Cloud Dynamics

The State of Art

Marius E. Penteliuc
West University of Timișoara
Faculty of Mathematics and Computer Science
Department of Computer Science
Romania

Advisor:

prof. Marc Frîncu

Committee:

prof. Marius Păulescu

prof. Dana Petcu

prof. Daniela Zaharie

November 23, 2020

Abstract

Research in the field of cloud identification, cloud advection, and solar radiation is vast. This report presents and describes the state of the art for the three respective study fields. The strengths and weaknesses of each methods are accentuated to reveal improvement areas. The report draws conclusions for each topic and states objectives for research focus.

Contents

List of Tables

1 Introduction

Energy is crucial to human expansion and industrial development. It can be harnessed from various sources using multiple methods. Some of the sources are finite (coal) and will be consumed, others are renewable (wind, water, sunlight) and can not be depleted. One of these sources of energy is sunlight which can be captured by photovoltaic (PV) solar panels that transform exposure to sunlight into electrical power by a physical and chemical phenomenon. This process proved to be highly efficient and favored by power plants, the effect being that a greater percent of energy is generated through photovoltaics year after year.

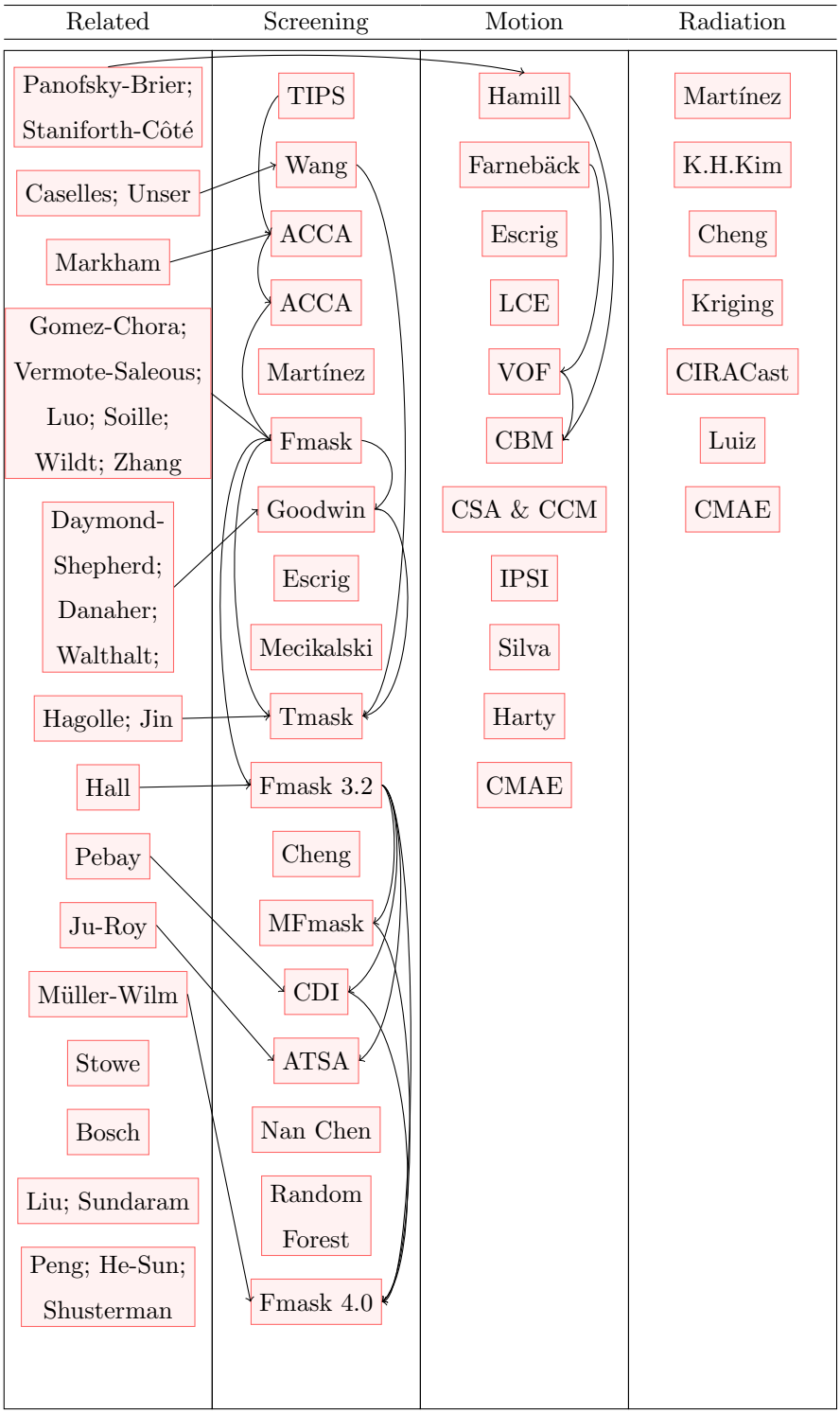
However powerful and inexhaustible this resource is, solar radiation is fluctuating and can cause power grid instability which means there is an imbalance between energy input and consumption. The main reason solar radiation is inconsistent are transient clouds that obstruct sunlight from hitting PV panels. Efforts to identify clouds and predict their position in time to help diminish their impact on power output variation arose and many of them succeeded under certain conditions.

The ramifications of transient clouds don't apply just to PV systems, but to remote sensing systems as well. Change monitoring is the process that involves the analysis of spectral information from satellite snapshots of Earth with the purpose to detect change in landscape features over time. Such activities are hindered by the presence of clouds that block land information from reaching the satellite's sensor. Identification of cloud contaminated images and masking of cloud and shadow pixels is an important preprocessing step for any remote sensing activity.

In this report, I present and describe state of the art methods and techniques in cloud detection and solar radiation variability estimation. In figure 1, the presented papers are grouped into four categories and ordered by year of publishing, and are connected by arrows that reveal connection and relevancy of each study. Each section begins with a brief description of the importance of the topic, a table displaying succinct contributions of the studies included, followed by a description of each method in chronological order of publishing, and conclusions. Section 2 contains techniques and methods regarding detection

of clouds in satellite and ground images. Section 3 examines the research that went into identification of movement and estimation of cloud advection. Section 4 reviews approaches that correlate solar radiation with cloud movement. Then, in section 5, a proposal of my objectives is stated, and the report is concluded with section 6.

Table 1: State of Art connections



2 Cloud Detection

2.1 Importance of Cloud Detection

Since the first satellite snapshots of Earth were made available by multi-spectral sensors, people started to investigate them and identify features of interest. Researchers started monitoring environment change such as biophysical trends, rates of change, quantifying disturbance of fire events, attack of insects, the influence of human agricultural and/or industrial activities like forest harvesting, urbanization and expansion. The problem imposed with using satellite imagery was the contamination with aerosols and clouds. They are a great deal of trouble because 2/3 of the Earth's surface is covered in clouds, therefore information about surface change is blocked by them. Clear pictures of Earth were necessary for monitoring activities and the volume of data was growing larger at a fast pace, but the manual, human observation process of tagging pictures with clouds was slowing studies workflows. From the desire of automating the process of removing cloud contamination, several cloud assessment techniques were developed and improved to the point of drawing a contour over cloudy regions and masking them out. Most representative techniques are shown in Table 2 and are described in subsection 2.2.

2.2 History of Cloud Detection

TIPS

In 1984, Su Jih-Jui [73] published a paper in which he described one of the earliest steps towards cloud screening, an algorithm that would automatically assess cloud cover. It was incorporated into the Thematic Mapper Image Processing System (TIPS), which was the ground system of the Thematic Mapper sensor aboard Landsat 4 and Landsat 5 satellites. Due to the computing limitations at the time of development, the algorithm was constrained and run under reduced computational load. These hardware limitation made it be insensitive to warm clouds, not be able to discriminate cloud from snow as expected, misclassify bright features (like deserts) as clouds, and have difficulties when the sun was at a low elevation.

Table 2: Remote sensing timeline

1984	The Thematic Mapper Image Processing System incorporates an algorithm to detect presence of clouds in a scene.
1999	Bin Wang proposes a fusion technique to automatically detect and remove clouds from Landsat images.
2000	ACCA is developed to aid in distinguishing cloudy images from clear scenes.
2006	Richard Irish documents and validates ACCA's performance.
2011	Martínez-Chico makes a cloud classification based on radiation data and sky images.
2011	Fmask, an object-based method for screening clouds and shadows, is introduced.
2013	Goodwin developes and automated method for screening cloud and cloud shadow.
2013	Escrig detects clouds by averaging surface albedo.
2013	Mecikalski quantifies growing cumulus obscured by cirrus clouds using properties derived from satellite data.
2014	Tmask is introduced as a variant of Fmask that uses time series data.
2015	Fmask 3.2 expands compatibility to Landsat 8 and Sentinel 2 images.
2015	Cheng develops a system to automatically classify clouds and predict solar irradiance.
2017	MFmask is launched to better detect cloud and cloud shadow in mountainous areas.
2018	David Frantz presents a Cloud Displacement Index to separate clouds from bright surfaces.
2018	ATSA is a new screening method for regions where cloud-free observations are infrequent.
2018	A machine learning-based masking algorithm is presented by Nan Chen.
2018	Nafiseh Ghasemian introduces two Random Forest methods for cloud detection.
2019	Fmask 4.0 is introduced and integrated different versions into a single release.

ACCA

Later, the Automatic Cloud Cover Assessment (ACCA) algorithm was improved to use the more capable resolution and additional bands of the new Enhanced Thematic Mapper Plus (ETM+) sensor found aboard the Landsat 7 satellite. Richard Irish presented the improved algorithm in the year 2000, and validated its performance in 2006 [29,30]. Landsat 7 was launched with the mission to provide cloud-free images on the Earth's landmasses, and ACCA was implemented to aid in distinguishing cloud-contaminated images and discard them. The new version used more bands than it's predecessor, specifically bands 2 through 5 converted to top of atmosphere (TOA) reflectances and band 6 converted to at-sensor temperature [43], and processed an image in two passes:

1. The first pass applied 8 filters to determine a precise cloud signature to isolate them from other elements in the scene. These included a Brightness Threshold, Normalized Snow Difference Index, Temperature Threshold, two Composites, and three Ratios between pairs of bands to filter out pixels that are unambiguous.

