

**NOAA NESDIS
CENTER for SATELLITE APPLICATIONS and
RESEARCH**

ALGORITHM THEORETICAL BASIS DOCUMENT

Convective Initiation

*Wayne M. MacKenzie, Jr., University of Alabama in Huntsville
John R. Walker, University of Alabama in Huntsville
John R. Mecikalski, University of Alabama in Huntsville*

Version 1.0

June 28, 2010

1	INTRODUCTION	4
1.1	Purpose of This Document.....	4
1.2	Who Should Use This Document	4
1.3	Inside Each Section.....	4
1.4	Related Documents	4
1.5	Revision History	4
2	OBSERVING SYSTEM OVERVIEW.....	5
2.1	Products Generated	5
3	ALGORITHM DESCRIPTION.....	6
3.1	UAH Algorithm Overview	6
3.2	Processing Outline	6
3.3	Algorithm Input	7
3.3.1	Primary Imager Data.....	7
3.3.2	Ancillary Data.....	7
3.4	Theoretical Description.....	8
3.4.1	Physics of the Problem.....	14
3.4.2	Physical/Mathematical Description	14
3.4.3	Algorithm Output.....	17
4	Test Data Sets and Outputs	17
4.1	Simulated/Proxy Input Data Sets	17
4.2	Output from Simulated/Proxy Inputs Data Sets.....	18
5	Practical Considerations.....	18
5.1	Numerical Computation Considerations.....	18
5.2	Programming and Procedural Considerations	19
5.3	Quality Assessment and Diagnostics	19
5.4	Exception Handling	20
5.5	Algorithm Validation	20
6	ASSUMPTIONS AND LIMITATIONS	21
6.1	Performance	22
6.2	Assumed Sensor Performance	22
6.3	Pre-Planned Product Improvements	22
7	REFERENCES	23

List of Figures

Figure 1. High Level Flowchart of the CI Algorithm illustrating the main processing sections.....	7
Figure 2. Threshold for object ground speed for a given object diameter which is parallel to the wind flow.....	11
Figure 3 Schematic diagram showing how a single object can be tracked through time via the temporal overlapping technique.....	11
Figure 4. Schematic showing that the summing of object arrays at T1 and T2 result in values of “-2” for all overlap region pixels, when all object pixels are initially assigned values of “-1”. The ellipse on the left represents an object at T1, while the ellipse on the right represents the same object at T2.....	11
Figure 5. Illustration of how the overlap regions are identified. Object array pixels from T1 (a) and from T2 (b) are summed. The result is a single array of integers (c) with values of “0” where no objects exist (grey), values of “-1” where objects exist but there is no overlap (green or red), and values of “-2” where there is overlapping between T1 and T2 (yellow).....	12
Figure 6. Continuing the example from Fig. X.3, the overlap regions have all been assigned unique ID numbers.....	12
Figure 7. Illustration showing the result after “spreading” ID numbers from the overlap regions to the rest of the space occupied objects both at T1 and at T2.....	13
Figure 8. Final output from the object tracking algorithm. Each object has been assigned a unique ID number that remains consistent from T1 (a) to T2 (b).....	13
Figure 9. CI Algorithm output examples from 5 minute MSG SEVIRI data from 08 June 2007. The top two images are at time 1 (1024 UTC), the middle two images are time 2 (1029 UTC). For the top two rows, the left side is the 10.7 micron channel image and the images on the right are the defined objects. The bottom row on the right is algorithm output valid at 1029 UTC and the bottom image on the left is the actual IR image in the future at 1124 UTC.).....	19

List of Tables

Table 1. CI Algorithm Requirements.....	5
Table 2. Channel numbers within GOES-R ABI and wavelengths used for the CI algorithm.....	6
Table 3. Current CI indicators currently being tested for use within MSG operations.....	9
Table 4. Current GOES infrared interest fields used within the current CI algorithm.....	15
Table 5. Comparison of the number of spectral tests triggered and impact on the accuracy	17
Table 6. Product Quality Information.....	18
Table 7. Metadata Information.....	18
Table 8. Dichotomous forecast verification 2x2 contingency table.....	20
Table 9. Contingency table from validation using MSG SEVIRI.....	20
Table 10. Contingency table from validation using RAMS simulated datasets.....	21

