Cloud Detection Using Geometry Point Square Pixel Algorithm to Support Aviation for Flight Safety

Yenniwarti Rafsyam
Department of Electrical Engineering, Politeknik Negeri Jakarta

Eri Prasetyo Wibowo
Department of Information Technology, Gunadarma University

Danang Surya Candra
National Research and Innovation Agency (BRIN)

Aini Suri Talita
Department of Information Technology, Gunadarma University

他

https://doi.org/10.5109/7326957

出版情報：Evergreen. 11 (4), pp.3213-3232, 2024-12. Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Japan
バージョン：

Cloud Detection Using Geometry Point Square Pixel Algorithm to Support Aviation for Flight Safety

Yenniwarti Rafsyam[1,*], Eri Prasetyo Wibowo[2], Danang Surya Candra[3], Aini Suri Talita[2], Shita Fitria Nurjihan[1], Arief Rinaldi[1]

[1]Department of Electrical Engineering, Politeknik Negeri Jakarta, Indonesia

[2]Department of Information Technology, Gunadarma University, Indonesia

[3]National Research and Innovation Agency (BRIN), Indonesia.

*Author to whom correspondence should be addressed:
E-mail: yenniwarti.rafsyam@elektro.pnj.ac.id

(Received July 4, 2024; Revised October 7, 2024; Accepted November 12, 2024).

Abstract: Cumulonimbus clouds (CB) are a specific type of cumulus cloud characterized by their towering, dense formations and their association with thunderstorms, resulting from unstable atmospheric stratification. In aviation, it is crucial to avoid Cumulonimbus clouds because passing through them can pose several dangers. These include lightning, which can disrupt electrical components, hail that can cause structural damage, strong convection and wind shear that create unsafe flight conditions or even lead to stalling, and icing, which alters the aircraft's aerodynamics. Monitoring and steering clear of Cumulonimbus clouds is, therefore, vital for ensuring flight safety. The detection of Cumulonimbus (CB) and non-Cumulonimbus (NonCB) clouds can be accomplished using satellite imagery, particularly the NOAA satellite data. In this study, we employed the Geometry Point Square Pixel algorithm (GPSP) to distinguish between these cloud types in the Indonesian region. To facilitate this distinction, we first determined the geographical location of airports within the satellite imagery using the geometry point algorithm. The airport area was defined with a 0.04-degree distance for each latitude and longitude using the rectangle polygon method. Subsequently, a pixel search was conducted, employing a threshold algorithm to identify CB and NonCB clouds at the detected airport points within the satellite imagery. The results of this algorithm yielded an average accuracy of over 86,6% in cloud detection for the Indonesian region, thus proving its effectiveness in differentiating between Cumulonimbus and non-Cumulonimbus clouds from NOAA satellite imagery.

Keywords: cloud; detection; geometry point; pixel; rectangle polygon; threshold

1. Introduction

Clouds appear in an unlimited variety of shapes and sizes, and meteorology categorizes them into distinct primary conditions. In his 1803 essay, 'Modifications of Clouds,' Luke Howard introduced a classification system that divides clouds into three main types: cirrus, cumulus, and stratus. Specifically, cumulus clouds are characterized by their generally flat bases, formed at altitudes where rising air vapor condenses[1)]. Cumulonimbus clouds are a type of cumulus cloud, characterized by their towering vertical formation, dense structure, and involvement in thunderstorms and other severe weather conditions. These clouds are also one of the earliest formations of tornadoes, with a rare phenomenon known as chaos[2)]. Tornadoes are winds caused by the impact of Cumulonimbus clouds, typically occurring during the rainy season[3)].

In aviation, understanding Cumulonimbus (CB) clouds is highly relevant for flight safety. Theories developed about CB clouds highlight various potential hazards that may be faced by aircraft and passengers. Factors such as strong turbulence, hail, intense lightning activity, unstable weather conditions in tropical regions, extreme altitudes, and phenomena like microbursts and fog formation are all aspects that need to be seriously considered by pilots and airlines. Recognizing these risks is crucial in making decisions regarding flight routes and implementing appropriate preventive measures. With a better understanding of these theories, we can enhance flight safety and reduce potential risks associated with CB clouds[4)].

Cumulonimbus clouds are dangerous to aviation as they can cause severe thunderstorms in the atmosphere. Therefore, monitoring CB clouds is crucial to avoid these dangerous zones during flights. Detection of CB clouds can be done in various ways, using satellite images or ground-based images. One effective way to monitor cloud and weather conditions is through the use of NOAA (National Oceanic and Atmospheric Administration) satellites. NOAA satellites are equipped with various sensors and remote sensing equipment capable of capturing high-resolution images of Earth's atmosphere. Data collected by these satellites, such as infrared images and sea surface temperature measurements, allow researchers and meteorologists to track cloud patterns, identify developing CB clouds, and monitor overall weather conditions. With the help of NOAA satellites, the weather information obtained can be used to provide early warnings about potential weather hazards, including CB clouds that pose a risk to aviation[5]. Thus, the use of NOAA satellites is a valuable tool in efforts to maintain flight safety and protect lives and property from the impacts of extreme weather.

However, despite the many advantages of NOAA satellites, one major drawback in weather prediction is their inability to provide direct information about atmospheric conditions at very specific locations. NOAA satellites are generally used to observe larger weather patterns and climate phenomena, such as tropical storms, global wind patterns, and long-term climate changes. Although the data collected by NOAA satellites can provide valuable insights for meteorologists, additional information from other weather monitoring systems, such as weather radar and ground observation stations, is often needed to accurately predict weather at local or regional levels. Identifying characteristics that indicate potential tornadoes is often challenging, especially when assessing the impact of Cumulonimbus clouds. Satellite imagery is a valuable tool for this purpose, as it can reveal key indicators. For example, significant temperature differences and the presence of Cumulonimbus clouds with striking gray boundaries that exceed the determined intensity color limits can be prominent signs[4),5)].

In this study, we enhanced satellite image results using the Geometry Point Square Pixel (GPSP) algorithm, which allows for distinguishing cloud types. To facilitate this distinction, we first determine the geographic location of airports in satellite images using a geometric point algorithm. The airport area is defined with a distance of 0.04 degrees for each latitude and longitude using the rectangular polygon method. Subsequently, pixel searches are conducted using the threshold algorithm to identify CB and NonCB clouds at detected airport points in satellite images. The test results are strictly validated using data from Meteorological Reports (METAR). With this approach, we can increase the accuracy in detecting clouds that may pose a risk to aviation, thereby aiding in making flight safety decisions in Indonesia. This not only contributes to improving aviation transportation safety by identifying hazardous zones but also lays the groundwork for potential applications in automatic detection and reporting of CB clouds through METAR in the future

2. Literature Survey

Many studies have explored cloud detection using various methods, including satellite and ground-based image analysis[6)-11)]. Qingyong Li introduced a practical ground-based cloud image detection approach using the Hybrid Thresholding Algorithm (HYTA). This algorithm converts input color cloud images into normalized blue/red channel images that maintain contrast differences, even with noise and outliers. HYTA then identifies the image ratio as unimodal or bimodal using fixed and minimum cross entropy (MCE) thresholding algorithms. The study found that HYTA achieved an accuracy of 88.53%, significantly higher than improved thresholding or MCE accuracy. HYTA has proven to outperform other advanced cloud detection approaches[6)].

Kohei Arai introduced a threshold-based method to detect rain clouds using NOAA/AVHRR data with the Jacobi iteration method. The proposed method was compared with data from the Japan Meteorological Agency derived from radar data. However, experimental results showed that the regression performance was not very good. The Root Mean Square Difference between the two binary images of detected radar rainfall rates and NOAA/AVHRR detected rainfall areas was 13.727. Thus, the proposed method has marginal rainfall detection accuracy[7)].

Radha Rani et al. developed a threshold-based technique to detect clouds in Bangladesh using NOAA-AVHRR satellite time series data. This algorithm was applied to daytime NOAA-17 for significant seasons in Bangladesh to detect clouds. The method's efficiency in detecting cloudy pixels is sufficiently high for operational use[8)]. Several studies related to Cumulonimbus cloud detection have also been presented to avoid hazardous zones[12-21)]. Wanayumini et al. presented tornado occurrence predictions using an edge detection algorithm approach to obtain gray boundaries to extract irregular pattern intensification on Cumulonimbus clouds[12)]. Other research shows that early detection of Cumulonimbus cloud layers, which often precede turbulent weather phenomena, can aid in tornado prediction. This is achieved using a Supervised Image Detection Algorithm with a Spectral Angle Mapper approach to establish minimum and maximum pixel intensity interval values based on spectral angles associated with Cumulonimbus clouds[13)].

Cumulonimbus cloud detection from infrared (IR) satellite images is conducted using a spatiotemporal detection method that leverages the distinct spatial characteristics of CB clouds. The presented technique does not require labeled data such as shape, temperature, texture, and gradient. The study incorporates temporal features into the multiscale spatial anomaly detection algorithm. The addition of temporal features in this study is significant and produces better results compared to spatial anomaly detection algorithms alone[14)]. Cumulonimbus cloud prediction has also been carried out using the European Medium-Range Weather Forecasting Integrated Forecast System (IFS) model for the aviation information region in Jakarta and Ujung Pandang using convective cloud (CC) cover calculations from precipitation products[18]. Additionally, radiosonde observations can also be used to predict Cumulonimbus clouds in short-term weather forecasting periods. Radiosonde data prediction models are trained using machine learning[19].

3. Proposed Methodology

3.1 Geometry Point Square Pixel Algorithm (GPSP)

2D digital images, whether captured or created on a flat surface, are typically defined using a limited data structure represented as a planar orthogonal grid with regularly spaced units. Pixels in these images can store various types of values, including integers, floating-point numbers, vectors, or even sets. For example, in color images, the significance of pixels is often represented by three scalar values such as red, green, and blue (RGB), or hue, saturation, and intensity (HSI).

The Geometric Pixel Square Point (GPSP) algorithm plays a crucial role in the processing and analysis of these 2D digital images. This algorithm enables efficient calculation and manipulation of pixel positions and values, considering the structure of the orthogonal planar grid. With GPSP, various geometric operations such as rotation, scaling, and translation can be performed with high precision, which is essential in computer graphics and image processing applications.

At a basic level, pixels have coordinates based on the grid, which is our way of representing points in two-dimensional space. Each pixel in a digital image is placed at the intersections of these grid lines and given integer coordinates (x, y) that uniquely determine its position. To facilitate data processing and analysis, the default use of integer coordinates has become standard. This is because integer coordinates make it easier to identify patterns in regular orthogonal arrangements.