2. Data too ambiguous to be determined as cloud or non-cloud features were analyzed further in the second pass by looking at band 6's thermal data. Based on the presence of snow or desert areas in the scene, a new cloud signature is determined and threshold comparisons are made.

Non-cloud pixels with more than 5 neighboring cloud pixels are labelled as cloud to extend the cloud cover and reflect the extent of unusable data. The output is a cloud cover percentage score determined by a grid overlay.

Wang

In 1999 Wang [80] proposed a scheme to automatically detect and remove cloud and shadow information from Landsat images. It used an image fusion technique to remove cloud and shadow contamination and integrate complementary information into a composite image of two separate snapshots. This enabled the production of a cloud-free image for a given geographic location using two relative closely captured snapshots. A pre-processing step was needed before

detecting clouds because the brightness of the two images are likely to be inconsistent due to the different solar irradiance and atmospheric effects. This correction was possible to a degree assuming a linear relation [6] between the two brightness values.

The study used a thresholding method to distinguish clouds from ground regions, by isolating pixels above a threshold determined from the image's histogram. To ensure the detection was correct, a second lower threshold was used when comparing the brightness values for absolute brightness difference. This outputs a binary map for cloudy regions that is to be used later in the fusion of the two images.

Threshold values were sufficient for detecting clouds because their reflectance greatly differs from that of the ground, but shadows have much closer brightness values to the ground. In contrast, regions covered by shadow are much smoother in brightness change than regions not covered in shadow. Therefore, a wavelet transform can be used to detect shadows because its absolute coefficients corresponding to shadow regions are much lower than those of other regions. The specific transformation used was a discrete wavelet frame (DWF) [77]. Lower absolute values between the two images correspond to shadow regions and are output to another binary map.

To make a composite of the two images, the algorithm decides which regions to integrate by using the two binary decision maps. The result is a fused image with as much of the cloud and shadow regions removed (except for the parts were they are overlapping in both images), and because of the wavelet transformation artifacts are invisible in the final result.

Martínez

In 2011, Martínez et al. [44] used an approach based on the ratio of direct solar radiation incident in the surface to extraterrestrial radiation and ground sky images from a whole-sky camera TSI-800 to classify sky conditions in over 96 000 images taken between August 2009 and July 2010.

Typically, when processing radiometric data, clouds are characterized by measuring global radiation and using a dimensionless index such as the hourly clearness index k_t, defined as the ratio of global irradiance to the extraterrestrial

irradiance. This study used a different beam transmittance k_b index computed as the ratio $k_b = B_n/I_{0n}$, where B_n is direct solar radiation, and I_{0n} is extraterrestrial radiation, high values indicating clear sky conditions and low values signaling overcast skies or aerosols. Several attenuation groups were computed from the ranges of the k_b index values and clouds in each image were classified according to these attenuation levels instead of the more traditional cloud classification by base-height. The images underwent a visual check to separate clouds that are not blocking the sun disk and to extract representative cloud features for each level of attenuation. The final output is a cloud classification for different classes of sky conditions, each class having it's specific features.

Fmask

The following year, 2012, a new object-based algorithm for screening clouds and their shadows is introduced by Zhu and Woodcock [84]. It was named Fmask and was built on earlier approaches and detected cloud and cloud shadow for Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+). Currently this algorithm is the most widely known technique for cloud screening.

Fmask is threshold-based and identifies layers of potential clouds, shadows, and snow, then tries to match cloud objects to potential shadow based on the similar shapes the two have.

1. The potential cloud layer is identified in two passes. The first pass includes several spectral tests designed to isolate pixels that sometimes are cloudy and sometimes are clear: an universal Basic Test similar to the one used in the first pass of ACCA, which separates most of cloudy pixels from snow and vegetation; a modified version of the Whiteness Test proposed by Gomez-Chova et al. (2008) [22] to exclude pixels not white enough to be clouds; a Haze Optimized Transformation (HOT) Test proposed by Zhang et al. (2002) [82] to separate haze and thin cloud from clear-sky pixels; a spectral test similar to one used in ACCA to separate most of bright rocks from clouds; a Water Test that selects pixels in certain NDVI ranges based on the land and water values observed by Vermote & Saleous (2007) [78] to use for the calculations in pass two. Potential cloud pixels are identified

as pixels captured by the first four tests in pass one. Only if all potential cloud pixels represent less than 99.9% of the image, will the algorithm run the second pass, otherwise it will use the pixels directly for the cloud mask. The second pass consists of using the rest of the pixels to compute cloud probability over land and water separately because of the variable nature of the temperature distributions and the range of reflectances.

2. The potential cloud shadow layer is identified by looking for pixels that exhibit a darkening effect relative to their surroundings in the NIR band (Luo et al., 2008) [42]. A flood-fill morphological transformation (Soille, 1999; Soille et al., 2003) [68, 69], is applied to extract the difference of the original Band 4 (NIR) reflectance and the filled Band 4 reflectance as potential shadow area.

3. The potential snow layer is depicted by the use of several spectral test borrowed from the MODIS snow mapping algorithm (Hall et al., 2001), but with a lower NDSI threshold to include more pixels that are covered in snow less than 50%, such as forest areas. This threshold was also tested by Wildt et al., 2007 [81] for snow mapping in data from the Meteosat Spinning Enhanced Visible Infra-Red Imager.

Clouds and cloud shadows are matched by taking a cloud object and calculating a projected direction of its shadow with the use of the satellite sensor's viewing angle, the solar zenith angle, and the solar azimuth angle. Because cloud height is not known, iterations with increasing values for cloud height are made while calculated shadow and potential shadow layer similarity increases or stays over 98%. When the iteration stops, the cloud shadow is determined to be identified.

Goodwin

Many studies detect clouds by using methods that involve analysis of a single date image. Time series approaches differ from single date methods in that they highlight changes in reflectivity over time. In 2013, Goodwin et al. [21] developed a four-pass processing method for automatically screening clouds by analyzing multiple satellite images spread over time. The images were sources

from Landsat's TM and ETM+ sensors that captured them between 1999 and 2012. Available snapshots counted as low as 4 images per year in the beginning and amounted up to about 37 images per year. A few pre-processing steps are necessary to reduce reflectance variation in time series images over time. These correction steps include the use of a topographic correction [15] and a modified bidirectional reflectance distribution function (BRDF) [13, 79].

After pre-processing, the proposed four-pass framework can begin processing the temporal image sequence. Each pass consists of three steps:

- Sample a reference pixel from the time series using running statistics. Compare all pixels to the reference and identify potential cloud/shadow related outliers (that exceed a change threshold). The set of outliers together with a layer of differences of each pixel and its reference form the output.
- Consider the outliers are seeds and use them together with the difference layer to enable the application of a flood-fill technique. This is applied until a threshold is met to map the extent of cloud/shadow areas.
- Separate the cloud/shadow pixels from other possible land changes. A filter ratio of bands 1 and 7 is used for cloud identification and a cloud-shadow matching algorithm for shadows.

Identified regions were expanded slightly to incorporate cloudy pixels at the edges of clouds and shadows. The four passes of the algorithm are as follows:

1. Pass one identifies most of the cloud affected pixels. It uses a window of 365 days for selection of the reference pixel.
2. Pass two refines the identification of cloudy pixels with lower magnitude of change. It uses a window of 4 years for selection of the reference pixel.
3. Pass three identified most fo the shadow affected pixels It involved a similar approach to classification as clouds, but located significant low reflectance outliers instead of high reflectance.
4. Pass four attributed shadow layers to their respective clouds using a simplified adaptation of Fmask's matching method [84].

A question regarding interpretation of data arose from the cases where a pixel as flagged as both cloud and shadow. As noted by Zhu & Woodcock [84], the shadow of a high altitude cloud might fall on a cloud underneath it, so there is need for a standard definition for what is to be considered a cloud pixel and a shadow pixel. The study lead by Goodwin decided to categorize such a pixel as cloud.

Escrig

Many studies used multispectral tests [9,14,36,62,63] to discriminate cloud from land and water by using threshold values to separate pixels into groups, but in 2013 Escrig et al. [17] developed an algorithm to detect clouds using several multispectral tests with one key difference being the use of variable thresholds instead of fixed ones. These new thresholds for infrared and visible tests were based on monthly averaged values determined from the analysis of mean minimum surface temperature and surface albedo [47]. Surface temperature data was obtained from the Satellite Application Facility on support to Nowcasting (SAFNWC) database for sea pixels and from the Spanish Meteorological State Agency (AEMET) stations for land pixels, and surface albedo was synthesized from two years of satellite image data.

Several tests were performed using the new threshold values and identified clouds were classified by their top height – low, medium, or high altitude – and opacity. This, together with the determination of motion vectors by binary cross correlation, is done for forecasting purposes as power plants can highly benefit from prediction of solar radiation over short-term periods.

The results of the multispectral tests were compared with data captured by a whole-sky camera TSI-800 (same one as used by Martínez) and they showed the algorithm is reliably detecting clouds.

Mecikalski

As observed by others [10,17,21,46] clouds are not necessarily at the same altitude, but are separated on different layers of altitude. A consequence of this is that satellite images may contain overlapping clouds which impede the infrared observations of low level clouds when the view is blocked by higher altitude

clouds, thus making cloud monitoring difficult. Escrig et al. [17] categorized clouds into low, medium, and high altitudes using sounding air measurements, but did not address the overlapping issue. However, in 2013 Mecikalski et al. [45] developed a method to quantify growing cumulus clouds beneath higher cirrus clouds.