1 INTRODUCTION

1.1 Purpose of This Document

The convection initiation theoretical basis document (ATBD) provides a high level description of and the physical basis for the assessment of convection initiation derived from the Advanced Baseline Imager (ABI) flown on the GOES-R series of NOAA geostationary meteorological satellites. The Convection Initiation (CI) algorithm provides an assessment of the clouds that may precipitate. The CI algorithm is designed to monitor the growth of non-precipitating clouds, and once a series of spectral and temporal thresholds are met, that cloud is identified as likely to have a radar reflectivity greater than 35 dBZ within 0-2 hours.

1.2 Who Should Use This Document

The intended user of this document are those interested in understanding the physical basis of the convection initiation algorithm and how to use the output of this algorithm to determine clouds which may produce radar reflectivities greater than 35 dBZ. This document also provides information useful to anyone maintaining or modifying the original algorithm.

1.3 Inside Each Section

This document is broken down into the following main sections.

- **System Overview:** Provides relevant details of the ABI and provides a brief description of the product generated by the algorithm.
- **Algorithm Description:** Provides all the detailed description of the algorithm including its physical basis, its input and its output. Validation will also be addressed.
- **Assumptions and Limitations:** Provides an overview of the current limitations of the approach and gives the plan for overcoming these limitations with further algorithm development.

1.4 Related Documents

This document currently does not relate to any other document outside of the specifications of the F&PS and to the references given through out.

1.5 Revision History

Version 1.0 of this document was created by Wayne M. MacKenzie, Jr., John R. Walker and John R. Mecikalski of the University of Alabama in Huntsville and its intent was to accompany the delivery of the version 1.0 algorithm to the GOES-R AWG Algorithm Integration Team (AIT).

2 OBSERVING SYSTEM OVERVIEW

Overview of the algorithm(s), including the objectives, characteristics of the instrument(s) referencing rather than repeating requirements that provides the input data and retrieval strategies.

2.1 Products Generated

The convection initiation (CI) algorithm produces a binary field at 2 km spatial resolution of areas where CI has a high likelihood of occurring. The product uses a spectral thresholding technique, which tracks clouds within their early stages of development, and monitor their spectral characteristics. If a large majority of the spectral “interest fields” thresholds are exceeded, then the pixels within the cloud object are flagged for having a high likelihood for CI.

Name	User & Priority	Geographic Coverage (G, H, C, M)	Vertical Resolution	Horizontal Resolution	Mapping Accuracy	Measurements Range	Measurements Accuracy	Product Refresh Rate/Coverage Time	Vendor Allocated Ground Latency	Product Measurement	Temporal Coverage Qualifiers
Convective Initiation	GOES-R	C	N/A	2 km	1 km	Binary Yes/No detection	70% Probability of Correct Detection	5 min	159 sec	N/A	Day and night
Convective Initiation	GOES-R	M	N/A	2 km	1 km	Binary Yes/No detection	70% Probability of Correct Detection	5 min	159 sec	N/A	Day and Night

Table 1. CI Algorithm Requirements.

2.2 Instrument Characteristics

The algorithm will use the various spectral channels within GOES-R (listed in Table 1). Table 1 summarizes the projected channel.

Channel Number	Wavelength (μm)	Projected to be used in CI processing
1	0.47	
2	0.64	
3	0.86	
4	1.38	
5	1.61	
6	2.26	
7	3.9	
8	6.15	X
9	7.0	X
10	7.4	X

11	8.5	X
12	9.7	X
13	10.35	
14	11.2	X
15	12.3	X
16	13.3	X

Table 2. Channel numbers within GOES-R ABI and wavelengths used for the CI algorithm.

The algorithm relies on the infrared channels only for the algorithm to have both day and night continuity. The performance of the algorithm may be sensitive to any instrument noise.

3 ALGORITHM DESCRIPTION

This is a complete description of the algorithm at the current level of maturity (which will improve with each revision).

