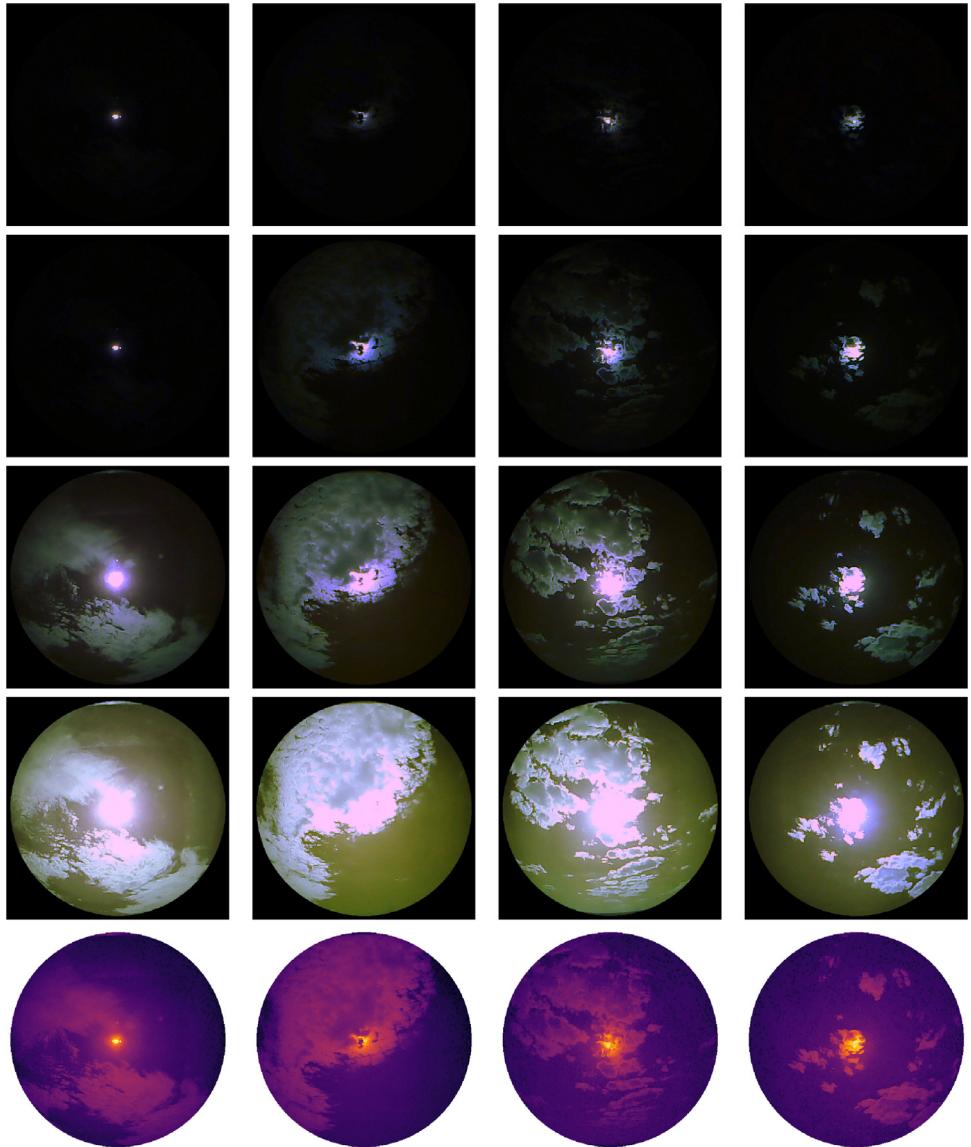
$$\mathcal{G}(x, y) = \frac{1}{C} \exp \left\{ -\frac{x^2 - y^2}{2\sigma^2} \right\}, \quad (17)$$

where  $\sigma = 7.5$ ,  $C$  is a normalization constant, so the kernel integrates to 1, and has dimensions  $15 \times 15$ . This operation is done to blur the mask pixels in the edges with the objective of smoothing out the transition between masked image regions in the resulting merged image. We define  $\tilde{\mathbf{M}}_e$  as each one of the masks that were processed using the Gaussian filter.

The different frames are then merged with the formula

$$\mathbf{X}^k = \alpha_e^k \cdot (\tilde{\mathbf{M}}_1 \odot \mathbf{I}_1^k) + \sum_{e=2}^4 \frac{\alpha_e^k}{e} \cdot [\tilde{\mathbf{M}}_{e-1} \odot (\mathbf{1} - \tilde{\mathbf{M}}_e) \odot \mathbf{I}_e^k], \quad \mathbf{X}^k \in \mathbb{R}^{D \times D} \quad (18)$$

where the Hadamard or element-wise product  $\odot$  is used to create the concentric rings with radii  $r_{e-1}$  and  $r_e$ . Coefficients  $\alpha_e^k$  are used to force the intensity of pixels of the outer edge of



**Fig. 5.** Result of the image fusion procedure. The four first rows of each column show the average of 10 sky images captured at a given exposure time (from top to bottom, 1, 4, 12 and 28 ms). The last row shows the images after completing the fusion algorithm. The resulting images are displayed in logarithmic-scale because of the high intensity of the pixels in the circumsolar region. Their size is  $450 \times 450$ . Pixels without information are not displayed.