The proposed technique would use cloud derived parameters to detect cumulus blocked by cirrus clouds in pre-convective environment. Convective initiation is the development of convective clouds before a thunderstorm. To do this, they retrieved the visible optical depth, emittance, liquid water path, and effective particle size cloud-derived parameters from the Geostationary Operational Environmental Satellite (GOES) and began categorizing convective initiation events into several visible optical depth (τ) 10-increment bins from 1 to 51+ (0 meaning cirrus free sky, 51 and over meaning sky is heavily blocked by cirrus). After analyzing the time rates of change for the derived parameters, results showed the following:

- when $\tau < 31$ at least some of the information from the underlaying clouds pass through cirrus.
- some parameters can be monitored for values reaching up to $\tau = 30$ because of linear τ changes, while other only up to $\tau = 10$, and some only in the range $11 < \tau < 20$.
- for values $\tau < 20$ over 90% of information is available as compared to clear-sky situations. Cloud parameters for values bigger than this are invariant and more noisy.
- in situations of $\tau > 40$ cumulus is largely obscured by thick cirrus and no information can be reliably captured.

Tmask

Most multi-temporal algorithms for detecting clouds assume that no land cover change is taking place during the span of time series images so the referenced image and the observed image are very similar ground-wise [21,23,33,80]. If the two images are close in time, this is a reasonable assumption, but if the acquisition dates are distant, then land cover change due to agricultural or industrial

development is causing problems in analysis of temporal images. Another assumption is that no snow is present in the scene [21,80]. To address these issues and to improve the Fmask algorithm [84], Zhu & Woodcock developed a new algorithm called Tmask (multiTemporal mask) [85] that achieves more accurate detection, takes land cover change into consideration, and better separates clouds and snow.

Before running the masking steps, pre-processing is done on the time series to extract TOA reflectances from the original Digital Number (DN) values. To detect clouds and snow, Tmask uses the green band (Band 2) instead of the traditionally used blue band (Band 1) for cloud and snow detection because it is less sensitive to atmospheric influences. The thermal band was not used because it is sensitive to non cloud-related phenomena and could produce commission errors.

The first step in Tmask is to retrieve an initial cloud, cloud shadow, and snow mask using the Fmask algorithm on each image in the time series. If for any pixel the number of cloud-free observations is less than 15, Tmask will use a backup algorithm to substitute Fmask for that pixel time series.

Based on the results of the previous step, Tmask is applied to the time series to estimate TOA reflectances and identify pixels that differ dramatically from the estimation. It will iterate up to five times create the estimations and even if Fmask made at most five consecutive mistakes in the first step, the calculations will not be influenced. Then the observed-to-estimated value difference is assessed and pixels are identified as cloud or snow if difference is higher than a threshold, and labelled as shadow if lower. To separate snow, Tmask uses a modified version of the NLR algorithm [1] to generate a new threshold that discriminates cloud from snow in Band 5 TOA.

This algorithm will better identify cloud shadows, and better separates cloud and snow.

Fmask 3.2

In 2015, Zhu & Woodcock [83] improved and expanded on the original Fmask [84] algorithm. Major improvements were made to cloud detection, cloud detection over water, potential shadow detection, cloud shadow detection, and snow

detection. Expansion of the algorithm was made to include support for the Landsat 8 and Sentinel 2 images.

The criterion used to label a pixel as cloud if probability was higher than 99% was removed to prevent overestimation of the cloud layer. The fixed threshold for cloud detection over water is replaced with a dynamic threshold derived statistically to improve detection over turbid or highly sediment-concentrated water. The flood-fill transformation used for the Near Infrared (NIR) is also applied on the Short Wave Infrared (SWIR) band to better extract potential cloud shadows in areas with many dark objects. The cloud and shadow matching iterations would stop if similarity reached 98% of the maximum measured similarity. It was observed that premature stops would course matching errors, so the stopping condition was decreased to 95%. The temperature threshold used for snow detection was increased from 277 K to 283 K according to Riggs and Hall's revised MODIS snow algorithm [61]. Also, as pixels surrounding clouds and cloud shadows can be influenced by thin edges, surrounding pixels of snow covered areas can be problematic for remote sensing activities, therefore the snow mark is dilated to include those pixels.

The algorithm was also expanded to offer support for cloud, cloud shadow, and snow masking for images captured by Landsat 8 and Sentinel 2. Landsat 8 includes a new cirrus band that the improved Fmask algorithm will use to generate the potential cloud layer by applying a simple cirrus test to identify potential cloud pixels. The calculation of cloud probability will also use the cirrus band to identify cirrus clouds. Sentinel 2 has no thermal band, so many tests used on Landsat images can not be applied here. However, it has an extra cirrus band that can be used with the same cirrus test that was developed for Landsat 8's cirrus band to detect clouds. Cloud shadows locations are predicted from cloud object matching, but as no thermal band is available to determine cloud height range and three-dimensional size, the clouds height is considered fixed and objects are considered flat. Despite the absence of the thermal band, the algorithm can still identify cloud, cloud shadow, and snow pixels.

MFmask

Because Fmask did not perform very well on images from mountainous regions, a new algorithm was developed by Qiu et al. (2017) [58] to better detect clouds and especially cloud shadows in mountainous terrain. Together with TOA reflectance and BT, MFmask uses Digital Elevation Model (DEM) data extracted from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER).

Water detection is done to enable detection of clouds over land and water separately, but the original Fmask algorithm sometimes identified terrain shadows as water. This lead to misclassification of water and terrain pixels in mountainous areas. To remove false positive water pixels from the water mask, MFmask integrates a 10^{o} slope threshold into the Water Test because bodies of water usually have a lower slope, compared with higher slopes for terrain shadows.

A dynamic threshold values is used by Fmask to detect clouds based on the temperature probability derived from an entire image's pixels' BT. The problem with this is that the same threshold is used within the entire scene, but in mountainous regions elevation changes greatly impact environment temperature, meaning that clear-sky land pixels on top of mountains are cooler than pixels at low elevations. MFmask builds a linear lapse rate model for the environmental temperature and normalizes BT based on the DEM, assuming the elevation linearly affects temperature.

Cloud shadow is detected using a double projection technique that projects a cloud shadow on a reference plane, then back projects the reference shadow along the sunlight direction onto the DEM map. This will result in the cloud shape as viewed from above. Then the shadow location is found by iterating cloud height and matching clouds with their shadows. MFmask employs a topographic correction model and an estimation of neighboring clouds base height to eliminate terrain shadow, lakes, wetlands, and other dark features that might confuse the search for the NIR and SWIR bands.

Cloud Displacement Index

Fmask is susceptible to miss low altitude clouds in Sentinel 2 imagery, because they are hard to catch using any of the available bands. The cirrus band can easily detect high altitude cirrus clouds, but other clouds are indistinctive. Moreover, artificial materials (man-made structures) are included in the potential cloud pixel layer and cannot be separated from actual clouds. Thus, industrial and residential areas might result in an abundance of false positives, as artificial materials can be bright in the complete spectrum of Sentinel 2's bands.

David Frantz (2018) [19] proposed a new Cloud Displacement Index (CDI) that exploits Sentinel 2's sensor configuration and by relying on viewing angle effects, the CDI approach is demonstrated to be superior to the probabilistic approach used in Fmask. Because of the sensor configuration, Band 7, 8, and 8A are highly correlated, but also a parallax effect can be seen for them (objects at high altitudes are shifted in relative to the ground).

There are two ratios computed for Band 8 and 8A (similar for land surfaces): $R_{8A,8}$, and Band 7 and 8A (similar for cloud tops): $R_{8A,7}$. These ratios are highly indicative of cloud locations because land surface is spatially smooth in $R_{8A,8}$, but granular in $R_{8A,7}$, and clouds are opposite, granular in $R_{8A,8}$, but smooth in $R_{8A,7}$. To highlight image contrast, a 1-pass variance filter [50] is applied on each ratio ($V_{8A,7}$ and $R_{8A,8}$), which shows clouds as highly contrasted in $V_{8A,8}$, but having much lower contrast in $R_{8A,7}$. The CDI is computed from a normalized differenced variance ratio of the two textures, and easily separates low to mid clouds from artificial materials. He proposed this index to be implement into existing Fmask algorithm replacing the cloud probability module for Sentinel 2 processing.

ATSA

Despite the improved accuracy of using multi-temporal images in cloud detection, these methods rely on several cloud-free observations, which are hard to obtain in tropical and subtropical regions [34] where the persistence of cloud cover greatly reduces the frequency of clear observations. In an attempt to improve cloud and shadow detection in areas of persistent cloud cover, the Automatic Time-Series Analysis (ATSA) method was developed in 2018, that does

not need as many parameters, nor as many bands as other methods.

ATSA runs in five main steps:

1. Calculate a cloud and shadow index from image bands. The cloud index is calculated separately over land and water using two HOT transformations specifically optimized for each scenario, then combined into a single cloud index map. The shadow index is calculated using the NIR and SWIR bands over land, and Bands 1 and 2 over water.

2. Detect initial cloud mask by running an unsupervised k-means classifier to pixel samples. The classifier separates the samples into three classes: clear surfaces, thin clouds, and thick clouds.

3. Remove non-cloud bright pixels and find very thin cloud pixels. This is done by analyzing the cloud index time series for variations that are influenced by clouds, but not by changes in land cover.

4. Estimate potential shadow by matching them to clouds using the geometric relationship between cloud height, shadow location, and the sun's position. However, because estimating cloud height from a thermal sensor is not possible on historical images without such a band, a range of possible heights is used instead to estimate a cloud's shadow location.

5. Detect shadow by predicting potential shadow zones using an inverse distance weighted interpolator. The darkness of the predicted pixels are estimated as the difference between the calculated shadow index and the predicted values. After that, a k-means classifier is used to separate pixels into clear observations and cloud shadow. A time series analysis is applied to refine the initial shadow mask.

