An Assessment of Convective Initiation Nowcasting Algorithm within 0-60 Minutes using Himawari-8 Satellite

I F P Perdana¹, and D Septiadi¹

¹Department of Meteorology, State College of Meteorology Climatology and Geophysics, Perhubungan I Street, South Tangerang, Banten, Indonesia, 15521

ilham.fajar@stmkg.ac.id

Abstract. Convective cloud monitoring since its growth stage primarily related to location and time of the first convective cloud initiated, called convective initiation (CI), could be the primary key in providing an earlier heavy rainfall event prediction. This study aimed to assess the accuracy and lead time of CI nowcasting using Satellite Convection Analysis and Tracking (SATCAST) algorithm in predicting the CI event within 0-60 minutes over Surabaya and surrounding area using Himawari-8 satellite during June-July-August (JJA) period in 2018. Three main processes used in this study were cloud masking, cloud object tracking, and CI nowcasting. Twelve interest fields were utilized as predictors based on six bands of Himawari-8 satellite, which represented cloud physics attributes such as cloud-top height, glaciation, or cooling rate. The verification was conducted by comparing CI prediction to CI location and time based on Surabaya weather radar within the next 0-60 minutes. The algorithm resulted that the prediction could achieve 87.3% of accuracy from the 3449 cloud objects. The prediction had POD, FAR, and CSI scores of 57.1%, 52.2%, and 35.2%, respectively. The 32.3 minutes of averaged lead time prediction indicated that CI nowcasting could detect growing cumulus about 30 minutes prior to the CI event.

Keywords: convective initiation, SATCAST, Himawari-8, nowcasting

1. Introduction
Hydrometeorological hazards such as strong wind, heavy rain, and lightning are often associated with convective clouds. Meanwhile, hydrometeorological hazards can impact socio-economic losses or even threaten human safety [1]. Therefore, it is necessary to predict the heavy rain event, especially those caused by convective clouds, to suppress the impact that probably can be happened. The accurate prediction regarding the location and time of the first convective event, called convective initiation (CI), could be the primary key in providing better heavy rain events prediction due to convective clouds [2].

Significant effort has been conducted to achieved better CI development and process understanding, especially its physical processes. [2,3]. One of the challenging parts of the CI process is the short time scales of convective cloud evolution which ranges from minutes to few hours. Therefore, remote sensing utilization with high spatial and temporal resolution could help the convective cloud evolution monitoring. Weather radar becomes widely used in predicting the location of CI occurrence and the propagation of the convective system due to the changes in weather radar...
reflectivity that can indicate a change of humidity in the vertical column of the atmosphere [4]. Some weather radar-based algorithms for predicting the initiation and propagation of convective cloud had developed, such as Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN) [5], the Warning Decision Support System-Integrated Information [6], and Radar Tracking and Monitoring (RadTRAM) [7]. However, most significant weather radar echoes were detected when the convective systems started to precipitate or reach the mature stage [7]. These issues probably happened because the standard operational weather radar network configuration focused on catching a considerable hydrometeor particle size. As a result, the nowcasting will get a shorter lead time prediction of convective cloud development.

The geostationary satellite, which has comparable temporal resolution with weather radar, is commonly used to monitor cloud location, distribution, and development. The geostationary satellite could be used to detect CI by identifying cloud-top brightness temperature with rapid cooling rate up to more than 30 minutes before weather radar displayed a significant echo [8]. Therefore, the geostationary satellite can be used to monitor convective cloud development since its growth phase and provide an earlier prediction than weather radar. Some geostationary satellite-based algorithms have been developed for CI nowcasting, such as Satellite Convection Analysis and Tracking (SATCAST) [9] and University of Wisconsin Convective Initiation (UWCI) [10] using GOES-12 satellite, and Rapid Developing Cumulus Area (RDCA) [11] using MTSAT-2 satellite. Those developed algorithms detected CI using single criterion or multiple criteria of brightness temperature, which represented the cloud physics attributes such as single spectral, differences between two spectral, or temporal variation of the spectral. The validation of algorithms resulted in a great potential in CI nowcasting using a geostationary satellite with lead time prediction of about 20-40 minutes.