The use of integer coordinates has several advantages. First, it allows for faster and more efficient data

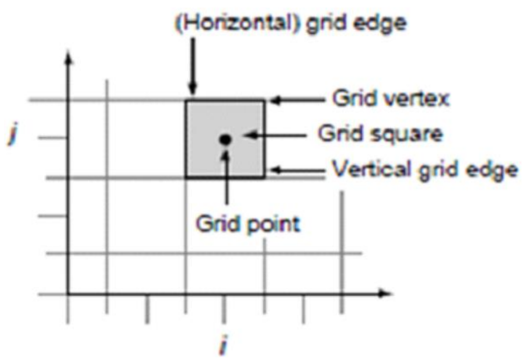


Fig. 1: Grid point and grid square.[22]

Furthermore, the use of integer coordinates also supports the implementation of simpler and more understandable algorithms. Many classical image processing algorithms, such as the Geometric Pixel Square Point (GPSP) algorithm, rely on the regularity and simplicity offered by integer coordinates. For example, GPSP leverages the structure of the orthogonal planar grid to perform geometric operations with high efficiency. In this way, various operations such as rotation, scaling, and translation can be carried out with high precision, which is crucial in computer graphics and image processing applications.

Overall, integer coordinates in pixel representation provide a strong foundation for various image processing techniques. They not only facilitate the implementation and efficiency of algorithms but also ensure that the results of image processing remain accurate and consistent. In the context of the continuously evolving development of graphic technology and image analysis, a deep understanding of the importance of integer coordinates becomes increasingly relevant.

The equation represents the set of all integer coordinates (see Equation 1-3). Each grid point in this space is significant as it marks the center of a grid square, whose sides have a length of one unit. These grid squares are parallel to the Cartesian coordinate axes, facilitating the visualization and analysis of available data (see Fig. 1). Additionally, the corner points of each grid square, known as grid nodes, play a crucial role in understanding the overall grid structure.

A solid understanding of the concept of grid coordinates is key to effective data manipulation and in-depth analysis in the context of digital representation. In digital representation, each pixel in an image can be viewed as a point in, with integer coordinates determining its position. The use of grid coordinates ensures that each pixel has a unique and clear location, which is crucial for image processing operations such as edge detection, segmentation, and geometric transformation.

Integer coordinates facilitate various mathematical and algorithmic operations. For example, in image transformations such as rotation or scaling, integer coordinates enable simpler and more efficient calculations. Additionally, patterns in images, such as lines or geometric shapes, can be more easily recognized using these integer coordinates.

The use of grid coordinates of also allows for consistent and accurate representation of digital images. In the context of image processing, each pixel carries important information that can be color, intensity, or other related data. With integer coordinate representation, this information can be accessed and manipulated in a structured and systematic way.

Furthermore, the corner points of grid squares, or grid nodes, provide important references in image processing. For example, in edge detection algorithms, grid nodes can be used to identify sudden changes in intensity or color, marking the edges of objects in the image. Overall, integer coordinates provide a strong foundation for various image processing and analysis techniques. They enable efficient and accurate data processing, ensuring that the results of image analysis remain consistent and reliable. In the development of graphic technology and image analysis, a deep understanding of the importance of grid coordinates becomes increasingly relevant and crucial.

A 2D grid of size m × n is a rectangular array of grid points, expressed as:

$$G_{m,n} = \{(i,j) \in \mathbb{Z}^2 : 1 \le i \le m \wedge 1 \le j \le n\} \qquad (1)$$

or a rectangular set of grid squares

$$G_{m.n} = \{\text{grid square } c : (i,j) = center\ of\ c \wedge 1 \le i \le m \wedge 1 \le j \le n\} \qquad (2)$$

where m, n >> 1 (read: "is much larger than 1").

The set of grid points in 2D images is defined in (i, j) ∈ , with i the longitude coordinate and j the latitude coordinate, while a grid square is defined by the four-square edges that make up the square. The squares are defined by quadruple edge squares, Ag, where successive edges (modulo 4) share vertices, using the equation:

$$Ag = (i - \theta_W, j - \theta_S)\ (i + \theta_E, j - \theta_S)\ (i + \theta_E, j + \theta_N)\ (i - \theta_W, j + \theta_N) \qquad (3)$$

With θ_W a and θ_E the longitude angles to the West and the East sides, respectively, and θ_S and θ_N the latitude angles to the South and the North sides, respectively.

The Digital Geometry Point algorithm, rectangular polygon, and threshold pixel algorithm are used to detect Cumulonimbus (CB) and NonCB clouds. CB and NonCB clouds are detected in the Indonesian region using NOAA satellite data, which is validated with METAR data from BMKG (Meteorology, Climatology, and Geophysics Agency). The cloud detection process is carried out with the following steps: first, NOAA satellite image data is collected for the Indonesian region, including visual and infrared information. Then, airport points are determined using the geometry point algorithm to identify the airport positions on the satellite image. Next, the area around the airport is determined by creating a rectangular polygon based on the airport center coordinates, with $\theta_W = \theta_E = \theta_S = \theta_N = 0.04$ degrees from the airport center.

After the airport area is determined, the threshold algorithm is used to examine pixels in that area. Pixels that meet certain threshold criteria are identified as part of CB or NonCB clouds. Based on the detected pixels, classification is performed to determine the type of cloud. The cloud detection results are then validated using METAR data from BMKG, which includes actual weather reports from meteorological stations to verify the accuracy of cloud detection from satellite images. In detail, the detection process for CB and NonCB clouds is explained as follows:

3.2 Algorithm 1 Dataset creation (satellite imagery overlay)

Clouds are the main indicator in determining weather conditions in an area and each type of cloud has a different meaning. Researchers apply various methods for cloud detection. Cloud detection methods based on remote sensing images are divided into three categories, namely methods based on band thresholds, methods based on machine learning and methods based on deep learning[32].

To detect clouds, researchers have developed various methods using remote sensing image technology. These methods can be categorized into three main groups: threshold-based band methods, machine learning-based methods, and deep learning-based methods. The choice of the appropriate detection method depends heavily on the needs and objectives of the desired weather analysis[33][34].

Threshold-based band methods are approaches that utilize specific threshold values to separate areas considered as clouds from the background of the image. Meanwhile, machine learning-based methods involve the use of machine learning algorithms to recognize patterns associated with the presence of clouds in the image. On the other hand, deep learning-based methods utilize artificial neural networks or other techniques that allow the system to automatically learn from data to detect and classify clouds with higher accuracy[35].

This research applies the threshold method using NOAA satellite image data based on the results of research conducted by previous researchers who also used the threshold method, namely Osaka[32] which was used to obtain information on rapidly developing cumulus areas (RDCAs) where cumulus clouds are estimated to have the potential to develop into storms lightning. The results of the research statistically show that it is necessary to re-determine the threshold value so that it can provide a picture that can show the life cycle situation of Cumulonimbus clouds. Yasuhiko[20], identified cumulus, Cumulonimbus and cirrus areas by developing the Convective cloud information (CCI) method using a grid template measuring 2t+1 and 2s+1 with center points x0, y0 using Himawari image data for aviation safety. The research results show that the method used is able to detect cloud areas ranging from developing cumulus to mature Cumulonimbus by 69%.

The focus of this research is on airport locations in Indonesia, in accordance with guidelines set by the International Civil Aviation Organization in 2018 (ICAO)[36]. This research begins with creating a dataset to overcome the unavailability of datasets that can be directly processed and labeled datasets, namely labeling datasets. Making the first dataset from satellite imagery requires data input in the form of NOAA satellite image L1B, point coordinates from the airport location, and METAR data from BMKG. From NOAA's satellite imagery, supporting data is collected from the BMKG's METAR and airport coordinates data. Then overlay the satellite imagery with the squares around the airport points. The output of creating the first dataset is a collection of satellite images at each airport along with METAR data. The dataset is made following the block diagram shown in Fig. 2.

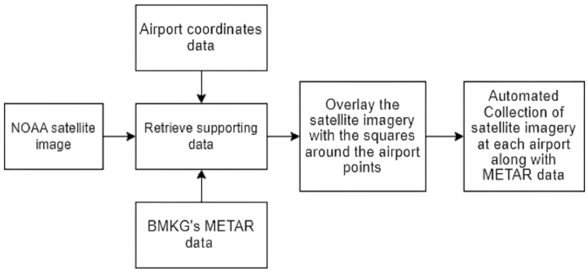


Fig. 2: Satellite image overlay block diagram.

The first step in this process is to acquire NOAA satellite imagery as the starting point. This imagery is the primary data that will be used in creating the dataset. Next, necessary supporting data, such as airport coordinates and METAR data from the Meteorology, Climatology, and Geophysics Agency (BMKG) or other weather data sources, are gathered. Airport coordinate data is important for determining the exact location where the satellite imagery will be overlaid later, while METAR data will provide actual information about the weather conditions at specific times.

After the supporting data is collected, the next step is to overlay the satellite imagery on the relevant airport location. This is done taking into account the previously established coordinates. This overlaying process ensures that the imagery covers the necessary area around the airport for further analysis.

The final step is to merge the satellite imagery with the corresponding METAR data for each airport. This involves combining information from both data sources so that the resulting dataset includes satellite imagery and relevant METAR weather data for each analyzed airport location. Thus, the first dataset creation process successfully integrates information from various sources to provide a comprehensive and relevant dataset for analyzing cloud patterns around airports and current weather conditions.

Based on Fig. 2, the output obtained is a directory containing METAR files (.txt), raster image files (.tif) and airport vector files (.gpkg). Based on the block diagram presented in Fig. 2, the output of the process is a directory containing several types of files, including METAR files (.txt), raster image files (.tif), and airport vector files (.gpkg). The first step in creating this dataset involves extracting airport coordinate data from Airport Meteorological Reports (METAR) provided by the Meteorology, Climatology, and Geophysics Agency (BMKG). This coordinate data is crucial as it will be used to determine the geographical locations of each airport focused on in this study.

The process of extracting airport coordinate data (latitude, longitude) from Airport Meteorological Reports (METAR) issued by the Meteorology, Climatology, and Geophysics Agency (BMKG) is carried out utilizing geometry points in raster data, using the following algorithm:

$$y\_lat = y\_lat * \left(y\_lat\_d + \left(\frac{y\_lat\_m}{60}\right) + \left(\frac{y\_lat\_s}{3600}\right)\right)$$

$$x\_lat = x\_lat * \left(x\_lat\_d + \left(\frac{x\_lat\_m}{60}\right) + \left(\frac{x\_lat\_s}{3600}\right)\right) \quad (4)$$

In both equation, we convert latitude and longitude coordinates from the degree-minute-second (DMS) format to the decimal format. In the DMS format, coordinates are represented by three components: degrees (D), minutes (M), and seconds (S). For example, a latitude coordinate can be expressed as y_lat_d y_lat_m, y_lat_s and a longitude coordinate as x_lat_d, x_lat_m. x_lat_s. To convert from DMS format to decimal, we use the following equation:

$$Decimal = Degrees + \frac{Minute}{60} + \frac{Second}{3600} \quad (5)$$

In the context of latitude coordinates:

$$\mathcal{Y}_{\_lat} = \mathcal{Y}_{\_lat} * \left(y\_lat\_d + \left(\frac{y\_lat\_m}{60}\right) + \left(\frac{y\_lat\_s}{3600}\right)\right) \quad (6)$$

In the context of longitude coordinates:

$$x\_lat = x\_lat * \left(x\_lat\_d + \left(\frac{x\_lat\_m}{60}\right) + \left(\frac{x\_lat\_s}{3600}\right)\right) \quad (7)$$

In these equations, we multiply the degree value by 1, the minute value by 1/60, and the second value by 1/3600 to obtain the value in decimal format. Then, we multiply this result by the original coordinate value, which was in the DMS format. The result of both equations is latitude and longitude coordinates in decimal format.