Compared to Fmask [84], ATSA can deliver robust cloud and cloud shadow masks in areas of persistent cloud cover. A limitation of ATSA is that it struggles in snow covered regions, were it would classify snow as cloud, but this is rare in tropical and subtropical images for which the algorithm was designed.

SCM

Threshold based methods for detecting clouds are susceptible to misclassify cloud and snow pixels due to similar optical properties in visible and near infrared bands. These methods become increasingly complicated with many satellite bands to achieve snow-cloud separation. Chen et al. (2018) [8] developed a machine learning algorithm to combat the problem of discriminating clouds and snow.

The algorithm is a multilayer perceptron neural network with one hidden layer that was trained using 20 million samples generated from radiative transfer simulations. The TOA reflectance is simulated using a DISORT radiative transfer model [38,70,75]. The advantages of using simulations to produce training data are: saving human effort in identifying images, availability of as much data as needed, having a full range of climate possibilities, and fast adaptability of the algorithm for new satellite sensors.

The Snow-ice Cloud Mask (SCM) algorithm is able to classify pixels after training using fewer satellite bands than other methods. It has low misclassification rates of clear-sky pixels and performs similar to the MODIS cloud mask over vegetation images.

Random Forest

Another method for detecting clouds is using Random Forest (RF) based methods. Ghasemian and Akhoondzadeh (2018) [20] proposed two such methods that take into consideration spectral characteristics as well as textural. Spectral features are the TOA reflectance and BT values of Terra MODIS and Landsat 8 sensors. Textural features represent the spatial distribution of spectral information. Clouds are variable, but their texture is different from ground and snow cover texture due to influence of terrain, vegetation, and human activity. The type of textural feature used by Ghasemian is Gray Level Co-occurrence Matrix (GLCM) [7,65,76]. The spectral and textural features are used as input data.

The two proposed methods for classification are:

- The Feature Level Fusion Random Forest (FLFRF) algorithm which fuses visible, infrared, and thermal features (both spectral and textural) into a

single set, adjusts RF parameters, and identifies clouds, snow and ice, and background pixels by classification using RF.

- The Decision Level Fusion Random Forest (DLFRF) algorithm which separates visible, infrared, and thermal features intro three sets, runs them through the RF algorithm sequentially, adjusts each RF classifier, and predicts the classification map using a voting matrix.

They also added Robust Extended Local Binary Pattern (RELBP) descriptors to input features to assist the algorithms if the input set is incomplete. This will enable them to improve cloud detection and more precisely identify snow and ice pixels. In the absence of infrared data, adding RELBP helps substantially.

FLFRF is more accurate than DLFRF, but not as efficient when the set of input features is large. Compared to traditional methods, the two methods produce good results and does not overestimate cloud pixels, nor needs a threshold to be set.

Fmask 4.0

The single date Fmask algorithm processes an image with no prior knowledge of the scene. Inclusion of auxiliary data could make the algorithm produce better results by relying on ground truth instead of generating intermediary masks. Qiu et al. (2019) presented version 4 of Fmask [59] which tackles problems that could be solved by knowing the geography of a scene, omission error in cloud detection for Sentinel 2 images because of the lack of a thermal band, and confusion with built-up objects and clouds.

While the cloud shadow detection module of Fmask 4.0 is brought directly from MFmask [58], the cloud detection module is developed and calibrated using the afore mentioned auxiliary data. They include the Global Surface Water Occurrence (GSWO) dataset which provides pixel-level water occurrence in percentages from no water to permanent water. This is great for better separating land and water surfaces as cloud detection is done separately over each surface. Previously only several spectral tests discriminated land from water, which could have been erroneously done. The DEM Version 2 dataset from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) has

global availability and is accurate enough to generate terrain slope and aspect data. It is useful for normalizing the cirrus band in Landsat 8 and Sentinel 2 and attenuate elevation impact on this band.

The cloud probability for pixels over land has been improved from the previous version of FMask by adding a contrast test for the NIR and SWIR bands which greatly reduces cloud commission errors over built-up areas. To replace the temperature probability for Sentinel 2, a new cloud probability that calculates based on a new HOT transformation is invented. The weights that favored cirrus cloud probability had been adjusted because previously it sometimes dominated the entire cloud detection. Also, calibration is done to update the global optimal cloud probability thresholds to include more sensors.

To not confuse clouds with snow, ice, and urban areas, this version of the algorithm combines spectral and contextual information to better separate snow and ice pixels from cloud. Because clouds and snow have different textures, a new Spectral-Contextual Snow Index (SCSI) is computed from Band 2 Standard Deviation and NDSI to distinguish smooth homogenous snow from clouds. Commission errors caused be urban areas and snow/ice present in mountainous regions are removed by applying a morphology method that uses simple spectral tests to identify potential false positive cloud pixels. The elimination process for these pixels involves an erosion and dilation method that will remove isolated pixels associated to built-up and recover cloud shapes.

Fmask 4.0 demonstrated that it is more accurate then the previous version of Fmask and also achieved higher accuracies for Sentinel 2 imagery than the Sen2Cor algorithm by Müller-Wilm et al. (2018) [48].

2.3 Conclusions

Remote sensing activities such as monitoring land change and human influence are having an important boost because of cloud screening schemes. Numerous workflows based on multiple techniques have been developed to identify cloudy pixels and mask out contaminated information. Both satellite and ground images are useful in the field of cloud detection and a great amount of research was done to develop new methods and is ongoing to further improve upon current approaches.

The strength of the techniques described in this section is their adaptability to an abundance of sensor configurations with different spectral ranges. Algorithm will compensate even for the lack of a thermal band by using ratios of the remaining bands. Auxiliary data inputs such as sun-cloud-satellite geometric relationship, global constructed water masks, and DEMs have a strong positive impact on the ability to accurately isolate clouds and their shadow from other land features. The collections of satellite imagery are larger than ever and are growing constantly with data from newer higher resolution sensors. This data can be used to train and test algorithm and even simulate different sensor configurations before launch.

The issues that still exist in this field are the difficulties these algorithms have in detecting cloud shadow. The most prevalent assumption is that shadows have a similar shape to the clouds that cast them, which is not invariably true. Slope angles and terrain shadow can confuse algorithms into producing errors.

Algorithms trained on simulations are broad, but not localized to specific sites. They could be better suited for certain locations if trained on simulations using local measurements as parameters.

Time series analysis is well suited for images from geostationary satellites, but not from orbiting satellites. These can not capture relatively frequent snapshots of the same location because of their orbital trajectory. Geostationary satellites do capture frequent snapshots of the same area, but the output is at a low spatial resolution compared to satellites in orbit and such their applicability to a confined area is limited.

3 Cloud Motion

3.1 Importance of Cloud Motion

The sun is an unfailing source of energy and solar harvesting is growing in use as fossil fuel resources are shrinking and public opinion is in favor of power generated using renewable resources. However, sun radiation is not constant, but is varying place-to-place, month-to-month, and day-to-day due to geography, change in seasonality, and weather conditions. Moreover, formation and dissipation of clouds, the passing of clouds advected by winds, cause a variability of solar irradiance observable in a minute-to-minute window. Photovoltaic (PV) power plants are hard to integrate into grid due to solar irradiance variability because power output is affected by transient clouds. Storage systems such as molten salts (for heat production and batteries), backup generators, and scheduling tables can be used to reduce the effect of inconstant output into the grid, but they need an accurate prediction of when power output will decline or increase. As clouds are the main cause of irradiance variability, cloud motion estimation to predict variability is vital for power plants, especially small-scale grids that don't have the capacity needed to absorb power fluctuations. Many techniques were proposed to detect motion and predict cloud movement as shown in Table 3. They range from using numerical weather forecast models to employing machine learning algorithms, from capturing data from ground sensors and sky cameras to incorporating satellite imagery with cloud screening procedures, and from cross-correlation with solar radiation to detecting speed and direction using block matching techniques. In subsection 3.2 several techniques are presented to show their individual strengths and weaknesses.

3.2 History of Cloud Motion

Hamill

In 1993, Hamill and Nehrkorn developed a scheme for short-term cloud forecast that skillfully output cloud advection [24]. Their technique was based on lag cross correlations [49] and generated displacement vectors looking at differences in two consecutive satellite images. A subset of pixels in the first image is

Table 3: Motion detection timeline

1993	Hamill described a short-term cloud forecast scheme using cross correlations and creating displacement vectors.
2003	Farnebäck presented a two-frame motion estimation algorithm based on polynomial expansion.
2013	Escrig employed binary cross correlations to determine motion vectors for cloud sectors.
2013	Bosch and Kleissl used a network of ground sensors to detect cloud speed using time delays.
2015	Variational optical flow is used by Chi Wai Chow to estimate cloud motion and stability, and forecast intra-hour cloud locations.
2016	Zhenzhou Peng proposes a hybrid approach between block matching and optical flow models to estimate complex cloud motion.
2018	Mohammad Jamaly estimates cloud motion from irradiance data using two methods: CSA & CCM.
2018	Fei Wang proposed a image phase shift invariance based calculation method using Fourier phase correlation theory to determine cloud motion vectors.
2018	Silva builds a cloud dynamics model simulation.
2019	Travis Harty forecasted cloud index by assimilating data from a 2D advection model with cloud motion vectors.
2020	The CMAE method is introduced to estimate cloud motion from small-scale irradiance sensor networks.

chosen and correlations are made to find the same subset of pixels at a different location in the second image. The centers of the two subset locations define a displacement vector. Through quality control, inaccurate displacement vectors generated because of analysis areas that are too small are eliminated.

Cloud free images will generate null displacement vectors that affect prediction by slowing down advective vectors (correctly generated displacement vectors) at cloud edges,so null displacement vectors are removed to avoid this issue. With the assumption that all clouds are on one single layer and cloud deformation is not severe, their cross-correlations scheme produces reliable displacement vectors even knowing higher winds advect clouds faster than lower winds.