SATCAST continued to be developed using GOES-13 and GOES-R Advanced Baseline Imager (ABI) [12,13]. The research enhanced the cloud tracking method by treating the clouds as a single object, which means the algorithm no longer needed satellite-derived mesoscale atmospheric motion vectors (MAMV) because the cloud object tracking would use only the two most recent datasets. Besides that, the algorithm used infrared channels only, so the CI could be detected during the whole day and not limited during daytime only as the previous SATCAST did. The validation reached the probability of detection (POD) of 85% and false alarm rate score (FAR) of 55%, with an average lead time of 27 minutes.

Himawari-8, with the primary payload of Advanced Himawari Imager (AHI), captures almost half of the whole Earth with 16 spectral channels every 10 minutes [14]. Himawari-8 AHI, which is launched and operated by Japan Meteorological Agency (JMA), has the highest spatial resolution for visible wavelength (band #3) at 0.5 km followed by 2 km of resolution for all near-infrared and infrared band except band #4 which has 1 km resolution. The Himawari-8 AHI spectral characteristics are comparable to the GOES-R ABI satellite operated by the National Aeronautics and Space Administration (NASA) and National Oceanic and Atmospheric Administration (NOAA) [15]. Therefore, the criteria used in CI nowcasting using GOES-R satellite can be adopted and applied in the Himawari-8 satellite. Some researchers had conducted an adoption of SATCAST algorithm and concept using Himawari-8 satellite to predict CI event in Korea and South-eastern China [16,17,18]. The validation resulted in better scores which achieved more than 75% in POD and less than 40% in FAR due to some improvements such as implementing machine learning or adjusting the threshold based on statistics of local cloud spectral. The average lead time also increased at 30-60 minutes.

Indonesia, as part of the Maritime Continent (MC), has a larger potential for convective development due to ocean-continent thermal contrast or complex terrain of local features [19]. It is necessary to provide a nowcasting system to reduce the potential losses. However, the SATCAST algorithm has not been assessed in the tropics, especially in Indonesia. The main goal of this study is to evaluate the accuracy and lead time of the CI nowcasting using the SATCAST algorithm in predicting the CI event within the 0-60 minutes timeframe. This paper is ordered as follows: Section 2 describes the area of interest and data used in this paper, Section 3 provides the information about CI
prediction and verification, Section 4 discusses the CI verification results, and Section 5 concludes the overall results of this paper.

2. Area of interest and data
This study took place over Surabaya, East Java, Indonesia during June to August (JJA) period in 2018 (Figure 1). Surabaya was chosen because there was a downtrend of annual precipitation amount, but it showed an uptrend of heavy rain event frequency over the last 30 years [20]. Himawari-8 data satellite from six bands (6.2, 7.3, 8.6, 11.2, 12.3, and 13.3 µm in central wavelength) with 10 minutes and 0.02° × 0.02° (2 km × 2 km) of temporal and spatial resolution, respectively, were used to examine the spectral characteristics of growing cloud. Those interest fields would determine whether the targeted cloud was predicted as CI or not. Meanwhile, Surabaya C-Band weather radar data with 10 minutes and 250 m of temporal and spatial resolution, respectively, were used to find the location and time of the CI event. The radar data would verify the CI prediction due to the definition of CI as the first occurrence of radar reflectivity greater than or equal to 35 dBZ [8,9,12,13]. Radar data was cropped to 120 km in range from the radar position to avoid the beam spreading of radar on low elevation scans at some distances [21]. In this paper, Column Maximum (CMAX) product was used and set to 1 and 15 km for the bottom and top height, respectively.

3. Methods used in CI determination and verification

3.1. CI determination
There were three main processes in CI determination in this research, (1) cloud masking, (2) object tracking, and (3) CI determination based on interest fields. Cloud masking in this process was aimed to select potential thick cloud or immature cumulus and removed cloudless area, cirrus, and mature cumulus. The original SATCAST used ancillary cloud type data in this procedure. However, this research applied cloud masking procedure based on the combination of previous research to make the algorithm became independent from ancillary data [17,18] with equations as follow:

\[
253.15 \, K < T_b(11.2) \, \mu m < 288.5 \, K
\]

\[
T_b(11.2) \, \mu m - T_b(12.3) \, \mu m < 6 \, K
\]

\[
T_b(6.2) \, \mu m - T_b(11.2) \, \mu m > -38 \, K
\]
Equation (1) was used to remove the cloudless area and matured clouds, and Equation (2) and (3) were used to remove cirrus clouds. This cloud pixel-based assessment would remain the pixels that were considered as a thick or growing cloud. Once the potential pixels were obtained, all those nearest pixels were grouped into a single entity called cloud object using the four-neighbourhoods connected-component labelling technique [22]. Besides, the cloud object which consists of less than four pixels were excluded to remove noises [16]. Figure 2 shows a comparison of the visible and infrared channel of the Himawari-8 satellite to the cloud masking result.