In this algorithm, y_lat and x_lat represent latitude and longitude coordinates that have been converted to decimal format. These coordinate data are then added to the lat and long arrays using the np.append() function. Afterward, the newly found coordinate values are inserted into the appropriate row of a DataFrame. Finally, the coordinate data is stored in list format for ease of further use.

The algorithm used to extract airport coordinate data utilizes the geometric points found in raster data. In the presented code, the variables lat and long are used to store arrays of latitude and longitude coordinates, respectively. Whenever a new coordinate is found, they are added to the corresponding array. Additionally, this data is also inserted into the station_df DataFrame for easier management and further use.

```
lat                 = np.append(lat, y_lat)
long                = np.append(long, x_lat)
row['lat']          = y_lat
row['long']         = x_lat
station_df['lat']   = lat.tolist()
```

station_df['long'] = long.tolist()

This code aims to extract and store latitude and longitude coordinate data from BMKG METAR reports into an appropriate data structure. The process begins by adding the converted latitude (y_lat) and longitude (x_lat) coordinates into existing arrays, `lat` and `long`, using the `np.append()` function from NumPy. This step allows for the incorporation of new coordinates with existing ones in the arrays.

Next, the newly added coordinates are also stored in the corresponding data row, where `row['lat']` is filled with the value (y_lat) and `row['long']` is filled with the value (x_lat). This helps maintain the association between latitude and longitude coordinate data within a single data entry. Finally, the updated `lat` and `long` arrays are converted into lists using the `tolist()` method from NumPy and stored in the DataFrame `station_df` in the 'lat' and 'long' columns. This step ensures that all latitude and longitude coordinates are stored in a format that can be accessed and further analyzed within the context of tabular data.

To determine the area of the airport, the method of rectangular polygon with the subdivision of the area into square vectors with a difference of 0.04-degrees in each latitude and longitude line is utilized. This approach is designed to address limitations in computation and expedite the process of retrieving Digital Number (DN) values from five NOAA satellite image bands. In this method, the area to be measured is divided into small squares called vector boxes. Each of these boxes has a 0.04-degree difference in latitude and longitude lines, ensuring that every area within the region is adequately covered. These geographic coordinates are then converted into map projection systems such as UTM (Universal Transverse Mercator) for ease of area calculation.

With the projection system, the area of each box can be calculated, and the total area of the airport region is obtained by summing the areas of all individual boxes. The selection of a 0.04-degree difference is crucial to balance between resolution and computational needs. This size is small enough to provide good detail but large enough to reduce the number of boxes that need to be processed, thereby reducing the computational load when processing satellite images.

DN (Digital Number), which represents the brightness level values of pixel satellite images in the five bands, can be retrieved more quickly because the number of pixels to be accessed and processed is reduced with this box size. By reducing the number of boxes through the selection of a 0.04-degree difference, the speed in reading and processing data increases, which is crucial in satellite image analysis requiring speed and efficiency.

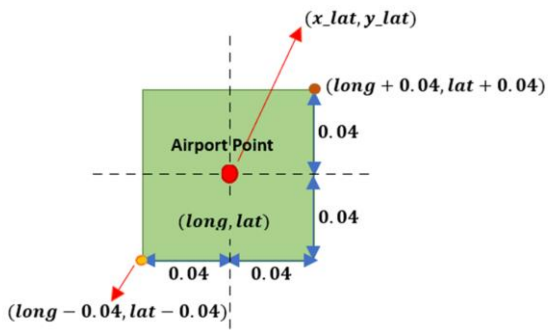


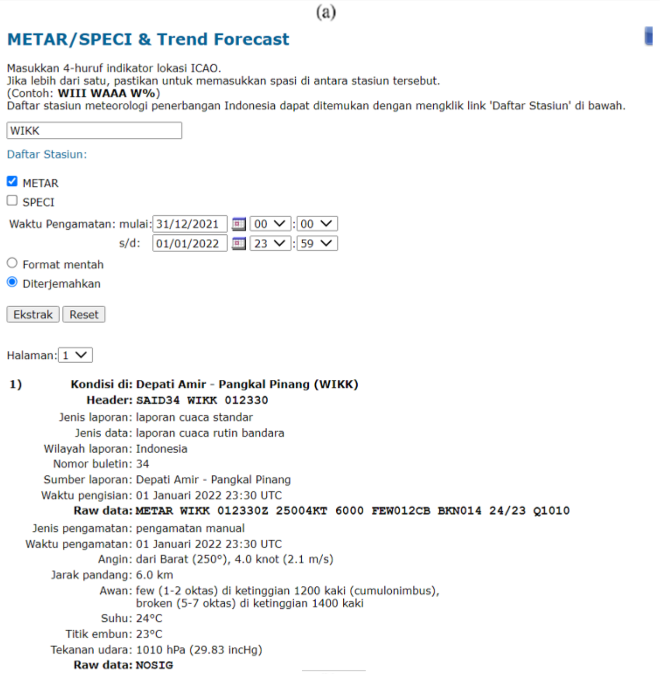
Fig. 3: Template the area of the airport with rectangle polygons.

This rectangular polygon method is based on the basic principles of geometry, where the area is divided into several simple shapes whose areas can be easily calculated. The use of geographic coordinates and projection systems ensures accuracy in area calculations. The utilization of DN values from satellite images to analyze various aspects of the area is also an integral part of this method, allowing for specific analysis of satellite data. With this method, the area of the airport region can be determined effectively and efficiently, considering computational limitations and the need to process satellite image data quickly.

Creating square vectors to determine the area of the airport with a 0.04-degree difference on each latitude and longitude line using the rectangular polygon method, as shown in Fig. 3. This 0.04-degree difference is considered to address computational limitations in reading and processing images and to expedite the process of retrieving DN values for the five NOAA satellite image bands.

After retrieving METAR data through web scraping from the BMKG website, The difference between update rates of METAR (usually every 30 minutes) and NOAA satellite imagery. Figure 4 shows an example of METAR data. Fig. 4. Example of METAR data with format (a) Raw data (b) Translated format at Depati Amir Airport - Pangkal Pinang where CB clouds are predicted on January 1, 2022, at 23:30 UTC or local Indonesian time on February 1, 2022, at 06:30 WIB.

(a)



(b)

Fig. 4: Example of METAR data with (a) Raw data format. (b) Translated format at Depati Amir – Pangkal Pinang airport where CB clouds are predicted on January 1 2022 at 23:30 UTC or Indonesian local time on 01-02-2022 at 06.30 WIB.

Overlaying satellite images with rectangular polygons around the airport, as illustrated in Fig. 5, aims to obtain a collection of satellite images for each airport point. The process above is repeated for each satellite image data, resulting in a set of satellite images at each airport, along with METAR data. The above process is repeated for each satellite image data so that a collection of satellite imagery at each airport, along with METAR data.

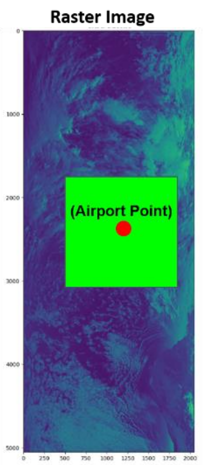


Fig. 5: Overlay satellite image with rectangle polygon.

3.3 Algorithm 2 Retrieving Digital Number (DN) values and cleaning data

The next process is creating a second dataset (training data frame/CSV). The block diagram of the second dataset is shown in Fig. 6.

Fig. 6: Block diagram of retrieving Digital Number values and cleaning data.

In the previous process, satellite image data, overlay data, and METAR data have been obtained. Collecting satellite image data, overlay data, and METAR data in the previous process is crucial for several reasons. First, satellite image data provides visual and spectral information about weather conditions and cloud cover over a wide area, enabling more comprehensive analysis. Second, overlay data provides important geospatial context, such as airport locations and observation area boundaries, allowing for the correlation of weather information with specific locations.

Third, METAR data, which are routine weather reports from airports, provide highly detailed and up-to-date weather observation data, including important information such as visibility, cloud conditions, and other weather phenomena. By combining these three types of data, we can enhance the accuracy and reliability of weather analysis. Satellite images provide broad and in-depth information, overlays assist in precise mapping, and METAR data provide detailed and verified observational data. This process enables better detection and analysis of weather phenomena, supporting more informed decision-making in various applications such as aviation, disaster planning, and climate research. Furthermore, integrating this data allows for the use of complex analytical algorithms to process and interpret data, resulting in deeper and more accurate insights into observed atmospheric conditions.

Raw METAR data is then converted into CSV format or a dataframe. Converting raw METAR data into CSV format or a dataframe is a crucial step in weather data analysis. Raw METAR data is usually in an unstructured text format, making it difficult to process and analyze directly. By converting this data into CSV format or a dataframe, we can utilize it more effectively in various data analysis and programming software such as Python or R.

The CSV or dataframe format has a well-organized tabular structure, where each row represents a set of observation data and each column represents a specific variable such as temperature, visibility, wind speed, and cloud conditions. This facilitates processes such as filtering, sorting, aggregation, and statistical analysis. Additionally, this format is compatible with various analytical and data visualization tools, allowing for deeper data exploration and the generation of informative visualizations.

By converting METAR data into CSV format or a dataframe, we can also integrate it with other data such as satellite images and geospatial overlays, supporting more holistic and comprehensive analysis. This conversion process also enables automation in data processing, allowing analyses to be conducted more efficiently and consistently. Overall, this conversion is an important step to ensure that weather data can be analyzed accurately, efficiently, and integrated with other data sources to generate deeper and more useful insights. From the airport overlay data, the largest data is taken as a representation of weather or clouds at each airport. The darkest or thickest color is considered to have a high DN value, as shown in Fig. 7.

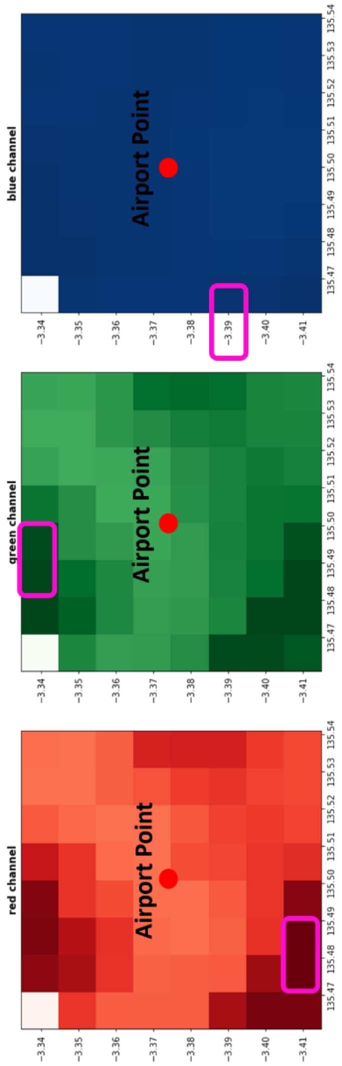


Fig. 7: Retrieval of pixel values on satellite imagery.