3.1 Algorithm Overview

Mecikalski and Bedka (2006) first showed that one can track growing cumulus, monitor their spectral properties, and using a set of threshold-based indicators to determine the likelihood that a particular cumulus will precipitate. The thresholds incorporate spectral differences that give information into the cloud-top phase of clouds as well as their respective location within the troposphere to determine cloud maturity. In addition to that information, the growth of the cloud through the troposphere over two successive images can be achieved using several of the spectral channels as outlined in Table 1. Knowledge of this information can provide information into the stage of development of a cumulus cloud, and thus identify whether a cloud will precipitate within a hour one time period.

3.2 Processing Outline

The processing outline of the CI algorithm is summarized in Figure 1. The current CI algorithm uses satellite data in netCDF format for input into Fortran based processing code. The ABI data (current data along with one previous time period) and the cloud-typing algorithm (current data along with one previous time period) are required to begin processing the CI algorithm.

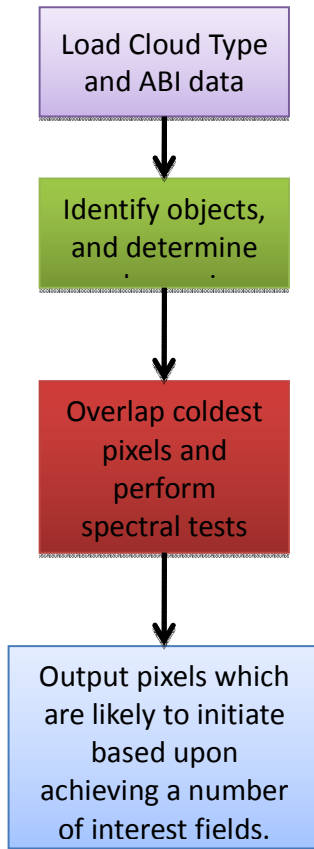


Figure 1. High Level Flowchart of the CI Algorithm illustrating the main processing sections.

3.3 Algorithm Input

This section describes the input needed to process the CI algorithm. Currently, the code is developed in Fortran 90.

3.3.1 Primary Imager Data

- In its current stage, the CI algorithm requires the use of the brightness temperatures from channels 8-12 and 14-16 within the IR (Table 1), as well as the current image and the previous image in order to process all stages of the algorithm outlined in figure 1.

3.3.2 Ancillary Data

The use of the ABI AWG Cloud Team Cloud Type Product is needed for the CI algorithm. The current time dataset along with the previous time dataset is required for processing. Any cloud type data dependencies as outlined in the Cloud Type ATBD are also inherently necessary for the CI algorithm.

3.4 Theoretical Description

3.4.1 Physics of the Problem

The CI algorithm tracks moving clouds using an object identification and tracking technique and monitors the growth of the clouds using a spectral thresholding technique using the thresholds listed in table 2. It is important to note that this algorithm is designed for identifying clouds, which have the potential for growth, thus mature clouds are omitted. It is important to note that other studies have used similar methods for monitoring mature mesoscale convective complexes (Carvalho and Jones 2001, Machado and Laurent 2004 and Vila et al. 2008).

Objects are identified using the AWG Cloud Typing algorithm. If clouds are identified as water, supercooled or mixed phase, those pixels are deemed immature for the purpose of CI identification. The algorithm searches around each pixel to determine a gap in pixels, and this is the method for determining whether a pixel is an independent object or part of a larger object. For the 100% delivery, the algorithm will take the larger objects and focus on the convectively active regions of the larger object. A size threshold will be used to determine whether an object is too large and a peak detection technique using the 11.2 μm channel to pull out the convectively active regions. This will also help with mitigating any false cloud detections by the AWG Cloud Typing algorithm since falsely identified large objects will be removed if there are no minimum temperature peaks.

The CI algorithm uses an object tracking technique, which is an overlap method. This overlap technique exploits the high temporal resolution of GOES-R. Currently, the tracking algorithm does not perform well for fast moving clouds and if the temporal resolution is greater than 5 minutes.