![Comparison of visible and infrared channel](image)

**Figure 2.** Comparison of (a) visible channel, (b) infrared channel, and (c) cloud masking product over Surabaya and around on January 14, 2018 at 12.00 local time (LT = UTC + 7 hours).

| Interest Fields | Cloud Attributes | Threshold |
|-----------------|------------------|-----------|
| $T_b$ 11.2 µm   | Cloud-top height and glaciation | -20º C to 5º C |
| $T_b$ 6.2 – 11.2 µm | Cloud-top height | -30º C to -10º C |
| $T_b$ 6.2 – 7.3 µm |                      | -25º C to -5º C |
| $T_b$ 13.3 – 11.2 µm |                      | -20º C to -5º C |
| $T_b$ 12.3 – 11.2 µm | Cloud glaciation | -3º C to 0º C |
| $T_b$ 8.6 – 11.2 µm |                      | -10º C to -1º C |
| $T_b$ (8.6 – 11.2 µm) – (11.2 – 12.3 µm) | | -10º C to 0º C |
| Temporal changes of $T_b$ 11.2 µm | Cloud-top cooling rate | < -2.66º C / 10 minutes |
| Temporal changes of $T_b$ 6.2 – 11.2 µm | Temporal changes on cloud-top height | > 1º C / 10 minutes |
| Temporal changes of $T_b$ 6.2 – 7.3 µm |                      | > 0º C / 10 minutes |
| Temporal changes of $T_b$ 12.3 – 11.2 µm | Temporal changes on cloud glaciation | > 1º C / 10 minutes |
| Temporal changes of $T_b$ (8.6 – 11.2 µm) – (11.2 – 12.3 µm) | | > 0º C / 10 minutes |

Once all individual cloud objects were identified, object tracking was performed to identify a particular cloud position and distribution over the period every 10 minutes. Object tracking was conducted by taking two consecutive cloud object data at $t_i$ and $t_{i+10 \text{ minutes}}$. Two cloud objects would be assigned the same unique number identifier (ID) if there was an overlapping area between two cloud objects at $t_i$ and $t_{i+10 \text{ minutes}}$. The high temporal resolution of the Himawari-8 satellite made MAMV was not essential in this algorithm [17]. Meanwhile, the new objects (new ID) was
defined if the object was identified at $t_{i+10\ minutes}$ but it did not exist at $t_i$. After all cloud objects were assigned with a unique ID, twelve interest fields based on SATCAST interest fields were evaluated by averaging the spectral value of 25% coldest pixels of $T_b\ 11.2\ \mu m$ within the cloud object. This calculation was intended to prevent the blurring signals and helped to focus on the updraft potential [12,13,17]. The CI prediction was determined based on previous interest fields and threshold (Table 1). The interest fields acted as cloud physics representation which gave information about cloud-top height, glaciation, or cooling rate. Cloud object was considered as positive CI (CI probably happen within the next 0-60 minutes) if at least 7 of 12 criteria were met and considered as negative CI (CI does not probably happen within the next 0-60 minutes) if only six or fewer criteria were met.

3.2. Verification

CI nowcasting was validated using Surabaya weather radar data which CI was defined as the first occurrence of reflectivity data greater than or equal to 35 dBZ [8,9,12,13]. Before the verification performed, the CI location and time were determined during the research period. The CI satellite-based prediction (positive and negative CI) were compared to CI location and time based on the radar within the next 0-60 minutes. Since the prediction result was dichotomous, a $2\times2$ contingency table was performed in this paper. Hit ($H$) was the number of correctly predicted CI event. Miss ($M$) was the number of negatives CI prediction followed by CI event observed within the next 10-60 minutes. However, the number of positives CI prediction with zero lead time is considered as Miss too. Meanwhile, false alarm ($F$) was the number of positives CI prediction without being followed by CI event within the next 10-60 minutes. Last, correct negative ($C$) was the number of correctly predicted non-CI event. Four quantitative skill scores, i.e. accuracy (ACC), probability of detection (POD), false alarm rate (FAR), and critical success index (CSI), were calculated to assess the CI nowcasting performance of CI prediction.