The process of selecting the largest data from airport overlays as a representation of weather or clouds at each airport aims to simplify and improve the accuracy of weather observations. In the overlay data, each airport may have various weather observations from several points around the area. By selecting the largest data, we ensure that we take the most significant value that likely represents the actual weather conditions at that airport.

Choosing the darkest or thickest color as an indication of high Digital Number (DN) values is based on principles of digital image processing. In satellite images, DN values represent the intensity of reflection or emission from the Earth's surface or atmosphere. Darker or thicker colors on the overlay typically indicate areas with higher reflection or emission, often associated with thick cloud formations or severe weather. Therefore, using these colors as indicators helps in identifying and measuring weather or cloud intensity around the airport more accurately.

This process is important as it provides a standardized and objective way to assess weather conditions based on satellite image data. By using high DN values as representations, we can leverage satellite data to support weather observations from airports, resulting in more comprehensive and reliable analyses. This is particularly useful in applications such as flight planning and disaster mitigation, where accurate and timely weather information is crucial.

After the data is collected in a CSV file, a preprocessing process is performed, where if the DN is outside the observation limits of the airport on the satellite image, then the data is cleaned by assigning a value of 0 and removing it. Algorithm 2 is used to find airport locations. Each directory produces a file containing METAR data with non-zero DN values representing detected airports. The output of Algorithm 2 is the formation of a dataset stored in the Huge.csv file, ready to be processed or tested using the Geometry Point Square Pixel algorithm.

In the analysis of cloud types, the provided algorithm utilizes the `matplotlib` library to display five image channels within a single figure using subplots.

```
fig, (axr, axg, axb,axi,ax5) = plt.subplots(1,5, figsize=(21,7))
show((src, 1), ax=axr, cmap='Reds', title='channel 1')
show((src, 2), ax=axg, cmap='Greens', title='channel 2')
show((src, 3), ax=axb, cmap='Blues', title='channel 3')
show((src, 4), ax=axi, cmap="Reds", title='channel 4')
show((src, 5), ax=ax5, cmap="Greens", title='channel 5')
plt.show()
```

First, a figure is created with one row and five columns of subplots, where each axis (represented as objects like “axr, axg, axb, axi, and ax5”) corresponds to a specific image channel. Next, the “show()” function is called to display each image channel from the source “src”, with the first channel showing the red component using the “Reds” colormap, the second channel displaying the green component with the 'Greens' colormap, the third channel representing the blue component with the 'Blues' colormap, the fourth channel using a red colormap, and the fifth channel also utilizing a green colormap.

Finally, the command "plt.show()" is executed to display the figure containing all the subplots, allowing for the simultaneous visualization of the image components for further analysis. In this study, this visualization is used to differentiate between cumulonimbus (CB) clouds and NonCB clouds, where the visual analysis of these channels can assist in identifying the characteristics and behaviors of both types of clouds. It is essential to ensure that the "show()"` function has been defined earlier in the context of the algorithm to process and display the images correctly. After executing the above algorithm, each channel in the image will be visualized with different color scales to aid in identifying specific characteristics of the image data, highlighting the further processing results of each channel. This approach allows patterns and differences in data distribution, such as the intensity and location of cumulonimbus (CB) and NonCB clouds, to become more apparent. By utilizing varying color scales, the visualization facilitates a deeper visual interpretation of the structure and composition of clouds based on data from each channel, as will be shown in Fig. 8.

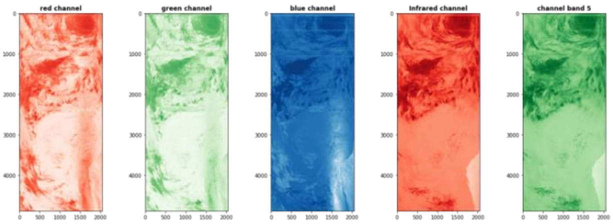


Fig. 8: Visual interpretation of the structure and composition of clouds based on data from each channel.

3.4 Algorithm 3 METAR Dataset Labeling

METAR data as a label is the data source for the presence of CB clouds or NonCB which will be used as data validation. The data frame .csv file does not contain CB and NonCB cloud detections yet. This detection is contained in the 'remarks' section/column. The data in the 'remarks' column is used to distinguish whether or not a CB cloud is detected by detecting the 'CB' code keyword in it or used as a label for CB cloud and NonCB cloud, which can be seen in Fig. 9.

medium_cloud_type	medium_cloud_level	high_cloud_type	high_cloud_level	highest_cloud_type	highest_cloud_level	cloud_coverage	remarks	band_0	band_1	band_2	band_3	band_4
0	0.0	0	0.0	0.0	0.0	4	NOSIG	546.0	482.0	842.0	597.0	576.0
0	0.0	0	0.0	0.0	0.0	2	0	302.0	314.0	812.0	591.0	577.0
0	0.0	0	0.0	0.0	0.0	2	0	319.0	351.0	863.0	618.0	604.0
FEW	1300.0	0	0.0	0.0	0.0	2	NOSIG	710.0	652.0	942.0	826.0	801.0
0	0.0	0	0.0	0.0	0.0	2	0	210.0	256.0	731.0	430.0	429.0
...	...	...	...	...	...	...	...	...	...	...	...	...
0	0.0	0	0.0	0.0	0.0	4	NOSIG	260.0	222.0	919.0	652.0	641.0
0	0.0	0	0.0	0.0	0.0	4	NOSIG	132.0	171.0	847.0	584.0	600.0
0	0.0	0	0.0	0.0	0.0	6	TEMPO FM0730 4000 RA	218.0	216.0	955.0	805.0	791.0
0	0.0	0	0.0	0.0	0.0	6	NOSIG	313.0	287.0	978.0	906.0	886.0
0	0.0	0	0.0	0.0	0.0	6	NOSIG	214.0	211.0	943.0	818.0	805.0

Fig. 9: CB and NonCB cloud labelling.

The CB cloud in the remark column is detected by visualizing the words contained in that column, which is shown in Fig. 9 by creating a WordCloud, a technique for showing which words appear most often in a given text to add a label to the resulting dataset. In Fig. 10, it can be seen that the word 'CB' is the second most common word after 'NOSIG' in the comments column, the number of 'CB' words is 0.185501, while 'NOSIG' is 1,000000. 'CB' is the METAR code for the CB cloud. Apart from that, in the top 5 words read, only CB Cloud is a cloud type.

NOSIG	1.000000
CB	0.184493
RMK	0.079230
TEMPO	0.056593
RA	0.046972
...	...
SCT012	0.000566
FEW015CB	0.000566
FM0640	0.000566
Q1016	0.000566
VC	0.000566

Fig. 10: The visualization result of the words contained in the remark column.

3.5 Algorithm 4 CB and NonCB clouds detection

The process of CB and NonCB cloud detection is explained in the block diagram Fig. 11 NOAA L1B satellite images are used as the data source in the analysis process. First, the images are converted to .tif format to further process them as raster images.

Subsequently, a process of clipping and enhancing the image using scaling methods is carried out. This scaling method utilizes functions in the Geospatial Data Abstraction Library (GDAL). GDAL is responsible for stretching pixel values in raster images to achieve the desired contrast enhancement, thus facilitating data visualization.

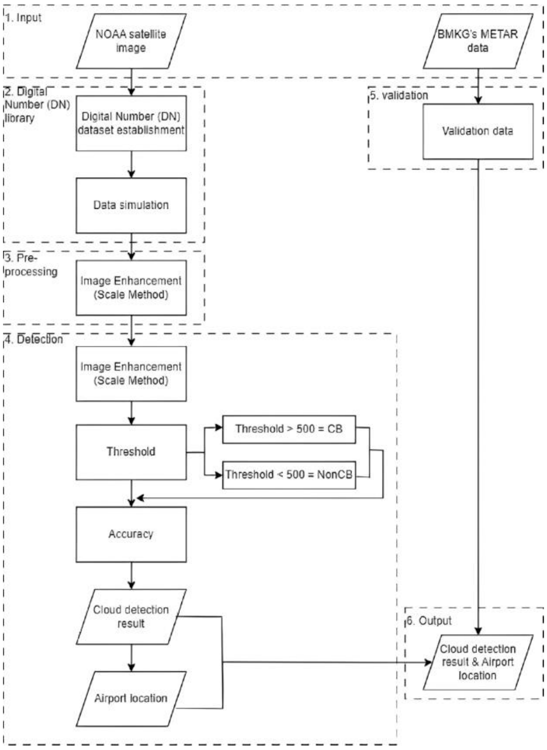


Fig. 11: CB and NonCB cloud detection block diagram.

The process begins by determining the minimum and maximum pixel values in the input image, and setting the desired minimum and maximum pixel value ranges in the output image. Subsequently, GDAL linearly scales the pixel values in the input image to fit the predetermined output range. This process ensures that the final result of the contrast enhancement has optimal pixel values for the specified visualization purposes.

The satellite imagery used as input is NOAA L1B satellite imagery which is converted into .tif format later (raster imagery). Furthermore, the process of image cutting and image enhancement is carried out using the scale method using function on Geospatial Data Abstraction Library (GDAL), which stretches the pixel value of the raster image and serves to enhance the contrast to facilitate visualization. In this case, it is to specify the minimum and maximum pixel values in the input image and the minimum and maximum pixel values in the output image. GDAL then linearly scales the pixel values in the input image to fit a specified output range.

Before the Cumulonimbus (CB) and non-CB cloud detection process, the original image, illustrated in Fig. 12(a), undergoes preprocessing. The initial L1B image is converted into a .tif format (raster image) and is then enhanced for brightness using the scaling method, as depicted in Fig. 12(b).

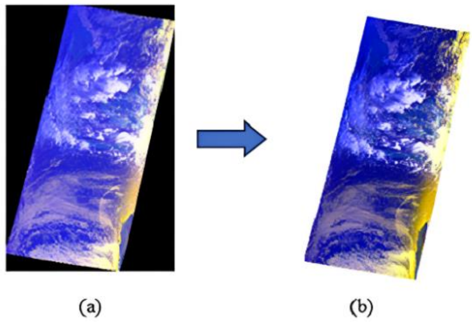


Fig. 12: Illustration of satellite image preprocessing results. (a) Original L1B image (b) image after preprocessing (raster. tif image).

The first step in preprocessing is converting the original L1B image into .tif format (raster image). This conversion allows for a more accurate representation of satellite image data, facilitating further analysis. Next, the brightness of the subject image is enhanced using scaling methods, as depicted in Fig. 13. This process aims to improve the contrast and clarity of the image, making it easier to identify and analyze clouds, including the important detection of Cumulonimbus clouds. Scaling can help highlight relevant features in the image, such as cloud formations, enabling the use of more effective and accurate detection algorithms.

Thus, this preprocessing step is an important initial stage in preparing satellite image data for further analysis, supporting better understanding of atmospheric conditions and more informed decision-making in various weather and climate applications.

The application of thresholding in cloud detection, as outlined in the research, is a key step in the process of classifying pixels or regions in satellite images as Cumulonimbus (CB) clouds or NonCB clouds. This threshold is a value used to separate pixels representing CB clouds from those that do not, based on specific characteristics measurable in the image. In this study, testing was conducted using several threshold values, namely > 800, > 500, and > 300, to evaluate algorithm performance.