Once the objects have been identified and tracked, the coldest 25% of the pixels within each of the object are averaged using a quick sorting routine in which all the brightness temperature pixels are listed, and then organized from coldest to warmest, and the coldest 25% of the pixels are used to average all spectral channels used within the CI algorithm.

Using the object average brightness temperature from each of the spectral channels used in table 1, a series of infrared spectral threshold tests will be performed (as listed in Table 2). The objects spectral information difference will occur using the object average temperature, and temporal differencing will occur over the previous time object average brightness temperature for each spectral channel.

If the object meets 7 of the 12 spectral tests, all pixels within the object will be highlighted as a high likelihood of initiation (initiation is defined as the object achieving a 35 dBZ radar echo).

In the following sections, the four main components of the algorithm will be discussed in detail. The four main components are:

- 1) Object Identification
- 2) Object Tracking
- 3) Spectral Interest Field Tests
- 4) Determining whether there is a high likelihood for Convective Initiation.

Interest Field	Physical Basis (Mecikalski et al., 2009)	Critical Value
6.2-10.8 μm	Cloud Depth	-30°C to -10°C
6.2-7.3 μm	Cloud Depth	-25°C to -5°C
11.2 μm	Cloud Depth/Glaciation	-20°C to 5°C
8.7-10.8 μm	Glaciation	-10°C to -1°C
Tri-channel Diff	Glaciation	-10°C to 0°C
5 min Tri-Channel	Glaciation Trend	>0°C
5 min 12.0-10.8 μm	Cloud Depth	>0.5°C
12.0-10.8 μm	Cloud Depth	-3°C to 0°C
5 min 10.8 μm	Cloud Growth	< -1.33°C
5 min 6.2-7.3 μm	Cloud Depth Trend	>0°C
5 min 6.2-10.8 μm	Cloud Depth Trend	>0.5°C
13.4-10.8 μm	Cloud Depth	-20° to -5°C

Table 3. Current CI indicators currently being tested for use within MSG operations.

Object Identification

The purpose of this algorithm is to take large objects, such as within a large cloud deck, and focus on the convective elements to perform the spectral and temporal tests to determine likelihood for convective initiation. This algorithm begins with a driver subroutine that handles the selection of cloud peaks and starts the “blobing” algorithm that defines the area of the object and defines a gap that no other object can begin to be defined in. The driver function takes in a blank array that is where the objects will be defined, the cloud mask, the brightness temperature array, and the maximum number of rows and columns. It then turns the 2D brightness temperature array into a 1D array that can be sorted and it creates a 1D list of the original indices of the temperature in question. Both of these arrays are sorted based on the brightness temperature. The algorithm takes the temperatures in order from coldest to warmest and compares the associated indices against the cloud mask and the output array to determine if the object is a cloud and if the area is not designated as something else already in the output. Finally it verifies that the point is actually a peak and then it passes the output array, the row and column in

question, a maximum temperature derived from an average of the temperatures within the cloud sections, an ID for the object, the cloud mask, the brightness temperature array, and the row and column in question again (meaning shown within the “blobing” algorithm.). Once the blobing algorithm returns the object ID is incremented and we look at the next coldest point.

Within the blobing algorithm it just calls a function that takes in the same arguments as the subroutine. Since this algorithm starts in the center then checks surrounding points, then the surrounding point's surrounding points in a “blobing” out style, It also compares against the cloud mask, then check if the point has already been visited. It then compares against how far the algorithm has come from the original point to keep it from spreading too far. If all that succeeds, the algorithm determines if the point should be given the objects ID number or be allocated as “skirt” based on distance from center. It then does the same thing for the points around it passing in the original point as the second point passed in that I referenced above. Once it has determined the while area it returns to the driver subroutine.

The above description will be included within the 100% delivery.

Object Tracking Methodology

The object tracking algorithm is based upon the simple concept of temporal overlapping. Because of this restriction, there is a weakness in that if the mean flow is fast and the object size is small, there may not be temporal overlap between the two times. This is somewhat mitigated from the fact that growing clouds will increase in horizontal size as well as vertically which minimizes this impact. Figure 2 shows the threshold where an object may be missed for a given object size and speed. This assumes the object size remains constant between time 1 and time 2.