An objective analysis is done to create a displacement vector for every pixel, which will results in a continuous flow pattern by applying multiple successive corrections. Before producing forecasts, the vectors are modified by a semi-Lagrangian displacement scheme [71] to better forecast trajectory in a curved flow. The scheme derives a compromise vector that advects cloud pixels to forecast location.

All pixel locations are forecast by a bilinear interpolation method. This scheme produced good displacement vectors for at least the first hour of forecast, which is sufficient for nowcasting small cloud features.

Escrig

Numerical weather forecast models use cloud motion vectors as parameters for predicting weather phenomena. Short-term solar radiation forecast can also benefit from generated cloud motion vectors. Escrig et al. (2013) [17] used a maximum cross-correlation method to identify motion vectors and estimate cloud movement by tracking them.

Winds have different directions at different altitudes and Escrig addressed the aspect of vertical wind layering by splitting motion vectors into three heights, after the classification of clouds as high, medium, and low altitude by applying Stowe's latitude criterion [72] on sounding air measurements data. At the spatial resolution the MSG imagery is provided, wind is observed to have horizontal variability. To account for that, the images are segmented into five horizontal

sectors: four sectors are obtained by dividing the image by its two bisections, and the fifth sector is the same size as the other, but centered on the image such that it overlaps equally on each of the other sectors' corners.

For each sector at each height, maximum cross-correlations are performed on the motion vectors of three consecutive images. Cross-correlations are also done for non height layered images and and non sector split images to avoid influence of clouds that occupy more than one layer and clouds that are exiting the scene. After all vectors are tested to be coherent, six quality tests are used to pick the most suitable cloud tracking vector for each height and sector.

This technique proved to determine good tracking vectors for more than 86% of the generated vectors and could be useful for predicting solar radiation. A TSI-800 whole-sky camera was used to visually check that the motion vectors the algorithm produced were correct and cloud tracking error was determined to be around 10% after discarding rejected vectors (which could be replaced with the last calculated vector in real time applications).

LCE

Another method to detect cloud speed and direction, and generate CMVs is tested by Bosch and Kleissl in 2013 [3]. Their method involves a network of ground sensors located on the site of a PV power plant. Using time delays between cloud edge passes over pairs of sensors, CMVs are generated and evaluated to assess cloud movement.

The PV power plant where tests were conducted has almost one million panels. Five of those were used as reference cells that combined represented two triplets of sensors (an origin and two more sensors). Assuming a linear cloud edge (LCE), a pair of sensors (excluding the origin) parallel to the cloud edge will be affected by the cloud's shadow at the same time. If the cloud moves in a direction that is parallel to the sensor pair, and the shadow is over the origin, it will never touch the other two sensors. To obtain a CMV, two cloud edge passes need to occur over the triplet.

Time lags are determined from measured local maxima (right before or after a shading event) and minima (cloud center) of measured irradiance. Preprocessing similar to Bosch et al. (2013) [4] is applied on the measurements to discard

noise induced by high frequency of data acquisition. The CMV is is determined by selecting the most frequent CMV in the last 60 minutes.

CCM is used to validate the LCE method starting from the assumption that pairs of sensors aligned with the CMV will be highly correlated when a time lag is applied to account for the distance. To obtain the speed of movement, a pair of inverters aligned with the direction of movement is selected from all the 4560 possible pairs of inverters. The CCM method is used to find the speed of the cloud.LCE is reliable for producing CMVs when the number of available sensors is reduced.

VOF

Chi Wai Chow et al. (2015) [12] proposed a variational optical flow (VOF) technique to determine CMVs from sky images. VOF identifies cloud deformations that happen in the span of a few minutes, which often are dismissed in cloud advection forecast models, and quantifies the stability of cloud formations.

The data used is comprised of snapshots from a sky camera located at UC San Diego that captures the sky every 30 seconds. The optical flow constraint (OFC) equation is a brightness constancy equation linearized by a first order Taylor expansion. However, this is not enough to recover specific motion as additional constraints are needed. Therefore, a simple, low computational algorithm proposed by Liu (2009) [39] that is flexible in parameters is used by the study. Inverse mapping is used to find a correct input pixel from the source image for each pixel in the output image. As opposed to forward mapping, where the output image could have holes and overlapped pixels, inverse mapping is guaranteed to map each output pixel, leaving no holes and making no overlaps.

Point trajectories are used to quantify cloud stability. An optical flow tracker based on Sundaram et al. (2010) [74] initializes tracking points across the entire first frame of an image (discarding points over hard to track homogenous areas). A point is tracked until one in three conditions is met:

- Point has moved out of the frame and can no longer be tracked.
- Forward and backward optical flow are inconsistent with each other.

- Point is no longer easily found due to loss of structure in neighboring pixels.

If too many tracking points were terminated in an area, new points are initialized in the next frame. The average time duration of terminated points for the current frame is a measure of cloud stability. Short tracking points (less than 1 minute) and those very close to the frame edge are not considered for the average.

The VOF method is compared with CCM forecasts by shifting cloud map pixels according to the motion vectors generated by the VOF method. This map is overlaid with the actual cloud map and a forecast error can be computed. The method demonstrated accurate cloud motion forecast is impossible with unstable clouds and averaged an error reduction of 39% to 19% for 0 to 15 minute forecasts compared to CCM.

Zhenzhou Peng

VOF [12] and block matching [24] techniques have their strengths in specific scenarios, but also weaknesses, such as image noise and brightness variation sensitivity, neglecting multiple cloud layers, and block segmentation sensitivity that can lead to inaccurate results. In an effort to address such issues, Peng et al. (2016) [52] described a new hybrid model that combines the two techniques such that the VOF model's output is refined by the block matching method.

They use the Cloud-block matching (CBM) technique [51] to determine and match cloud blocks dynamically. A binary classifier based on Support Vector Machines (SVM) is implemented to identify cloud pixels and generate a cloud mask. The mask is divided into homogeneous blocks using quad-tree structure decomposition method [67] which successfully groups color, texture, and structure to aid block-wise motion tracking.

The motion vectors obtained from block matching are refined to remove falsely estimated vectors due to image noise by utilizing histogram statistics to extract the most frequent motion [25]. Small -scale vectors are also removed to ignore sky pixels and slow clouds that affect dominant motion calculation. Usually there are no more than three cloud layers, so the three most common dominant motion vectors are picked to represent the dominant motion for each

cloud layer.

Filters are applied to remove outliers that cause noise in the motion field, find dominant motion patterns, and utilize as much cloud information as possible. The dense flow field is optimized by an iterative algorithm. A velocity threshold is used to identify clear sky pixels and eliminate small motion. Outliers to the dominant motion patterns are identified by another filter that greatly refines the motion field and a weighted median filter smooths and removes the noise of the motion field. Repeating the whole process up to three times best balances accuracy and computational cost.

The hybrid method extracts dominant motion patterns from block-wise motion tracking and estimates a dense motion field through the use of refinement filters. Results show a reduction of 30% of the angular error compared to other models [2, 5, 26, 27, 40, 51] and a 10% lower MAE.

CSA & CCM

CMV estimation using the methods described are not efficient and lack granularity, except for local ground measurements [4]. In 2018, Jamaly and Kleissl, motivated by the increasing availability of dense PV power output observations, proposed two new methods: cross-spectral analysis (CSA) and cross-correlation method (CCM) [32].

Engerer and Mills, 2014 [16] demonstrated that PV power output can be converted into clear sky index, meaning that PV systems could be used as a large irradiance sensor to detect cloud motion. CSA estimates cloud speed and direction by cross-spectral analyzing observations at given site locations. Instead of using the classical CSA [28, 66], cloud movement speed is detected as the median of velocities in the direction where the minimum time delay variation moment occurs.

Quality control to exclude particular sites and data makes CSA more accurate. This step excludes low quality data by removing conditions with low variability, pairs of sites that are too distant, and sites that are less correlated. Restriction of point selection to specific geographic regions reduces the computational costs, while still achieving same results.

The CCM method is applied to ground measured data, and finds the best

matches for any given subset of pixels in two images. It is applied to the whole domain, but also to smaller subsets to improve accuracy. A three-step search proposed by Li et al. (1994) [60] is used to reduce computational cost compared to a more traditional full search block matching algorithm.

Quality control is a step that also benefits the CCM method. By removing vectors that differ from the local average enough to pass a threshold, the results are more consistent. Several quality control parameters along with specific thresholds and local consistency make CCM produce more accurate results. The final velocity field is obtained by successive corrections as Hamill and Nehrkorn did [24].

Results show that the modified CCM method provides reliable cloud motion speed up to a horizon of 50 seconds. CSA has reduced computational costs because of restricted number of sites, but keeps the save level of accuracy. Its results are reliable in scenarios of low cloud cover fractions. Changes in cloud speed or direction will degrade CSA accuracy if they are often occurring, making CSA estimates best for short time intervals.

3.3 Conclusions

Displacement was observed since interest in motion emerged to detect movement and predict it. Cloud motion was estimated to produce forecasts of cloud advection, and compute motion vectors to use as input to numerical weather forecast models. Motion vectors that describe displacement of the cloud field can be generated using various techniques that rely on sky or satellite imagery, ground sensors, statistical methods, time lag correlations, etc.

The advantage is that displacement vectors can be computed relatively straightforward from two frames if the background is stable and the motion is sufficiently ample to be observable. An array of mathematical techniques can be used to match cloud blocks, track points, and cross-correlate sectors of an image to generate motion vectors. Options in detection are abundant as motion vectors can be detected using data from geostationary satellites and from ground irradiation sensors, and even PV modules power outputs.

Despite the data available for research, images from high spatial resolution sensors onboard orbiting satellites can not be used for motion estimation because

of the sporadic captures of confined areas of interests. For these areas, the best alternative is to use sky cameras or ground sensors as rich data sources.