4. Results and Discussions

4.1. CI result on dichotomous verification and its distribution

Based on cloud objects and CI events detection using Himawari-8 satellite and Surabaya weather radar, respectively, there were 3449 cloud objects and 415 CI events detected during JJA period in 2018 over Surabaya and around. CI nowcasting verification was conducted by evaluating the CI satellite-based prediction within the next 0-60 minutes compared to the CI event. The verification result is presented in Table 2.

| Predicted | Observed | Yes | No |
|-----------|----------|-----|----|
| Yes       | 237      | 259 |
| No        | 178      | 2775|

Table 2 shows that there are 237 events of CI prediction were followed by CI observed or categorized as hit event. Miss event has a higher number with 259 events. It means that more than half of the CI prediction was wrong. However, the false alarm event has a lower number compared to the hit event with 178 events. It means that most CI observed events are correctly predicted. Meanwhile, the correct negative event dominates in the verification distribution with 2775 objects.

Figure 3 presents the distribution of hit, miss, false alarm, and correct negative frequencies in the $0.2^\circ \times 0.2^\circ$ grid box. In general, hit events are dominantly distributed overland with almost 80% in percentage rather than overseas. Hit events are concentrated over the mountainous area which lined up over the southern part of East Java (the topographic map can be seen in Figure 1) with 44.3% in percentage. Hit events are also distributed over the northern coast of East Java and Madura with a lower percentage than mountainous area as it takes about less than 19% for all hit events. Meanwhile,
the miss events distribution presents an opposite portrait compared to hit events distribution. The most of miss events are focused over the northern coast of East Java and Madura with up to 40% in percentage, but the miss events only take about 25% over the mountainous area. It means that the algorithm predicts CI occurrence better over the mountainous area rather than coastal in this case. This might be caused by the different characteristic between land and sea cloud characteristic which clouds over sea tend to be shallower than inland cloud [23]. This type of cloud has a shorter lifetime rather than cloud lifetime over the mountainous area which can persist for a longer time [24]. On the other hand, the false alarm and correct negative events are distributed more balanced compared to the hit and miss events distribution. However, there are slightly concentrations of false alarm events over the southern part of East Java, particularly over the backside of the mountainous area. This might be caused by the beam blockage effect of the mountain, so some CI events cannot be detected well, especially a cloud with low to middle-level cloud-top.

Figure 3. Spatial distribution of (a) hit, (b) miss, (c) false alarm, and (d) correct negative events in 0.2° × 0.2° grid box over Surabaya and around during JJA period in 2018. Each box represents the accumulation number of events over a grid box divided by the total events of each category.

The temporal distribution of hit, miss, false alarm, and correct negative events are presented in Figure 4. Hit and miss events occur commonly during the afternoon to late evening (13-22 LT) with its peak ranges at 13-16 LT, which reach a percentage of 30.5% and 25.4% of the total hit and miss
events, respectively. The hit events distribution start to decrease significantly until it returns a slight increase at 22-01 LT with 13.1% in percentage. Meanwhile, the miss events distributions decrease more gradually compared to the hit events distribution. Hit and miss events have their lowest frequency at dawn with percentage of 4% for both categories. False alarm and correct negative events have different pattern compared to hit and miss events. False alarm and correct negative events are mostly distributed from morning to afternoon. False alarm events have the maximum frequency at 01-04 LT with 15.4% in percentage, while the correct negative has its maximum at 13-16 LT with 15.1% from 2775 events in total. The false alarm events distribution has the lowest frequency from afternoon to late evening, which has inverse distribution compared to hit and miss events. Meanwhile, the correct negative events distribution has the lowest percentage in the early morning.

![Figure 4](image.png)

**Figure 4.** Temporal distribution of hit, miss, false alarm, and correct negative events over Surabaya and around during JJA period in 2018.