The selection of the appropriate threshold value is important because it affects the accuracy of CB cloud detection. In this case, a threshold value of > 500 was chosen based on empirical testing results showing that this value provided the most satisfactory results in distinguishing CB clouds from NonCB clouds. This is adapted to the characteristics of NOAA satellite images from band 1 to band 5 as well as the pixel resolution in visible and infrared channels. By using the appropriate threshold value, the algorithm can provide more accurate and reliable detection results, which are crucial for weather and environmental monitoring applications requiring identification of clouds with high confidence.

The process described regarding the use of threshold values on the average of four bands to visualize the detection results on raster data is a technique to identify and differentiate Cumulonimbus (CB) clouds from non-Cumulonimbus clouds. The threshold value determined here, namely 500, is used as a separator between two conditions:

If the average pixel value of the four bands (band 1 to 4) exceeds 500, then the pixel is considered to represent a Cumulonimbus cloud and is assigned a value of 100. Conversely, if the average pixel value does not reach or equal 500, then the pixel is considered as a non-Cumulonimbus cloud and is assigned a value of 0.

This process is based on the assumption that Cumulonimbus clouds have different characteristics in satellite image data compared to non-Cumulonimbus clouds. By setting a threshold value and applying it to the average pixel value of four bands, we can roughly distinguish between the two types of clouds. The use of uint16 values or 16-bit data to represent pixels in the range of 0 to 65535 is standard in satellite image processing. This range covers all possible pixel values that can be observed in the four bands used, and therefore, is suitable for this purpose.

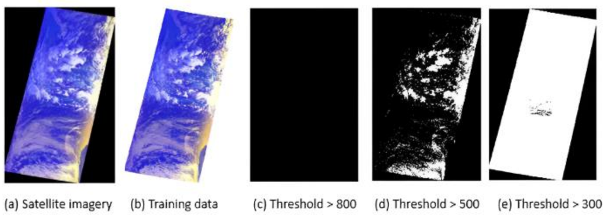


Fig. 13: Data training results.

The process of CB and NonCB cloud detection above can be written using the following mathematical equation:

Input: Tensor $T_{i,j}^{k}$ which means: the size is mxnx4 where 4 is the number of bands:

$$1 \leq i \leq m, \qquad 1 \leq j \leq n, \qquad 1 \leq k \leq 4$$

Output: Raster $R_{i,j}$ for $1 \leq i \leq m,\ 1 \leq j \leq n$.

For a value of i and j:

If $\frac{\sum_{k=1}^{4} T_{i,j}^{k}}{4} > 500$ then $R_{i,j} = 100$ else $R_{i,j} = 0$

3.6 Model Evaluation

The performance of the Geometry Point Square Pixel Algorithm (GPSP) is evaluated by calculating the level of accuracy, which is calculated in percentage (%)[23)24)]:

$$Accuracy\ rate\ (\%) = \frac{n}{N} x100\% \qquad (8)$$

n = The same number of predicted results as METAR data
N = the number of detected stations

Where n is the number of prediction results that match the METAR data, and N is the total number of detected stations. The variable n indicates the number of points where the algorithm's prediction matches the actual observed data, indicating correct predictions. Meanwhile, N represents the total number of points evaluated. By dividing n by N and multiplying by 100, this equation provides the accuracy value in percentage form.

This value indicates how well the GPSP algorithm predicts cloud conditions compared to actual METAR data, where 100% accuracy indicates perfect prediction, while lower values indicate the level of error in the algorithm's predictions. In other words, this equation provides a metric for evaluating algorithm performance in terms of prediction accuracy, with 100% accuracy indicating perfect prediction, and lower values indicating the level of error in the algorithm's predictions.

The final results of the data obtained in this study will be analyzed using the average equation to provide an overview of the performance results from the algorithms used. The average equation applied is as follows:

$$Average = \frac{X1 + X2 + X3 + \ldots + X_n}{n} x100\% \qquad (9)$$

Where $X1 + X2 + X3 + \ldots + X_n$ are the obtained values, and n is the total number of values. By applying this equation, the analysis will provide clearer insights into the performance of the algorithms in detecting clouds.

4. Result and discussion

4.1.Result algorithm CB and NonCB clouds detection

The detection of Cumulonimbus (CB) and non-Cumulonimbus (NonCB) clouds in Indonesia relies on NOAA satellite imagery captured across band 1 to 4, as detailed in Table 1[25)].

Table 1. Band characteristics.

Band	Area Wavelength (μm)	Wavelength	Specifications for Use
Band 1	0,58 – 0,68	Visible	Cloud mapping during the day and surface
Band 2	0,725 – 1,10	Near Infrared (NIR)	Land and water boundaries
Band 3A	1,58 – 1,64	Short Wave Infrared (SWIR)	Ice and snow detection
Band 3B	3,55 – 3,93	Mid-Wave Infrared (MWIR)	Night cloud mapping and sea surface temperature
Band 4	10,30 – 11,30	Thermal Infrared	Night cloud mapping and sea surface temperature

The Table 1 shows various spectral bands used in remote sensing, along with their wavelength ranges, wavelength types, and specifications for use. Band 1 (0.58 – 0.68 µm) is in the visible spectrum and is used for cloud and land surface mapping during daylight. This wavelength is within the visible light spectrum visible to the human eye. The main use is to identify and map cloud and land surface distribution during daytime conditions. Band 2 (0.725 – 1.10 µm) is near-infrared (NIR) useful for detecting the boundary between land and water. Near-infrared wavelengths are very useful in identifying vegetation and observing water boundaries because water absorbs NIR light while plants reflect it. Band 3A (1.58 – 1.64 µm) is shortwave infrared (SWIR) used for detecting ice and snow.

SWIR can penetrate thin layers of clouds and fog, making it suitable for monitoring ice and snow and can be used in land surface sensing. Band 3B (3.55 – 3.93 µm) is mid-wave infrared (MWIR) effective for nighttime cloud mapping and measuring sea surface temperature. MWIR is highly sensitive to heat, making it suitable for nighttime observations and detecting temperature variations on the sea surface. Band 4 (10.30 – 11.30 µm) is thermal infrared, also used for nighttime cloud mapping and sea surface temperature.

This wavelength detects thermal radiation emitted by the Earth and is very useful for monitoring surface temperature, both during the day and at night. Testing was conducted on three NOAA satellite image samples for Cumulonimbus (CB) and non-Cumulonimbus (NonCB) cloud detection using digital geometry point, rectangular polygon, and pixel threshold algorithms, validated using METAR data from BMKG to ensure cloud detection accuracy. This analysis utilizes various wavelengths for specific remote sensing applications, ensuring accurate and useful data for various meteorological and environmental purposes. Test results were compared or validated with BMKG METAR data. The complete detection results of CB and NonCB clouds with airport location points in the Indonesian region for the above 3 test samples are as follows:

Table 1 shows various spectral bands used in remote sensing, along with their wavelength ranges, wavelength types, and specifications for use. Band 1 (0.58 – 0.68 µm) is in the visible spectrum and is used for cloud and land surface mapping during daylight. This wavelength is within the visible light spectrum visible to the human eye. The main use is to identify and map cloud and land surface distribution during daytime conditions. Band 2 (0.725 – 1.10 µm) is near-infrared (NIR) useful for detecting the boundary between land and water. Near-infrared wavelengths are very useful in identifying vegetation and observing water boundaries because water absorbs NIR light while plants reflect it. Band 3A (1.58 – 1.64 µm) is shortwave infrared (SWIR) used for detecting ice and snow.

SWIR can penetrate thin layers of clouds and fog, making it suitable for monitoring ice and snow and can be used in land surface sensing. Band 3B (3.55 – 3.93 µm) is mid-wave infrared (MWIR) effective for nighttime cloud mapping and measuring sea surface temperature. MWIR is highly sensitive to heat, making it suitable for nighttime observations and detecting temperature variations on the sea surface. Band 4 (10.30 – 11.30 µm) is thermal infrared, also used for nighttime cloud mapping and sea surface temperature.

This wavelength detects thermal radiation emitted by the Earth and is very useful for monitoring surface temperature, both during the day and at night. Testing was conducted on three NOAA satellite image samples for Cumulonimbus (CB) and non-Cumulonimbus (NonCB) cloud detection using digital geometry point, rectangular polygon, and pixel threshold algorithms, validated using METAR data from BMKG to ensure cloud detection accuracy. This analysis utilizes various wavelengths for specific remote sensing applications, ensuring accurate and useful data for various meteorological and environmental purposes. Test results were compared or validated with BMKG METAR data. The complete detection results of CB and NonCB clouds with airport location points in the Indonesian region for the above 3 test samples are as follows:

4.2.Sample on 17-01-2021 at 23.30 UTC

On January 17, 2021, at 23:30 UTC, the detection of Cumulonimbus (CB) and non-Cumulonimbus (NonCB) clouds resulted in the identification of 20 airport points. The corresponding satellite imagery is displayed in Fig. 14 (a) and (b), while Fig. 14(c) shows the satellite imagery plotted on the map of Indonesia. Using the geometry point algorithm, the airport points detected within the satellite imagery area were located. Subsequently, the threshold algorithm was applied to determine the cloud type at each detected airport location, as presented in Fig. 14 (d). Each station location is written in ID form, displaying the ICAO airport code consisting of four alphanumeric digits assigned to airports worldwide. This code is regulated by the International Civil Aviation Organization (ICAO), which can also be seen in Table 2. These codes are crucial for ensuring consistent identification and communication across global airports.

The detection of Cumulonimbus and nonCumulonimbus clouds is significant because these cloud Types can affect flight safety. Cumulonimbus clouds, in particular, are known as storm clouds that can cause severe turbulence and extreme weather conditions hazardous to aviation. Therefore, timely and accurate identification of these cloud types around airports is critical for flight planning and aviation safety. The satellite imagery used in this analysis provides visual data that allows for the precise geographic positioning of airports and the type of clouds above them, thus enhancing the reliability and efficiency of this cloud detection system. In the context of international aviation, the consistent and globally recognized (ICAO) airport codes facilitate aviation professionals, including pilots and air traffic controllers, in accurately and quickly identifying and communicating airport locations. This coding system, as shown in Table II, supports efficient and safe coordination in the complex global aviation industry.

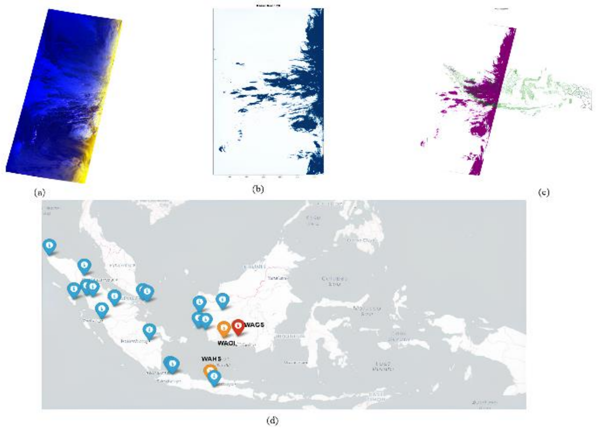


Fig. 14: (a) .tiff raster image, (b) Image of CB and NonCB cloud detection results, (c) Plot results of satellite imagery in Indonesian territory, (d) Visualization display of CB and NonCB cloud predictions at airports in Indonesian territory covered by NOAA satellite imagery.