Discussed research assumes no clouds are forming or dissipating when predicting motion. In reality clouds are not only changing locations and modifying their shape, but new clouds appear seemingly out of thin air and wind can break clouds into smaller parts that easily disappear. This issue should be addressed to have reliable prediction of future clouds.

Table 4: Solar radiation timeline

2011	Martínez-Chico classified clouds based on solar radiation.
2014	Kee Han Kim cross correlated solar stations data with information from NCDC & NOAA to evaluate best method for estimating hourly global solar radiation.
2015	Cheng used automatic cloud classification to predict short-term solar irradiance.
2017	Jamaly and Kleissl use Kriging method to forecast irradiance data at an arbitrary point.
2018	CIRACast introduced as a solar forcasting system that leverages satellite imagery, wind field data, and radiative transfer calculations.
2018	Luiz evaluated solar irradiance variability using satellite imagery for three sites in Brazil..
2020	CMAE is a new method for estimating cloud motion using small-scale irradiance sensors networks..

4 Solar Radiation

4.1 Importance of Solar Radiation

Solar radiation is an great resource of energy and photovoltaic (PV) power plants are increasingly used to generate electric power as they are environmental friendly and public opinion is viewing them as a better alternative to fossil fuel energy generation. Buildings that have HVAC systems could reduce operating costs significantly if the level of solar radiation hitting the building is known a priori. Cold air flow could start cooling before an increase of indoor temperature is detected and heating elements could start and stop depending on short-term prediction of irradiance levels. In Table 4 several studies that estimate or predict solar radiation levels are shown. They are described in subsection 4.2 and advantages and disadvantages of each method is discussed.

4.2 History of Solar Radiation

Multi-model prediction

In 2015, Cheng and Yu proposed a new multi-model technique to predict solar irradiance based on the type of clouds present in ground-based all sky images [10]. Three local pattern descriptors (LBP, LTP, and LDP) are used to extract texture characteristics from images divided into blocks. Those extraction, along with classical statistical features form a feature vector that a trained SVM classifier will get as input and then classify images into six classes depending on the observed cloud types.

Based on the classification results, multiple solar irradiance prediction models are constructed to accommodate the different weather conditions. The number of prediction models range from 3 to 6, where three classes consists of clear sky (model A), cirrus, cirrostratus, scattered cumulus, and altocumulus cloud types (model B), and cumulus, cumulonimbus, stratus cloud types (model C). Cases with more prediction models separate cloud types from models B and C into multiple, more specialized classes. Clear sky conditions, cirrus, and cirrostratus clouds correspond to low irradiance variation; medium to high irradiance changes can be detected from scattered cumulus and altocumulus clouds; and cumulus, cumulonimbus, and stratus cloud conditions manifest the highest irradiance variation.

Rather than predicting global horizontal irradiance directly, a scheme that converts irradiance into clearness index yields better forecast accuracy [11]. Irradiance is computed from the predicted clearness index using the Perez conversion model [56] with parameters that can be determined for a location at a specified time.

The proposed multi-model shows higher prediction accuracy for irradiance than other methods that are blind to cloud types. It was validated using a dataset with very high irradiance variation.

Kriging

Another method to forecast irradiance data is using Kriging to interpolate spatial and temporal data at locations with no observing sensors. Because ground

solar measurement sensors are sparse and satellite-derived data is of low temporal resolution, interpolation is required to produce data at arbitrary points in space and time where no observations were made. Kriging is a stochastic interpolation method that is superior to other techniques. It preserves irradiance data properties and can be applied to ground measurements and satellite data alike. Of the various Kriging methods available [55], Jamaly and Kleissl used the spatial and spatiotemporal ordinary Kriging method to make estimations of solar irradiance [31].

Passing clouds determine spatiotemporal variability in irradiance data. To model such events, it is required to use an anisotropic covariance function. This study proposed a new non-separable anisotropic covariance function that is based on Schlather's Lagrangian covariance function [64]. Assuming constant cloud motion speed, the proposed anisotropic spatiotemporal Kriging method accurately produces results for steady and slow varying cloud motion cases. For long time series where cloud motion is not constant, but has moderate to high variability, it can split the series into shorter intervals to provide better results.

An important function in the ordinary Kriging method is the semivariogram function that describes the dependence of solar radiation to space and time in degrees. It is the variance of two (spatiotemporally separate) observation points' solar irradiation difference. The empirical semivariogram is deducted from observed irradiance and pairs are grouped into bins. Then it is modeled with a parametric function which's coefficients are computed using a weighted least squares method and minimizing the empirical and parametric difference. Cloud motion effects (estimated by CSA or CCM [32]) are considered to fit the semivariogram function. Using the parametric spatiotemporal semivariogram function, irradiance can be estimated at arbitrary locations and times.

Validation of the Kriging method showed improvement over a persistence model. Parameter shrinkage was applied to obtain a 99.86% reduction in computational costs. The shrinkage was justified by the dismissible data in uncorrelated zones. Accurate estimation of cloud motion is necessary for reliable irradiance forecast, therefore shorter time intervals produce better results than long unstable time series with variate cloud speeds. This method is promising for satellite data, but is most fitted for use with ground data.

CIRACast

While the Kriging method used by Jamaly and Kleissl [31] is best suited for use with ground measured data, satellite-based solar forecasting is tackled by Miller et al. (2018) with a satellite/model coupling scheme [46]. CIRACast is a scheme that includes cloud properties, steering winds, and radiative transfers to compute surface radiation.

Cloud properties such as cloud-top height/pressure, optical depth, effective particle radius, and liquid/ice phase are obtained from NOAA's CLAVR-x code package and a NWP model uses these properties to model winds at multiple levels. This model also provides total column ozone and precipitable water to send as input to code responsible with radiative transfer calculations.

After advecting the cloud field according to the winds generated by the NWP, surface irradiance is determined for new cloud shadow locations. For this, Pinker's Satellite Algorithm for Shortwave Radiation Budget (SASRAB) radiative transfer code [57] is used in forecast mode. At each time step, surface irradiance is computed from cloud properties, reflectance in the visible spectrum, precipitable water, ozone, solar/satellite geometry, and surface reflectance. A mask is produced from composites of the second-darkest GOES VIS band reflectance over a two-week period to hide snow and other land features that could be picked up as false cloud.

This coupling of operational geostationary satellite imagery with a numerical model, produces short term solar irradiance ramp forecasting.

TODO: read methodology and conclusions for CIRACast – I think this subsection is too vague

Luiz

In 2018, Luiz et al. evaluated the variability of intra-day solar irradiance using three ground based solar irradiance measurement sites in Brazil and described a new method to evaluate surface solar irradiance using only satellite imagery [41].

At the three sites in Brazil, global irradiance data was captured for a one year period starting with July 2016 with a minute temporal resolution. Using this data, the clearness index (k_t) was calculated and empirically corrected to remove the diurnal and seasonal cycles, and the air mass effect [54]. Kleissl's

moving average approach [35] was used to diminish different time steps' influence over the evaluation of ramp rates for temporal resolutions of 30, 5, and 1 minute.

Using satellite imagery, the effective cloud cover coefficient can be computed for each pixel as the ratio Ceff $= (L_r - L_{clr})/(L_{cld} - L_{clr})$, where L_r is the radiance measured in the visible satellite channel, L_{clr} and L_{cld} are the clear sky and overcast radiances for a compositing of one month of satellite images. Usually, L_{clr}/L_{cld} is the lowest/highest radiance level for each pixel during the one month duration. The ramp rates based on Ceff are produced at a temporal resolution of 30 minutes.

To understand the behavior of ramp rates, they use Lave's proposal of a Variability Score [37] that takes the largest value of the product of a ramp rate for a certain temporal resolution multiplied by its probability, and scales it by 100. It is well correlated with a transformer's number of tap changes. A larger variability score means more variability in the solar irradiance.

By comparing both satellite observations and ground measurements for all three locations, it was observed that there is a linear variability relationship between the two methods. Cloud coverage variability can represent k'_t variability. Satellite observations proved to be a relative simple method for evaluating solar irradiance variability without requiring large computational resources and a good provider of information necessary for PV power grid balancing.

4.3 Conclusions

Irradiance data is used as input in many fields of research and getting accurate predictions of solar activity is the difference between having efficient systems and increased operational costs. Solar activity has great impact buildings' energy consumption and public health, but the greatest impact is on PV power plants. Therefore, much attention is concentrated towards collecting radiation data and predicting hourly irradiation forecasts for PV sites.

Irradiance data can be captured using ground sensors and irradiance at locations where no observations are made can be estimated by applying interpolation methods to two neighboring sites that do have active ground sensors. Geostationary satellites can also be used with radiative transfer calculations to advect the cloud field and estimate solar irradiance for new shadow locations.

Calculations using simulated irradiance data or information from typical year weather constructs are possible, but in order to accurately make predictions on solar activity, data collection is needed, which means additional costs for installing and maintaining irradiance sensors or sky cameras.

5 Objectives

Cloud detection algorithms are practically sufficiently many and further research should be focused on improving current methods by tackling niche cases were cloud screening accuracy is not as high as desired.

In [53] I trained a feed forward back propagation neural network on motion vector fields generated by applying Farnebäck's variant of Optical Flow [18] on sequential geostationary satellite images. My findings show that motion vectors can be predicted using neural networks and are sufficiently reliable to forecast future cloud locations. As concluded in 3.3, cloud advection prediction techniques assume no clouds are forming or dissipating during the forecast period, but rather just translate positions and change shapes slightly. It should be interesting to analyze when and where new clouds are forming in satellite images and observe how cloud shapes are evolving according to the estimated motion vector field. This could give prediction algorithms (including my own) increased accuracy by allowing not only forecast of cloud advection in the scene, but also the emergence of new clouds and breaking apart other clouds where this behavior is expected.