4.2. CI nowcasting performance

The performance of CI nowcasting was assessed using four skill scores (Table 3). Out of four scores, ACC achieves the highest scores at 87.3%, which means more than three-fourths of the prediction, including positive and negative CI predictions, are correctly predicted during the JJA season in 2018. However, this score is lower compared to corresponding research that can achieve more than 90% [13,17]. POD and FAR scores reach 57.1% and 52.2%, respectively. This value means that more than half of the CI event observed are correctly predicted for POD and more than a half of positive CI predictions are not followed by the CI event observed within the next 0-60 minutes. Both POD and FAR score in this paper are slightly better compared to the average of POD and FAR from all location of Walker et al. in the United State (54.6% for POD and 54.3% for FAR), even the highest POD value can reach 85% [13]. However, both scores represent that this algorithm has lower performance compared to corresponding research in Korea and Southern China which reaches more than 75% for POD and lower than 45% for FAR [16-18].

| Skill Score | Value     |
|-------------|-----------|
| ACC         | 87.3%     |
| POD         | 57.1%     |
| FAR         | 52.2%     |
| CSI         | 35.2%     |

Table 3. The performance of CI nowcasting
Although the ACC achieves up to 87.3%, this value is not good enough in representing the overall algorithm performance in predicting rare event (CI event in this case). This is because the non-rare events (correctly predicted of negative CI prediction) dominates in the scores. Therefore, the CSI score is calculated to show a more representative score. CSI score which excludes the correct negative achieves at 35.2%. In this paper, the CSI score gives a lower result compared to the CSI score of CI nowcasting in South Korea which can reach more than 60% [16]. The higher CSI score in that research can be caused by the implementation of machine learning which uses a training data set to calculate a new dynamic threshold of the interest fields. That process will make the thresholds more suitable for the research area instead of using the original static threshold of SATCAST. However, this value means that some factors need to be concerned in CI nowcasting development, especially in Surabaya during JJA period in 2018. Those concerns could be about the threshold values, appropriate interest fields, implementation of current development such as machine learning, or appropriate choice of radar location due to some radar limitations.

The lead time prediction frequencies in CI nowcasting from 237 correctly predicted CI events is shown in Figure 5. It seems that the 10 minutes lead time dominates in the distribution with 29.2% in percentage (70 CI events). The distribution starts to decrease significantly on 20 minutes lead time and reach the minimum frequency on 40 minutes lead time with 11.4% in percentage (27 CI events). The lead time frequency starts to climb again up to 18.2% from 237 events (43 CI events) on 60 minutes lead time. CI nowcasting over Surabaya and around has 32.3 minutes in average with 30 minutes in median which means that CI nowcasting over Surabaya and around has potential to signal the development of convective cloud up to 30 minutes prior to CI event. This value is similar to the average lead time in United State for 30 minutes [13], but relatively lower compared to the average lead time in Korea for 30-40 minutes [16,18] and South-eastern China for up to 60 minutes [17].

The distribution of each lead time values at different altitudes is presented using a stacked bar chart in Figure 6. In general, there is a decreasing frequency over higher altitudes on shorter lead time. On the lead time prediction of 40-60 minutes prior to the CI event, there is more than a half of correctly predicted CI event for each lead time occur in the location with altitudes greater than 500 m or highland. Meanwhile, on the lead time prediction of 30 minutes or less, correctly predicted CI events are dominated in the location with altitudes less than 500 m. For example, CI event happens at more than 500 m altitudes takes only 24.5% from 70 correctly predicted CI events on 10 minutes lead time. On that lead time, the frequency is dominated by CI event over water and 0-200 m with 34.3% (24 CI events) occurs at altitude 0-50 m. Therefore, there is a correlation between the location altitudes and the lead time prediction which at the higher altitudes, the lead time tends to be longer. It can be caused by the barrier effect which slows down the upstream wind, so the cloud can strengthen and persist for
9

a longer time before it begins to precipitate over the mountainous area [24]. In the other hand, the distribution of each lead time prediction over different altitudes may answer why the number of correctly predicted CI events start to climb again at longer lead time up to 43 events at 60 minutes of lead time. That number is dominated by the CI events over higher altitudes which also impacts on the 32.3 minutes of average lead time.