In Fig. 14 (d), the color on the visualization display for predictions of CB (Cumulonimbus) and NonCB clouds is used to indicate the match or mismatch between prediction results and METAR (Meteorological Aerodrome Report) data. Specifically, the red color indicates that both the predicted results and the METAR data have detected CB clouds. The blue color signifies that NonCB clouds were detected in both the prediction results and the METAR data at the airport point. Meanwhile, the orange color shows a discrepancy, where the prediction results detected CB clouds, but the METAR data detected NonCB clouds.

Table 2 presents the detection results of CB and NonCB clouds at various airport locations, with validation data sourced from BMKG METAR. Based on the detection results, the clouds detected at airports other than WAGI, WAGS, and WAHS are NonCB clouds. In the WAGS area, there is a match between the prediction results and the METAR data, both indicating the presence of CB clouds. However, when compared with BMKG METAR data, there are discrepancies at two airport locations, namely WAGI and WAHS airport. In these locations, the METAR data indicates the absence of CB clouds, differing from the detection results, which identified the presence of CB clouds.

This difference in detection results at the two airport locations highlights the challenges and limitations in accurately predicting CB clouds. Despite these discrepancies, the overall accuracy of the detection results, when compared to the BMKG METAR data, was calculated to be 18 out of 20, or 90%. This suggests that while the detection system is generally reliable, there is room for improvement in enhancing its accuracy, particularly in distinguishing CB clouds.

Table 2. The results of Cloud Detection on 17-01-2021 at 23.30 UTC.

No	Station ID (ICAO)	Airport Name	Predicted CB	METAR BMKG
1	WAGI	Iskandar - Pangkalan Bun	CB	NonCB
2	**WAGS**	**Haji Asan - Sampit**	**CB**	**TCU**
3	WAHQ	Adi Soemarmo - Surakarta	NonCB	NonCB
4	WAHS	Jenderal Ahmad Yani International Airport	CB	NonCB
5	WIBB	Sultan Syarif Kasim II - Pekanbaru	NonCB	NonCB
6	WIDD	Hang Nadim International Airport	NonCB	NonCB
7	WIDN	Raja Haji Fisabilillah - Tanjung Pinang	NonCB	NonCB
8	WIEE	Minangkabau International Airport - Padang	NonCB	NonCB
9	WIHH	Halim Perdana Kusuma - Jakarta	NonCB	NonCB
10	WIII	Soekarno-Hatta International Airport	NonCB	NonCB
11	WIMB	Binaka - Gunung Sitoli	NonCB	NonCB
12	WIME	Aek Godang - Padang Sidempuan	NonCB	NonCB
13	WIMM	Kualanamu International Airport	NonCB	NonCB
14	WIMS	Dr. Ferdinand Lumban Tobing - Pinangsori (Sibolga)	NonCB	NonCB
15	WIOH	Paloh - Sambas	NonCB	NonCB
16	WIOK	Rahadi Oesman - Ketapang	NonCB	NonCB

17	WIOO	Supadio International Airport	NonCB	NonCB
18	WIOS	Tebelian - Sintang	NonCB	NonCB
19	WIPP	Sultan Mahmud Badaruddin II International Airport	NonCB	NonCB
20	WITT	Sultan Iskandar Muda International Airport	NonCB	NonCB

4.3. Sample on 29-01-2021 at 22.52 UTC

The results of CB and NonCB cloud detection on 29-01-2021 at 22.52 UTC resulted in 33 airport points. With the geometry point algorithm, the detected airport points are found in the satellite imagery area. Fig. 15 (a) and (b), while Fig. 15(c) shows the satellite imagery plotted on the map of Indonesia. Then, using the threshold algorithm, the detected airport locations are obtained. Each station location is written in ID form.

In Fig. 15 (d), the color on the visualization display for predictions of CB and NonCB clouds is a blue color indicating that the airport point has detected the prediction results and METAR data for NonCB clouds, the orange color indicates the predicted results have detected CB clouds, but METAR data has detected NonCB clouds, the red color indicates that the predicted results have detected CB clouds, and the METAR data has also detected CB clouds and the green color indicates the predicted results have detected NonCB clouds, but METAR data has detected CB clouds.

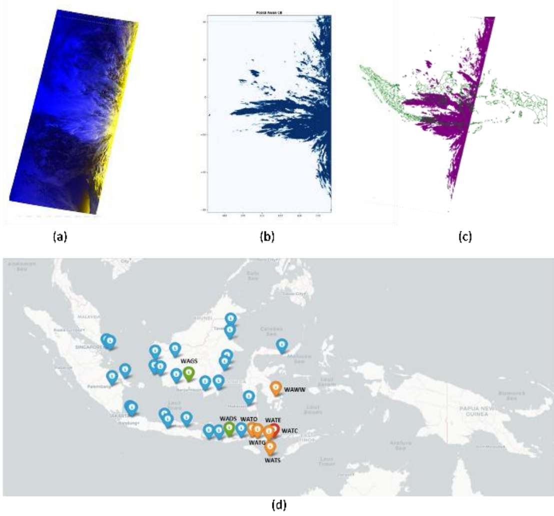


Fig. 15: (a) .tiff raster image, (b) Image of CB and NonCB cloud detection results, (c) Plot results of satellite imagery in Indonesian territory, (d) Visualization display of CB and NonCB cloud predictions at airports in Indonesian territory covered by NOAA satellite imagery.

Table 3. The results of Cloud Detection on 29-01-2021 at 22.52 UTC.

No	Station ID (ICAO)	Airport Name	Predicted CB	METAR BMKG
1	WAAA	Sultan Hasanuddin International Airport	NonCB	NonCB
2	WADB	Sultan Muhammad Salahuddin - Bima	NonCB	NonCB
3	WADD	Ngurah Rai International Airport	NonCB	NonCB
4	WADL	Bandara Internasional Lombok - Praya	NonCB	NonCB
5	WADS	Sultan Muhammad Kaharuddin - Sumbawa Besar	NonCB	CB
6	WAGI	Iskandar - Pangkalan Bun	NonCB	NonCB
7	WAGS	Haji Asan - Sampit	NonCB	CB
8	WAHQ	Adisumarmo International Airport	NonCB	NonCB
9	WAHS	Ahmad Yani - Semarang	NonCB	NonCB
10	WALL	Sultan Aji Muhammad Sulaiman International Airport	NonCB	NonCB
11	WALS	Aji Pangeran Tumenggung Pranoto - Samarinda	NonCB	NonCB
12	WAMG	Djalaluddin - Gorontalo	NonCB	NonCB
13	WAOK	Gusti Sjamsir Alam - Kota Baru	NonCB	NonCB
14	WAOO	Syamsudin Noor International Airport	NonCB	NonCB
15	WAQQ	Juwata International Airport	NonCB	NonCB
16	WAQT	Kalimarau - Tanjung Redeb	NonCB	NonCB
17	WARR	Juanda - Surabaya	NonCB	NonCB
18	**WATC**	**Fransiskus Xaverius Seda - Maumere**	**CB**	**CB**

19	WATE	H. Hasan Aroeboesman - Ende	CB	NonCB
20	WATG	Frans Sales Lega - Ruteng	CB	NonCB
21	WATO	Komodo - Labuan Bajo	CB	NonCB
22	WATS	Tardamu - Sabu	CB	NonCB
23	WAWW	Haluoleo - Kendari	CB	NonCB
24	WIDD	Hang Nadim International Airport	NonCB	NonCB
25	WIDN	Raja Haji Fisabilillah - Tanjung Pinang	NonCB	NonCB
26	WIHH	Halim Perdana Kusuma - Jakarta	NonCB	NonCB
27	WIII	Soekarno-Hatta International Airport	NonCB	NonCB
28	WIKK	Depati Amir - Pangkal Pinang	NonCB	NonCB
29	WIOH	Paloh - Sambas	NonCB	NonCB
30	WIOK	Rahadi Oesman - Ketapang	NonCB	NonCB
31	WIOO	Supadio - Pontianak	NonCB	NonCB
32	WIOS	Tebelian - Sintang	NonCB	NonCB
33	WIPP	Sultan Mahmud Badaruddin II - Palembang	NonCB	NonCB

Table 3 presents the results of Cumulonimbus (CB) and NonCB cloud detection at each airport location, validated using BMKG METAR data. The findings indicate that CB clouds were detected at 6 airport points, specifically WATC, WATE, WATG, WATO, WATS, and WAWW, while 27 points were identified as NonCB clouds. It's worth noting that, based on the BMKG METAR validation data, CB clouds were detected at 6 different airport locations. Consequently, variations were observed in the detection outcomes of Cumulonimbus (CB) and NonCB clouds at nine specific location points. The accuracy of the detection results, when compared to the BMKG metadata, stands at 79% (26 out of 33 points).

4.4.Sample on 02-01-2022 at 12.30 UTC

The results of CB and NonCB cloud detection on 02-01-2022 at 12.30 UTC resulted in 11 airport points. The satellite imagery display is shown in Fig. 16 (a) and (b).

Table 4. The results of Cloud Detection on 02-01-2022 at 12.30 UTC.

No	Station ID (ICAO)	Airport Name	Predicted CB	METAR BMKG
1	WARR	Juanda - Surabaya	NonCB	NonCB
2	WIDD	Hang Nadim International Airport	NonCB	NonCB
3	WIEE	Minangkabau International Airport - Padang	NonCB	NonCB
4	WIII	Soekarno-Hatta International Airport	NonCB	NonCB
5	WIKK	Depati Amir - Pangkal Pinang	NonCB	NonCB
6	WIMM	Kualanamu International Airport	NonCB	NonCB
7	WIMS	Dr. Ferdinand Lumban Tobing - Pinangsori (Sibolga)	NonCB	CB
8	WIOK	Rahadi Oesman - Ketapang	NonCB	NonCB
9	WIOO	Supadio - Pontianak	NonCB	NonCB
10	WIPP	Sultan Mahmud Badaruddin II - Palembang	NonCB	NonCB
11	WITT	Sultan Iskandar Muda International Airport	NonCB	NonCB

The results of the overall detection of CB and NonCB clouds on 3 test samples with validation data from BMKG METAR data can be seen in Table 4. The accuracy of the detection results for CB and NonCB clouds is above 86.6%, so the proposed algorithm is the digital geometry point algorithm, Rectangle Polygon, and pixel threshold produces satisfactory detection of CB and NonCB clouds for NOAA satellite imagery.

The satellite imagery is plotted on the map of Indonesia shown in Fig.16 (c). With the geometry point algorithm, airport points are detected in the satellite imagery area. Then, using the threshold algorithm, the detected airport locations are obtained which is shown in Fig. 16 (d). Each station location is written in ID form.