Geostationary satellites are locked in space relative to the Earth's rotation, but don't capture high spatial resolution images as well as low-orbit satellites do. However, orbiting satellites don't capture the same scene continuously and can not be used for motion estimation of clouds. A single frame is frozen in time and is void of any movement information. A workaround to this fact could be taking advantage of the sensor configuration and time delays between the acquisition timestamps of each sensor. Because clouds are visible in more than one spectral band, the acquisition time delay, coupled with the geometry of the sensor, could be sufficient information for motion estimation on fast moving clouds.

Solar irradiance is predictable from sky images that were correlated with ground sensors, but satellite imagery was correlated with computed solar irradiance rather than observed solar irradiance, which is objectively inferior to true data. Correlations between clouds detected in satellite imagery and solar irradiance levels recorded by ground sensors should be valuable information

for establishing if solar radiation can be predicted from satellite imagery alone making use of cloud type classifications.

Taking into consideration all of the above, I state my objectives as follow:

1. I shall correlate cloud types over city of Timișoara with solar radiation recorded by sensor on the premise of West University of Timișoara. This correlation will generate **radiation brackets** that could be used to train a neural network to predict radiation from satellite imagery alone.

2. I will proceed to analyze when and where are clouds **forming** and **dissipating** in satellite imagery. Observing this and taking into account a corresponding vector field, can result in improving my algorithm to better advect clouds and reflect real world scenarios by deforming, dissipating, and forming new clouds in the scene.

3. I will also investigate using a different technique to predict cloud movement based on the **Boids Flocking Behavior** algorithm. It will be receiving as input the detected clouds at their current position and a time series of the corresponding motion vector field.

4. I will try to extract cloud motion information from orbiting satellite sensors by taking advantage of the **time delay** between acquisitions of different bands. Because of the configuration of satellite sensors, not all bands are capturing simultaneously. The time difference between two bands could be sufficient to detect large movement of clouds because clouds are visible in more than one single spectral band.

5. My final goal is to create a **cloud platform** that delivers cloud masking abilities, motion estimation, and solar radiation predictions. Users of the system will be able to choose or plug in their own methods for detection of clouds, estimation of motion, essentially making the platform customizable to fit the specific needs of each use case.

6 Conclusions

This report presented the state of the art in cloud screening techniques (Section 2), motion estimation methods (Section 3), and solar irradiance research (Section 4). The study areas are vast and new findings that try do deal with each of the three topics are frequently published. The merits of each study are acknowledged and put forward.

Limitations of these methods still impose obstacles that need to be overcome. They are detailed in the conclusions of each section and my objectives are set out in Section 5.

7 Acknowledgement

This work was supported by a grant of the Romanian Ministry of Education and Research, CNCS - UEFISCDI project number PN-III-P1-1.1-TE-2019-0859, within PNCDI III.

References

[1] Tom Andersen. Operational snow mapping by satellites. Number 138, pages 149–154, 1982. cited By 21.

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

[3] J.L. Bosch and J. Kleissl. Cloud motion vectors from a network of ground sensors in a solar power plant. *Solar Energy*, 95:13 – 20, 2013.

[4] J.L. Bosch, Y. Zheng, and J. Kleissl. Deriving cloud velocity from an array of solar radiation measurements. *Solar Energy*, 87:196 – 203, 2013.

[5] T. Brox and J. Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):500–513, 2011.

[6] V. CASELLES and M. J. LÓPEZ GARCÍA. An alternative simple approach to estimate atmospheric correction in multitemporal studies. *International Journal of Remote Sensing*, 10(6):1127–1134, 1989.

[7] Y Changhui, Yuan Yuan, Miao Minjing, and Zhu Menglu. Cloud detection method based on feature extraction in remote sensing images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2:W1, 2013.

[8] Nan Chen, Wei Li, Charles Gatebe, Tomonori Tanikawa, Masahiro Hori, Rigen Shimada, Teruo Aoki, and Knut Stamnes. New neural network cloud mask algorithm based on radiative transfer simulations. *Remote Sensing of Environment*, 219:62 – 71, 2018.

[9] P. Y. Chen, R. Srinivasan, G. Fedosejevs, and B. Narasimhan. An automated cloud detection method for daily noaa-14 avhrr data for texas, usa. *International Journal of Remote Sensing*, 23(15):2939–2950, 2002.

[10] Hsu-Yung Cheng and Chih-Chang Yu. Multi-model solar irradiance prediction based on automatic cloud classification. *Energy*, 91:579 – 587, 2015.

[11] Hsu-Yung Cheng, Chih-Chang Yu, and Sian-Jing Lin. Bi-model short-term solar irradiance prediction using support vector regressors. *Energy*, 70:121 – 127, 2014.

[12] Chi Wai Chow, Serge Belongie, and Jan Kleissl. Cloud motion and stability estimation for intra-hour solar forecasting. *Solar Energy*, 115:645 – 655, 2015.

[13] T. Danaher, Peter Scarth, John Armston, Lisa Collett, J. Kitchen, and S. Gillingham. Remote sensing of tree-grass systems: The eastern australian woodlands. *Ecosystem Function in Savannas: Measurement and Modeling at Landscape to Global Scales*, pages 175–194, 01 2010.

[14] M. Derrien, B. Farki, L. Harang, H. LeGléau, A. Noyalet, D. Pochic, and A. Sairouni. Automatic cloud detection applied to noaa-11 /avhrr imagery. *Remote Sensing of Environment*, 46(3):246 – 267, 1993.

[15] J. R. Dymond and J. D. Shepherd. Correction of the topographic effect in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2618–2619, 1999.

[16] N.A. Engerer and F.P. Mills. Kpv: A clear-sky index for photovoltaics. *Solar Energy*, 105:679 – 693, 2014.

[17] H. Escrig, Francisco Batlles, Joaquín Alonso-Montesinos, F.M. Baena, Juan Bosch, I. Salbidegoitia, and Juan Burgaleta. Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast. *Energy*, 55, 06 2013.

[18] Gunnar Farnebäck. Two-frame motion estimation based on polynomial expansion. In Josef Bigun and Tomas Gustavsson, editors, *Image Analysis*, pages 363–370, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[19] David Frantz, Erik Haß, Andreas Uhl, Johannes Stoffels, and Joachim Hill. Improvement of the fmask algorithm for sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment*, 215:471 – 481, 2018.

[20] Nafiseh Ghasemian and Mehdi Akhoondzadeh. Introducing two random forest based methods for cloud detection in remote sensing images. *Advances in Space Research*, 62(2):288 – 303, 2018.

[21] Nicholas Goodwin, Lisa Collett, Robert Denham, Neil Flood, and Daniel Tindall. Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM + time series. *Remote Sensing of Environment*, 134:50–65, 07 2013.

[22] Luis Gómez-Chova, Gustau Camps-Valls, Javier Calpe, Luis Guanter, and Jose Moreno. Cloud-screening algorithm for ENVISAT/MERIS multispectral images. *Geoscience and Remote Sensing, IEEE Transactions on*, 45:4105 – 4118, 01 2008.

[23] O. Hagolle, M. Huc, D. Villa Pascual, and G. Dedieu. A multi-temporal method for cloud detection, applied to formosat-2, venμs, landsat and sentinel-2 images. *Remote Sensing of Environment*, 114(8):1747 – 1755, 2010.

[24] Thomas M Hamill and Thomas Nehrkorn. A short-term cloud forecast scheme using cross correlations. *Weather and Forecasting*, 8(4):401–411, 1993.

[25] Kaiming He and Jian Sun. Statistics of patch offsets for image completion. In Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid, editors, *Computer Vision – ECCV 2012*, pages 16–29, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[26] Berthold K.P. Horn and Brian G. Schunck. Determining Optical Flow. In James J. Pearson, editor, *Techniques and Applications of Image Understanding*, volume 0281, pages 319 – 331. International Society for Optics and Photonics, SPIE, 1981.

[27] Hao Huang, Jin Xu, Zhenzhou Peng, Shinjae Yoo, Dantong Yu, Dong Huang, and Hong Qin. Cloud motion estimation for short term solar irradiation prediction. pages 696–701, 10 2013.

[28] Tsuyoshi Inoue, Tetsuo Sasaki, and Takashi Washio. Spatio-temporal kriging of solar radiation incorporating direction and speed of cloud movement. *The 26th Annual Conference of the Japanese Society for Artificial Intelligence*, JSAI2012:1K2IOS1b3–1K2IOS1b3, 2012.

[29] Richard Irish. Landsat 7 Automatic Cloud Cover Assessment. *Proceedings of SPIE - The International Society for Optical Engineering*, 4049, 08 2000.

[30] Richard Irish, John Barker, Samuel Goward, and Terry Arvidson. Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) algorithm. *Photogrammetric Engineering & Remote Sensing*, 72:1179–1188, 10 2006.

[31] Mohammad Jamaly and Jan Kleissl. Spatiotemporal interpolation and forecast of irradiance data using kriging. *Solar Energy*, 158:407 – 423, 2017.

[32] Mohammad Jamaly and Jan Kleissl. Robust cloud motion estimation by spatio-temporal correlation analysis of irradiance data. *Solar Energy*, 159:306 – 317, 2018.

[33] Suming Jin, Collin Homer, Limin Yang, George Xian, Joyce Fry, Patrick Danielson, and Philip A. Townsend. Automated cloud and shadow detection and filling using two-date landsat imagery in the usa. *International Journal of Remote Sensing*, 34(5):1540–1560, 2013.

[34] Junchang Ju and David P. Roy. The availability of cloud-free landsat etm+ data over the conterminous united states and globally. *Remote Sensing of Environment*, 112(3):1196 – 1211, 2008.

[35] Jan Kleissl. *Solar energy forecasting and resource assessment*. Academic Press, 2013.

[36] K. T. Kriebel, G. Gesell, M. Ka¨stner, and H. Mannstein. The cloud analysis tool apollo: Improvements and validations. *International Journal of Remote Sensing*, 24(12):2389–2408, 2003.