5. Summary and Conclusions
Using the Himawari-8 AHI satellite, a geostationary satellite-based CI nowcasting algorithm named SATCAST was evaluated over Surabaya and around during JJA period in 2018. Three major processes, i.e. cloud masking, object tracking, and CI predicting, were conducted to predict CI event within the next 0-60 minutes. Based on Surabaya weather radar which verified the prediction, a notable accuracy score as high as 87.3% was achieved from the 3449 cloud objects in total. However, that score was followed by 57.1%, 52.2%, and 35.2% of POD, FAR, and CSI score, respectively, which means that the algorithm performance still needs further improvement and development. The 32.3 minutes of average lead time indicates that CI nowcasting over Surabaya and the surrounding area can potentially predict up to 30 minutes prior to the CI event. However, the lead time prediction is quite affected by the altitudes which longer lead time events are more frequent in higher altitudes.

Regarding the period of this assessment, expanding the period of the assessment may result in giving more accurate reflection about this CI nowcasting performance. In the other hand, this algorithm is very dependent on the threshold value of the interest fields. In this paper, the SATCAST threshold value which developed over the middle latitudes region was used, but the algorithm was assessed over tropics. Therefore, adjusting threshold value based on local cloud physics characteristics or implementing dynamic threshold using machine learning in further research may increase the algorithm performance.

Acknowledgements
The authors would like to thank the satellite subdivision of BMKG and Djuanda meteorological station for providing and supporting the data used in this research and everyone who has helped this study.
References

[1] Karmakar S, Quadir D A, Alam R and Das M K 2016 Journal of NOAMI 33 19-33
[2] Weckwerth T M, Parsons D B, Koch S E, Moore J A, Lemone M A, Demoz B B, Flamant C, Geerts B, Wang J and Feltz W F 2004 Bull. Amer. Meteor. Soc. 85 253-278
[3] Weckwerth T M, and Parsons D B 2006 Mon. Wea. Rev. 134 5-22
[4] Weckwerth T M, Pettet C R, Fabry F, Park S, Lemone M A and Wilson J W 2005 J. Appl. Meteorol. 44 285-300
[5] Dixon M and Wiener G 1993 J. Atmos. Oceanic Technol. 10 785-797
[6] Lakhsmanan V, Smith T, Stumpf G and Hondl K 2007 Wea. Forecasting 22 596-612
[7] Kober K and Tafferner A 2009 Meteorol. Z. 1 75-84
[8] Roberts R D and Rutledge S 2003 Wea. Forecasting 18 562-584
[9] Mecikalski J R and Bedka K M 2006 Mon. Wea. Rev. 134 49-78
[10] Sieglaff J M, Cronce L M, Feltz W F, Bedka K M, Pavolonis M J and Heidinger A K 2011 J. Appl. Meteor. Climatol. 50 110-126
[11] Sobajima A 2012 Rapidly Development Cumulus Areas Derivation Algorithm https://cwg.eumetsat.int/res/pdf/ATBD_RapidlyDevelopingCumulusAreas_CWG.pdf
[12] Walker J R and Mecikalski J R 2011 Algorithm theoretical basis document (ATBD) for CI https://www.nsstc.uah.edu/SATCAST/docs/GOESR%20AWG%20ATBD%20Aviation%20ConvectiveInitiation%20v2.0.pdf
[13] Walker J R, Mackenzie W M, Mecikalski J R and Jewett C P 2012 J. Appl. Meteor. Climatol. 51 1931-1949
[14] Besho et al. 2016 J. Meteor. Soc. Japan 94 151-183
[15] Schmit T J, Gunshor M M, Menzel W P, Gurka J J, Li J and Bachmeier A S 2005 Bull. Amer. Meteor. Soc. 88 1079-1096
[16] Lee S, Han H, Im J, Jang E and Lee M 2017 Atmos. Meas. Tech. 10 1859-1874
[17] Zhuge X and Zou X 2018 J. Meteor. Soc. Japan 96 337-353
[18] Matsui T, Chern J, Tao W, Lang S, Satoh M, Hashino T and Kubota T 2016 J. Hydrometeor. 17 1425-1445
[19] Houze R A 2011 Rev. Geophys. 50 RG1001