each ring to be equal to the intensities of the inner edge of the next ring. We obtain a set of weights  $\alpha_e^k = \{1, \alpha_2^k, \dots, \alpha_E^k\}$  for each frame  $k$  by averaging the pixels of each frame outer edge as

$$\alpha_{e+1}^k = \alpha_e^k \cdot \frac{\sum_i \sum_j \mathbf{I}_e^k \odot \mathbf{R}_e^2}{\sum_i \sum_j \mathbf{I}_{e+1}^k \odot \mathbf{R}_e^1}, \quad \forall e \in \{1, \dots, E-1\}, \quad \alpha_e^k \in \mathbb{R}. \quad (19)$$

where  $\mathbf{R}_e^1$  and  $\mathbf{R}_e^2$  are rings of radius  $\epsilon = \sqrt{2}$  defined as

$$\begin{aligned}\mathbf{R}_e^1 &= \mathcal{M}(r_e) \oplus \mathcal{M}(r_e + \epsilon), \quad \forall r_e = \{r_1, \dots, r_E\} \\ \mathbf{R}_e^2 &= \mathcal{M}(r_e) \oplus \mathcal{M}(r_e - \epsilon), \quad \forall r_e = \{r_1, \dots, r_E\}.\end{aligned}\quad (20)$$

The images are then converted to 16 bit arithmetic for further processing, and stored. Since the maximum pixel amplitude after the fusion is 225, the 8 to 16 bit conversion equation is,

$$\mathbf{I}^k = \left( \frac{\mathbf{X}^k}{225} \right) \cdot 2^{16}, \quad \mathbf{I}^k \in \mathbb{R}^{D \times D}. \quad (21)$$

The outer circular region in the fused image is set to an intensity value of 0, as it displays artifacts produced by its own lens holder and case (see Fig. 5). Therefore, it does not include relevant information for the prediction.

## Ethics Statement

The research does not involve human subjects, animal experiments or data collected from social media platforms.

## CRediT Author Statement

**Guillermo Terrén-Serrano:** Software, Formal analysis, Methodology, Investigation, Data curation, Writing - original draft, Visualization; **Adnan Bashir:** Software, Data curation; **Trilce Estrada:** Writing - review & editing, Supervision; **Manel Martínez-Ramón:** Writing - review & editing, Conceptualization, Supervision, Project administration, Resources, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work has been supported by **NSF EPSCoR** grant number **OIA-1757207** and the King Felipe VI endowed Chair. Authors would like to thank the UNM Center for Advanced Research Computing, supported in part by the **National Science Foundation**, for providing the high performance computing and large-scale storage resources used in this work.

## References

- [1] W. Zappa, M. Junginger, M. van den Broek, Is a 100% renewable european power system feasible by 2050? *Appl. Energy* 233-234 (2019) 1027–1050.
- [2] D. Yang, J. Kleissl, C.A. Guemard, H.T. Pedro, C.F. Coimbra, History and trends in solar irradiance and pv power forecasting: a preliminary assessment and review using text mining, *Solar Energy* 168 (2018) 60–101. *Advances in Solar Resource Assessment and Forecasting*.
- [3] A. Mammoli, G. Terren-Serrano, A. Menicucci, T.P. Caudell, M. Martínez-Ramón, An experimental method to merge far-field images from multiple longwave infrared sensors for short-term solar forecasting, *Solar Energy* 187 (2019) 254–260.
- [4] M. Diagne, M. David, P. Lauret, J. Boland, N. Schmutz, Review of solar irradiance forecasting methods and a proposition for small-scale insular grids, *Renew. Sustain. Energy Rev.* 27 (Supplement C) (2013) 65–76.
- [5] H.-Y. Cheng, Cloud tracking using clusters of feature points for accurate solar irradiance nowcasting, *Renew. Energy* 104 (2017) 281–289.

- [6] C. Furlan, A.P. de Oliveira, J. Soares, G. Codato, J.F. Escobedo, The role of clouds in improving the regression model for hourly values of diffuse solar radiation, *Appl. Energy* 92 (2012) 240–254.
- [7] W. Kong, Y. Jia, Z.Y. Dong, K. Meng, S. Chai, Hybrid approaches based on deep whole-sky-image learning to photovoltaic generation forecasting, *Appl. Energy* 280 (2020) 115875.
- [8] J.-S. Chou, N.-S. Truong, Cloud forecasting system for monitoring and alerting of energy use by home appliances, *Appl. Energy* 249 (2019) 166–177.
- [9] C. Deng, Z. Li, W. Wang, S. Wang, L. Tang, A.C. Bovik, Cloud detection in satellite images based on natural scene statistics and gabor features, *IEEE Geosci. Remote Sens. Lett.* 16 (4) (2019) 608–612.
- [10] L. Ye, Z. Cao, Y. Xiao, Z. Yang, Supervised fine-grained cloud detection and recognition in whole-sky images, *IEEE Trans. Geosci. Remote Sens.* 57 (10) (2019) 7972–7985.
- [11] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, A. Fouilloy, Machine learning methods for solar radiation forecasting: A review, *Renew. Energy* 105 (Supplement C) (2017) 569–582.
- [12] O. Garca-Hinde, G. Terrn-Serrano, M. Hombrados-Herrera, V. Gmez-Verdejo, S. Jimnez-Fernndez, C. Casanova-Matteo, J. Sanz-Justo, M. Martnez-Ramn, S. Salcedo-Sanz, Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting, *Eng. Appl. Artif. Intell.* 69 (2018) 157–167.
- [13] F. Crespi, A. Toscani, P. Zani, D. Snchez, G. Manzolini, Effect of passing clouds on the dynamic performance of a csp tower receiver with molten salt heat storage, *Appl. Energy* 229 (2018) 224–235.
- [14] X. Chen, Y. Du, E. Lim, H. Wen, K. Yan, J. Kirtley, Power ramp-rates of utility-scale pv systems under passing clouds: Module-level emulation with cloud shadow modeling, *Appl. Energy* 268 (2020) 114980.
- [15] C. Poynton, *Digital Video and HD: Algorithms and Interfaces*, 2, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2012.