[37] Matthew Lave, Matthew J. Reno, and Robert J. Broderick. Characterizing local high-frequency solar variability and its impact to distribution studies. *Solar Energy*, 118:327 – 337, 2015.

[38] Z Lin, S Stamnes, Z Jin, I Laszlo, S-C Tsay, WJ Wiscombe, and Knut Stamnes. Improved discrete ordinate solutions in the presence of an anisotropically reflecting lower boundary: Upgrades of the disort computational tool. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 157:119–134, 2015.

[39] Ce Liu et al. *Beyond pixels: exploring new representations and applications for motion analysis*. PhD thesis, Massachusetts Institute of Technology, 2009.

[40] Bruce Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision (ijcai). volume 81. Vancouver, British Columbia, 04 1981.

[41] Eduardo Weide Luiz, Fernando Ramos Martins, André Rodrigues Gonçalves, and Enio Bueno Pereira. Analysis of intra-day solar irradiance variability in different brazilian climate zones. *Solar Energy*, 167:210 – 219, 2018.

[42] Yi Luo, Alexander Trishchenko, and Konstantin Khlopenkov. Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment - REMOTE SENS ENVIRON*, 112:4167–4185, 12 2008.

[43] B.L. Markham and J.L. Barker. Landsat MSS and TM Post Calibration Dynamic Ranges, Exoatmospheric Reflectance and at-satellite Temperatures. *Landsat MSS and TM Post-Calibration Dynamic Ranges, Exoatmospheric Reflectances and At-Satellite Temperatures*, pages 3–8, 01 1986.

[44] Maria Martinez-Chico, Francisco Batlles, and Juan Bosch. Cloud classification in a mediterranean location using radiation data and sky images. *Energy*, 36:4055–4062, 07 2011.

[45] John Mecikalski, Patrick Minnis, and Rabindra Palikonda. Use of satellite derived cloud properties to quantify growing cumulus beneath cirrus clouds. *Atmospheric Research*, s 120–121:192–201, 02 2013.

[46] Steven D. Miller, Matthew A. Rogers, John M. Haynes, Manajit Sengupta, and Andrew K. Heidinger. Short-term solar irradiance forecasting via satellite/model coupling. *Solar Energy*, 168:102 – 117, 2018. Advances in Solar Resource Assessment and Forecasting.

[47] G. MOUSSU, L. DIABATE, D. OBRECHT, and L. WALD. A method for the mapping of the apparent ground brightness using visible images from geostationary satellites. *International Journal of Remote Sensing*, 10(7):1207–1225, 1989.

[48] U Müller-Wilm. Sen2cor configuration and user manual (s2-pdgs-mpc-l2a-sum-v2. 5.5), 2018.

[49] Hans Arnold Panofsky and Glenn Wilson Brier. *Some applications of statistics to meteorology*. Mineral Industries Extension Services, College of Mineral Industries . . . , 1958.

[50] Philippe Pierre Pebay. Formulas for robust, one-pass parallel computation of covariances and arbitrary-order statistical moments. Technical report, Sandia National Laboratories, 2008.

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

[52] Zhenzhou Peng, Dantong Yu, Dong Huang, John Heiser, and Paul Kalb. A hybrid approach to estimate the complex motions of clouds in sky images. *Solar Energy*, 138:10 – 25, 2016.

[53] M. Penteliuc and M. Frincu. Prediction of cloud movement from satellite images using neural networks. In *2019 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, pages 222–229, Sep. 2019.

[54] R. Perez, P. Ineichen, R. Seals, and A. Zelenka. Making full use of the clearness index for parameterizing hourly insolation conditions. *Solar Energy*, 45(2):111 – 114, 1990.

[55] Richard Perez, Mathieu David, Thomas E. Hoff, Mohammad Jamaly, Sergey Kivalov, Jan Kleissl, Philippe Lauret, and Marc Perez. Spatial and temporal variability of solar energy. *Foundations and Trends® in Renewable Energy*, 1(1):1–44, 2016.

[56] Richard Perez, Robert Seals, Pierre Ineichen, Ronald Stewart, and David Menicucci. A new simplified version of the perez diffuse irradiance model for tilted surfaces. *Solar Energy*, 39(3):221 – 231, 1987.

[57] R.T. Pinker, I. Laszlo, J.D. Tarpley, and K. Mitchell. Geostationary satellite parameters for surface energy balance. *Advances in Space Research*, 30(11):2427 – 2432, 2002.

[58] Shi Qiu, Binbin He, Zhe Zhu, Zhanmang Liao, and Xingwen Quan. Improving fmask cloud and cloud shadow detection in mountainous area for landsats 4–8 images. *Remote Sensing of Environment*, 199:107 – 119, 2017.

[59] Shi Qiu, Zhe Zhu, and Binbin He. Fmask 4.0: Improved cloud and cloud shadow detection in landsats 4–8 and sentinel-2 imagery. *Remote Sensing of Environment*, 231:111205, 2019.

[60] Reoxiang Li, Bing Zeng, and M. L. Liou. A new three-step search algorithm for block motion estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 4(4):438–442, 1994.

[61] George Riggs and Dorothy K Hall. Improved snow mapping accuracy with revised modis snow algorithm. 2012.

[62] R. W. SAUNDERS. An automated scheme for the removal of cloud contamination from avhrr radiances over western europe. *International Journal of Remote Sensing*, 7(7):867–886, 1986.

[63] R. W. SAUNDERS and K. T. KRIEBEL. An improved method for detecting clear sky and cloudy radiances from avhrr data. *International Journal of Remote Sensing*, 9(1):123–150, 1988.

[64] Martin Schlather. Some covariance models based on normal scale mixtures. *Bernoulli*, 16(3):780–797, 08 2010.

[65] Zhenfeng Shao, Juan Deng, Lei Wang, Yewen Fan, Neema S Sumari, and Qimin Cheng. Fuzzy autoencode based cloud detection for remote sensing imagery. *Remote Sensing*, 9(4):311, 2017.

[66] K. Shinozaki, N. Yamakawa, T. Sasaki, and T. Inoue. Areal solar irradiance estimated by sparsely distributed observations of solar radiation. *IEEE Transactions on Power Systems*, 31(1):35–42, 2016.

[67] E. Shusterman and M. Feder. Image compression via improved quadtree decomposition algorithms. *IEEE Transactions on Image Processing*, 3(2):207–215, 1994.

[68] P. Soille. *Morphological Image Analysis - Principles and Applications*, volume 49. 01 1999.

[69] Pierre Soille, Jürgen Vogt, and Roberto Colombo. Carving and adaptive drainage enforcement of grid digital elevation models. *Water Resources Research*, 39(12), 2003.

[70] Knut Stamnes, S-Chee Tsay, Warren Wiscombe, and Kolf Jayaweera. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Opt.*, 27(12):2502–2509, 1 1988.

[71] Andrew Staniforth and Jean Côté. Semi-Lagrangian Integration Schemes for Atmospheric Models—A Review. *Monthly Weather Review*, 119(9):2206–2223, 09 1990.

[72] L. L. Stowe, C. G. Wellemeyer, H. Y. M. Yeh, T. F. Eck, and The Nimbus-7 CLOUD DATA PROCecessing TEAM. Nimbus-7 Global Cloud Climatology. part I: Algorithms and Validation. *Journal of Climate*, 1(5):445–470, 05 1988.

[73] J.J. Su. Enhanced ACCA Algorithm. *Space Imaging Corporation Technical Memo IT81-LSD-SA&E Memo*, 01 1984.

[74] Narayanan Sundaram, Thomas Brox, and Kurt Keutzer. Dense point trajectories by gpu-accelerated large displacement optical flow. In Kostas Daniilidis, Petros Maragos, and Nikos Paragios, editors, *Computer Vision – ECCV 2010*, pages 438–451, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[75] Gary E Thomas and Knut Stamnes. *Radiative transfer in the atmosphere and ocean*. Cambridge University Press, 2002.

[76] Bin Tian, Mukhtiar A Shaikh, Mahmood R Azimi-Sadjadi, Thomas H Vonder Haar, and Donald L Reinke. A study of cloud classification with neural networks using spectral and textural features. *IEEE transactions on neural networks*, 10(1):138–151, 1999.

[77] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, 4(11):1549–1560, 1995.

[78] E. Vermote and N. Saleous. LEDAPS surface reflectance product description, version 2.0. 2007.

[79] C. L. Walthall, J. M. Norman, J. M. Welles, G. Campbell, and B. L. Blad. Simple equation to approximate the bidirectional reflectance from vegetative canopies and bare soil surfaces. *Appl. Opt.*, 24(3):383–387, 2 1985.

[80] B. Wang, A. Ono, K. Muramatsu, and N. Fujiwarattt. Automated detection and removal of clouds and their shadows from landsat tm images. *IEICE Transactions on Information and Systems*, E82-D(2):453–460, 1999. cited By 89.

[81] Martijn Wildt, Gabriela Seiz, and Armin Gruen. Operational snow mapping using multitemporal Meteosat SEVIRI imagery. *Remote Sensing of Environment*, 109:29–41, 07 2007.

[82] Y Zhang, B Guindon, and Josef Cihlar. An image transform to characterize and compensate for spatial variations in thin cloud contamination of landsat images. *Remote Sensing of Environment*, 82:173–187, 10 2002.

[83] Zhe Zhu, Shixiong Wang, and Curtis E. Woodcock. Improvement and expansion of the fmask algorithm: cloud, cloud shadow, and snow detection for landsats 4–7, 8, and sentinel 2 images. *Remote Sensing of Environment*, 159:269 – 277, 2015.

[84] Zhe Zhu and Curtis Woodcock. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118:83–94, 03 2012.

[85] Zhe Zhu and Curtis E. Woodcock. Automated cloud, cloud shadow, and snow detection in multitemporal landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152:217 – 234, 2014.