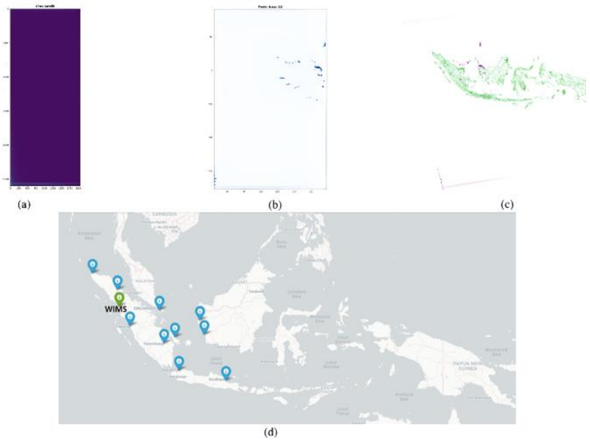


Fig. 16: (a) .tiff raster image, (b) Image of CB and NonCB cloud detection results, (c) Plot results of satellite imagery in Indonesian territory, (d) Visualization display of CB and NonCB cloud predictions at airports in Indonesian territory covered by NOAA satellite imagery.

In Fig. 16 (d), the color representation on the visualization display indicates that blue corresponds to the airport point's prediction results matching with METAR data for non-Cumulonimbus (NonCB) clouds, while green signifies instances where the prediction results align with NonCB clouds, but the METAR data registers Cumulonimbus (CB) clouds. Table IV provides a detailed breakdown of Cumulonimbus (CB) and NonCB cloud detection results at each airport location, all validated against BMKG METAR data. The detection results reveal that all airport locations were identified as having non-Cumulonimbus (NonCB) clouds. However, the BMKG METAR validation data showed that CB clouds were detected at one unique airport location, resulting in a discrepancy between the CB and NonCB cloud detection results at that specific point. The accuracy of these detection results when compared to BMKG data is 10/11 = 91%.

The presented Table 5, is the result of observations from the NOAA-19 satellite using the GPSP algorithm to detect cloud bands on certain dates. The "Samples" column shows the specific date and time the observation was made, such as on January 17, 2021, at 23:30, January 29, 2021, at 22:52, and January 2, 2022, at 12:30. The total number of detected cloud bands for each observation is displayed in the "Number of airports detected" column, with varying numbers such as 20, 33, and 11 for each respective date.

The next column explains the use of the GPSP algorithm in detecting cloud bands, with the detection results further categorized in the "METAR Accuracy" column. In this column, the detected cloud bands are divided into two categories: CB (Cumulonimbus) and NonCB (non-Cumulonimbus).

The last column shows the detection accuracy percentage, indicating how accurate the classification of the cloud bands was, with accuracy values of 90% for the observation on January 17, 2021, 79% on January 29, 2021, and 91% on January 2, 2022. This data provides an overview of the distribution of clouds and weather conditions observed by the NOAA-19 satellite, as well as the effectiveness of the GPSP algorithm in classifying these types of clouds.

Table 5. The Percentage of Successful Detection of CB and NonCB Clouds.

Samples	The number of airports detected	GPSP Algorithm		METAR		Accuracy	Average Accuracy
		CB	NonCB	CB	NonCB		
17-01-2021 2330 NOAA19_L1B	20	3	17	1	19	90%	86,60%
29-01-2021 2252 NOAA19_L1B	33	6	27	3	30	79%	
02-01-2022 1230 NOAA19_L1B	11	0	11	1	10	91%	

4.5.Discussion

This study relies on NOAA satellite imagery, known for its wide spatial coverage, high spectral resolution, and frequent observations. The AVHRR sensors aboard NOAA satellites are highly valuable to researchers across diverse fields, such as meteorology, environmental monitoring, and climate change. Numerous studies related to NOAA satellite imagery have been previously conducted[1)2)3)4)5)].

The NOAA satellite has 5 channels, and this study uses channel one (visible) to channel four (thermal infrared) to detect CB and NonCB clouds. A similar study was conducted by Kohei Arai[7)].

Detecting Cumulonimbus clouds involves utilizing both visible and thermal infrared channels. In a similar vein, Gautam[26)] harnessed channels one through four to identify hotspots, employing the Principal Component Analysis (PCA) method. The proposed method, known for its adaptability, has demonstrated significant success in analyzing most AVHRR images, boasting an average detection accuracy of 87.27% and a minimal false alarm rate of 0.201%. This accuracy was assessed by comparing results with ground truth data in the Jharia region of India. Cloud properties exhibit substantial spatial and temporal variation. Describing the visual appearance of clouds encompasses six key features: brightness, texture, size, shape, organization, and shadow effects. While spectral properties of clouds may change, their textural properties remain distinctive for specific cloud types[27)28)].

CB and NonCB cloud detection in this study takes the brightness features of the clouds. To find out the success rate of the algorithm in performing cloud detection, that is comparing it with METAR data. Similar research was also conducted by Tuomola (2021) and Rumi et al[11)]. Before carrying out the cloud detection process, it is necessary to create airport data in .csv format and airport boundary datasets in. gpkg format as well as polygon geometry data (template). Next, carry out the CB and NonCB cloud detection process using the threshold algorithm as follows.

a. Take image data with the attributes date, month, year, hour and minute to be tested with the L1B extension
b. Convert the L1B image to .tif format using the GDAL library to produce 2 .tif output (the section 1.1)
c. Retrieve METAR data from the L1B file (data to be tested) by accessing the METAR web
d. Opening the airport boundary file in. gpkg format (the section 1.1)
e. Iterate line by line from the METAR .txt file to find out the airport code (example WAAA) (the section 1.1)
f. The results of the line-by-line iteration are airport codes which are then compared with the data in the. gpkg airport boundaries (the section 1.1)
g. The box is attached to the raster.tif image
h. The raster.tif image is cut into squares (grid square)
i. Take the highest pixel of the grid (airport) for all five satellite raster bands (the section 1.2)
j. Calculate the average pixel value of 4 bands (Band 1 + Band 2 + Band 3 + Ban4) > 500, if it is above 500 it is initialized with a value of 100 otherwise it is 0. So the value 100 is CB, otherwise it is NonCB (the section 1.4)
k. From the average pixel calculation, the CB or NonCB category is validated with the cloud type originating from the Huge.csv file in the remark column (the section 1.2). If the remark column contains the word CB then the cloud type is CB (the section 1.3)

The success rate of cloud detection in this study was 86.6%, focusing on detecting CB and NonCB cloud areas around airports throughout Indonesia. Accuracy rates of 90%, 79%, and 91% were obtained using equation (8), while the overall accuracy of 86.6% was calculated by averaging the percentages from the three test samples using equation (9). The calculations were conducted by first calculating the average for each satellite image individually. Then, the average results from each image were further averaged to obtain a final value that is more representative. This approach ensures that variations between the images are comprehensively accounted for before deriving the overall average.

In this study, a box template size of 0.08° x 0.08° was used with a threshold algorithm shown in Fig. 3. A similar study was also conducted by Osaka et al (2011) with a threshold algorithm and using a grid template of 0.01° × 0.01° for cumulus cloud detection with MTSAT satellite imagery. The results obtained cannot be determined with the right threshold value to obtain optimal results. Yasuhiko[20)], identified cumulus, Cumulonimbus, and cirrus areas by developing the Convective cloud information (CCI) method using a box template (grid) measuring 2t+1 and 2s+1 with the center point x_o, y_o with Himawari image data for flight safety. From the obtained result, the data frame is used to classify CB and NonCB cloud detection.

The method used in this labeling is the bilinear interpolation method. The linear interpolation method is used to obtain a new value containing white pixels in a satellite image to identify the type of cloud (Cloud type) CB and NonCB clouds as labels for CB clouds and NonCB clouds, which will be used to predict CB clouds by taking pixels in band 0-4 for every station. METAR data labeling results can be seen in Table 6.

Table 6. METAR dataset labeling results.

Station_Id	**Latitude**	**Longitude**	**Remarks**	**Cloud_Type**
WAMG	0.52	123.07	CB TO NW	CB
WAGI	-2.7	111.6666667	NOSIG RMK CB TO W & NW	CB
WALS	-0.616666667	117.15	CB TO N	CB
WIKK	-2.17	106.13	NOSIG RMK CB TO NW	CB
WIOK	-1.85	109.97	CB TO SE	CB
WAHS	-6.98	110.38	CB TO S	CB
WALL	-1.27	116.9	CB TO W	CB
WIKK	-2.17	106.13	NOSIG RMK CB TO S	CB
WAAA	-5.07	119.53	TEMPO TL0850 4000 TSRA RMK CB IN APCH	CB
WATO	-8.483333333	119.8833333	CB IN CLIMB OUT	CB
WIDN	0.92	104.52	NOSIG RMK CB TO SE	CB
WIMB	1.17	97.7	CB TO SW	CB

The research results show that the method used can detect cloud areas ranging from growing cumulus to mature Cumulonimbus by 69%. And in research conducted by Jing et al 2016, namely detecting estimated rainfall, using a spatial resolution with a box template of 0.25° × 0.25° or a box size of 25 km longitude/latitude.

From the results of this research, the number of airports covered for each test sample was on average more than 11 airports, namely 11, 20, 33 and 35 airports. The number of airports covered depends on the satellite image coverage area in Indonesia based on maps of Indonesian cities/regencies. If there are a lot of satellite images covering Indonesia's territory, then the number of airports detected will also be large. Meanwhile, cloud detection depends on the Digital Number (DN) value.

The results of CB and NonCB cloud detection in this study were carried out using the METAR dataset labeling method by reading the results of the remark column. The dataset used for labeling uses data from 2019 to 2022. And the CB cloud detection results in the table for each test sample are in accordance with the BMKG METAR data. Table VI shows the results of labeling CB and NonCB clouds which consist of station ID, latitude, longitude, and remarks information, resulting in labeling of words that often appear, namely in the form of CB clouds.

The dataset used represents real-time data based on actual weather conditions, so the limited number of CB clouds is a natural occurrence. The proportion of CB clouds in the METAR data is indeed small, which inevitably affects the evaluation of CB cloud detection. However, it is important to note that this algorithm is not only designed to detect CB clouds but also to identify non-CB clouds. The detection capability that covers both types of clouds contributes to the accuracy of each satellite image, which is then averaged to obtain a more representative final result.

5. Conclusions

Given based on these further evaluation metrics, it can be concluded that the accuracy of detecting CB and NonCB clouds is 86.6%. Therefore, the proposed algorithm, which uses digital geometry points, rectangle polygons, and pixel thresholds, capable of effectively detects CB and NonCB clouds in NOAA satellite imagery.

By implementing the Geometry Point Square Pixel (GPSP) algorithm, we can generate more accurate and reliable detections of CB and NonCB clouds in satellite images. This improvement is supported by several factors:

a. High Accuracy: The GPSP algorithm has proven to provide detection results with high accuracy, with an average cloud detection accuracy exceeding 86.6%. This ensures that information obtained from NOAA satellite images can be relied upon for various purposes, such as weather monitoring, weather prediction, and aviation safety.
b. High Resolution: The GPSP algorithm can capture fine details in satellite images, enabling detection of clouds with high resolution. This allows for more accurate observation of cloud patterns, including detection of complex clouds and rapid weather changes.
c. Validation with METAR Data: Integration with METAR data provides strong validation of detection results. By comparing detection results with METAR data, we can ensure that the presented information is accurate and consistent with actual weather conditions in the field.
d. More Efficient Weather Monitoring: By using the GPSP algorithm, weather monitoring can be conducted more efficiently and effectively. The rapid and accurate cloud detection process enables authorities to respond to weather changes more timely, which can enhance safety and efficiency in various sectors, including aviation, agriculture, and fisheries.

The expansion of using these algorithms provides significant benefits in cloud monitoring from satellite imagery data. The digital geometry point algorithm is used to identify significant points in the image representing possible cloud locations. Meanwhile, the Rectangular Polygon allows clustering of these points to form more precise boundaries, facilitating further identification and analysis. The Pixel Threshold is then applied to distinguish between CB and NonCB cloud types based on observed radiation characteristics in the image. The success of these algorithms in detecting CB and NonCB clouds with high accuracy demonstrates remarkable capabilities in processing satellite image data efficiently and effectively. These satisfactory detection results have important implications in various applications, including weather monitoring, weather prediction, and aviation safety. Thus, the use of these algorithms has proven to make valuable contributions in enhancing our understanding of atmospheric and environmental conditions, as well as improving the reliability of satellite-based weather monitoring systems.

Acknowledgment

The author expresses his sincere thanks to the National Research and Innovation Agency (BRIN) which has provided support in terms of the availability of research data.

References

1) J. Wallach, “User’s Guide for Building and Operating Environmental Satellite Receiving Stations,” US Department of Commerce, National Oceanic and Atmospheric Administration, (1997).
2) J. Alvarez-Ramirez, C. Ibarra-Valdez, E. Rodriguez, “Fractal analysis of Jackson pollock’s painting evolution, Chaos, Solitons and Fractals” 83 pp. 97–104 (2016). doi:10.1016/j.chaos.2015.11.034
3) J. P. Cangialosi, A. S. Latto, R. Berg, “National hurricane center tropical cyclone report: Hurricane irma,” National Hurricane Center pp. 1–111 (2018). doi: 10.1175/WAF-D-19-0007.1
4) J. P. Cangialosi, J. L. Franklin, “National hurricane center forecast verification report,” National Hurricane Center (2010).
5) I. Dutta, S. Banerjee, M. De, “An algorithm for pre-

processing of satellite images of cyclone clouds," International Journal of Computer Applications 78 (15), (2013). doi:10.1520/13598-1317
6) Q. Li, W. Lu, J. Yang, "A hybrid thresholding algorithm for cloud detection on ground-based color images," Journal of atmospheric and oceanic technology 28 (10) pp. 1286–1296 (2011). doi: 10.1175/JTECH-D-11-00009.1
7) K. Arai, "Thresholding based method for rainy cloud detection with noaa/avhrr data by means of jacobi itteration method," Int. J. Adv. Res. Artif. Intell 5 (6) pp. 21–27 (2016). doi:10.14569/IJARAI.2016.050605
8) R. R. Ghosh, M. S. Ali, A. Hena, H. Rahman, "A simple cloud detection algorithm using noaa-avhrr satellite data," International Journal of Scientific and Engineering (IJSER) Research 3 (6) PP. 1-5 (2012).
9) W. Zhuo, Z. Cao, Y. Xiao, "Cloud classification of ground-based images using texture–structure features," Journal of Atmospheric and Oceanic Technology 31 (1) pp. 79–92 (2014). doi:10.1175/JTECH-D-13-00048.1
10) S. Mahajan, B. Fataniya, "Cloud detection methodologies: Variants and development—a review," Complex & Intelligent Systems 6 pp. 251–261 (2020). doi:10.1007/s40747-019-00128-0
11) E. Rumi, D. Kerr, A. Sandford, J. Coupland, M. Brettle, "Field trial of an automated ground-based infrared cloud classification system," Meteorological Applications 22 (4) pp. 779–788 (2015). doi:10.1002/met.1523
12) O. S. Sitompul, M. Zarlis, S. Suwilo, et al., "Automatic detection of chaos phenomenon in tornadoes prediction using edge detection," in: 2019 3rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM), IEEE, pp. 48–52 (2019). doi:10.1109/elticom47379.2019.8943845
13) W. Wanayumini, "Supervised Image Classification of Chaos Phenomenon in Cumulonimbus Cloud Using Spectral Angle Mapper," International Journal on Advanced Science Engineering Information Technology, pp. 987-992 (2020). doi:10.18517/ijaseit.10.3.11493
14) R. Dorfman, E. Wagner, A. Lahav, A. Amar, R.Talmon, Y. Halle, "Spatiotemporal detection of Cumulonimbus clouds in infrared satellite images," in: 2018 International Conference on the Science of Electrical Engineering in Israel (ICSEE), IEEE, pp. 1–5 (2018). doi:10.1109/ICSEE.2018.8646047
15) Q. Wang, W. Fan, "Cumulonimbus forecasting based on rough set and artificial immune algorithm," in: 2010 Sixth International Conference on Natural Computation, Vol. 6, IEEE, pp. 2856–2860 (2010). doi:10.1109/ICNC.2010.5584020
16) A. F. Rais, R. M. Putra, M. A. Fitrianto, T. Hermansyah, "A preliminary comparative study on the feasibility of a multipurpose numerical weather model for prediction of Cumulonimbus clouds in indonesia," in: 2022 International Conference on Science and Technology (ICOSTECH), IEEE, pp. 1–6(2022). doi:10.1109/ICOSTECH54296.2022.9829049
17) T. V. La, C. Messager, R. Sahl, P. Dupont, E. Prothon, M. Honnorat, "Combination of geostationary and polar satellite sensors to monitor Cumulonimbus and their winds at the ocean surface," in: IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, IEEE, pp. 5380–5383 (2020). doi:10.1109/IGARSS39084.2020.9323111
18) A. F. Rais, F. Setiawan, R. Yunita, E. Meinovelia, S. Soenardi, M. Fadli, B. Wijayanto, et al., "Prediction of Cumulonimbus (cb) cloud based on integrated forecast system (ifs) of european medium-range weather forecast (ecmwf) in the flight information region (fir) of jakarta and ujung pandang," Jurnal Sains & Teknologi Modifikasi Cuaca 21 (2) pp. 95-100 (2020). doi:10.29122/jstmc.v21i2.4100
19) R. Putra, E. Fibriantika, H. Herawati, Y. Kusumayanti, E. Afriani, A. Hidayanti, A. Wiujiana, W. Swastiko, D. Andrianto, "Cumulonimbus cloud prediction based on machine learning approach using radiosonde data in surabaya, Indonesia," in: IOP Conference Series: Earth and Environmental Science," Vol. 724, IOP Publishing, pp. 012047 (2021). doi:10.1088/1755-1315/724/1/012047
20) S. Yasuhiko, S. Hiroshi, I. Takahito, S. Akira, "Convective cloud information derived from himawari-8 data," Meteorological Satellite Center Technical Note No62, pp. 19-37 March (2017).
21) L. Sun, X. Mi, J. Wei, J. Wang, X. Tian, H. Yu, P. Gan, "A cloud detection algorithm-generating method for remote sensing data at visible to short-wave infrared wavelengths," ISPRS journal of photogrammetry and remote sensing 124 pp. 70–88 (2017). doi:10.1016/j.isprsjprs.2016.12.005
22) R. Klette, A. Rosenfeld, "Digital geometry: Geometric methods for digital picture analysis," Morgan Kaufmann, (2004). doi:10.1016/j.isprsjprs.2016.12.005
23) S. Soim, S. Sholihin, C.B. Subianto, "Optimizing Performance Random Forest Algorithm Using Correlation-Based Feature Selection (CFS) Method to Improve Distributed Denial of Service (DDoS) Attack Detection Accuracy," Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM), vol. 7, no. 2, pp. 220–228, Sep. 2024. doi:10.24014/ijaidm.v7i2.24783
24) Y. Liu, J. Xia, C.-X. Shi, Y. Hong, "An improved cloud classification algorithm for china's fy-2c multi-channel images using artificial neural network, Sensors 9" (7) pp. 5558–5579 (2009). doi:10.3390/s90705558
25) W.-H. Lee, J.-i. Kudoh, S. Makino, "A new cloud

detection algorithm for noaa avhrr imagery, in: Proceedings of IEEE. IEEE Region 10 Conference. TENCON 99." 'Multimedia Technology for Asia-Pacific Information Infrastructure' (Cat.No. 99CH37030), Vol. 1, IEEE, pp. 170–173 (1999). doi:10.1109/TENCON.1999.818377
26) R. S. Gautam, D. Singh, A. Mittal, "Application of principal component analysis and information fusion technique to detect hotspots in noaa/avhrr images of jharia coalfield, india," Journal of Applied Remote Sensing 1 (1) 013523 (2007). doi:10.1117/1.2771256
27) N. Lamei, K. D. Hutchison, M. M. Crawford, N. Khazenie, "Cloud-type discrimination via multispectral textural analysis," Optical engineering 33 (4) pp. 1303–1313 (1994). doi:10.1117/12.166920
28) G. Pankiewicz, "Pattern recognition techniques for the identification of cloud and cloud systems," Meteorological Applications 2 (3) pp. 257–271 (1995).
29) L. Tuomola, "Cumulonimbus cloud detection with weather radar at helsinki-vantaa airport". doi:10.1002/met.5060020309
30) I. Osaka, T. Imai, Y. Izumikawa, T. Matsumoto, "Detection of rapidly developing cumulus areas through mtsat rapid scan operation," Meteorological Satellite Center Technical Note, pp. 699-91 (2011).
31) W. Jing, Y. Yang, X. Yue, X. Zhao, "A comparison of different regression algorithms for downscaling monthly satellite-based precipitation over north china," Remote Sensing 8 (10) 835 (2016). doi:10.3390/rs8100835
32) L. Zili, Y. Jiajung, W. Wenjing, S. Zhenwei, Cloud detection methods for remote sensing images: a survey, Chinese Space Science and Technology 43(1), 1, (2023). doi:10.16708/j.cnki.1000-758X.2023.0001
33) Dascal, V., Dolea, P., Cristea, O., Palade, T., "Thresholding based method for rainy cloud detection with noaa/avhrr data by means of jacobi itteration method," Acta Technica Napocensis 53 (2012). doi:10.14569/IJARAI.2016.050605
34) Line, W.E., Schmit, T.J., Lindsey, D.T., Goodman, S.J, "Use of geostationary super rapid scan satellite imagery by the storm prediction center," Weather and Forecasting 31(2), pp. 483–494. (2016). doi:10.1175/WAF-D-15-0135.1
35) Tang, R., Tang, B., Wu, H., Li, Z.-L. "On the feasibility of temporally upscaling instantaneous evapotranspiration using weather forecast information." International Journal of Remote Sensing 36 pp. 1–18 (2015). doi:10.1080/01431161.2015.1029597
36) ICAO, "Safety Report 2018," International Civil Aviation Organization, Montreal, Canada pp. 1-51 (2018).