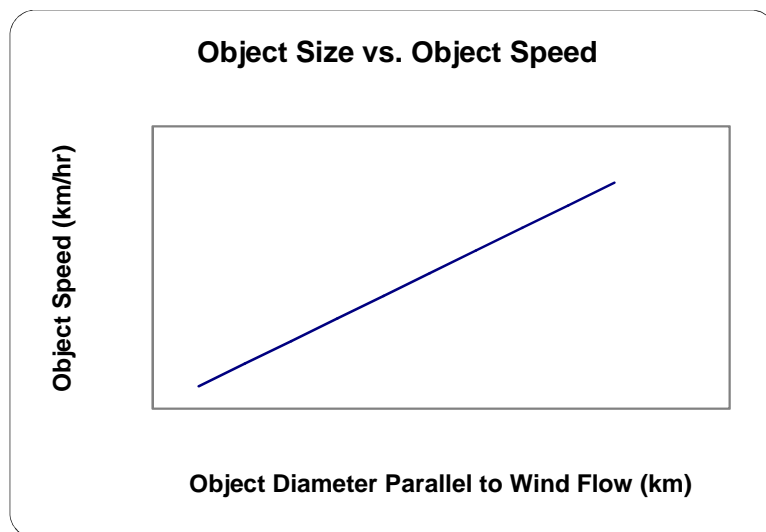


Figure 2.. Threshold for object ground speed for a given object diameter which is parallel to the wind flow.

Temporal overlap is where an object that occupies a space at Time 1 (T1) can be assumed to be the same object at Time 2 (T2) as long as its position at T2 coincides or “overlaps” with part of the space it occupied at T1 (Fig. 3).

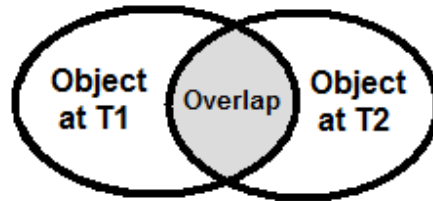


Figure 3 Schematic diagram showing how a single object can be tracked through time via the temporal overlapping technique.

The first step to applying this method to the tracking of cloud features is to identify all desired cloud types resulting from a trustworthy satellite-based cloud classification scheme, and then mask out all other irrelevant or undesired cloud types. In the given algorithm, all interesting cloud feature pixels from both T1 and T2 are assigned the integer, “-1”. Next, the resulting arrays are summed, so that all temporally overlapping regions can be identified wherever the integer “-2” is present (Fig. 4). Obviously, non-overlapping region pixels are left at values of “-1” after the arrays are summed. Fig. 5 better demonstrates this overlap region identification with real, scaled down output from the computer algorithm, using simplified input case data.

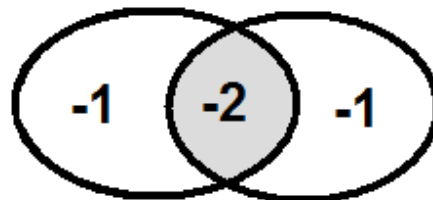


Figure 4. Schematic showing that the summing of object arrays at T1 and T2 result in values of “-2” for all overlap region pixels, when all object pixels are initially assigned values of “-1”. The ellipse on the left represents an object at T1, while the ellipse on the right represents the same object at T2.



Figure 5. Illustration of how the overlap regions are identified. Object array pixels from T1 (a) and from T2 (b) are summed. The result is a single array of integers (c) with values of “0” where no objects exist (grey), values of “-1” where objects exist but there is no overlap (green or red), and values of “-2” where there is overlapping between T1 and T2 (yellow).

The next step performed in this algorithm is to assign each individual overlap region a unique, positive integer identification number (ID number). The algorithm loops through the summed array, searching for overlap pixels, valued at “-2”. Then, whenever the first group of overlap pixels is encountered, a counter, which is initialized at “0”, is incremented up to “1”, and each pixel in that overlap group is assigned an ID number of “1”. The next time a group of overlap pixels is encountered, the counter is incremented up by one more integer value, and the process continues until each overlap region is assigned a unique integer ID number (Fig. 6).

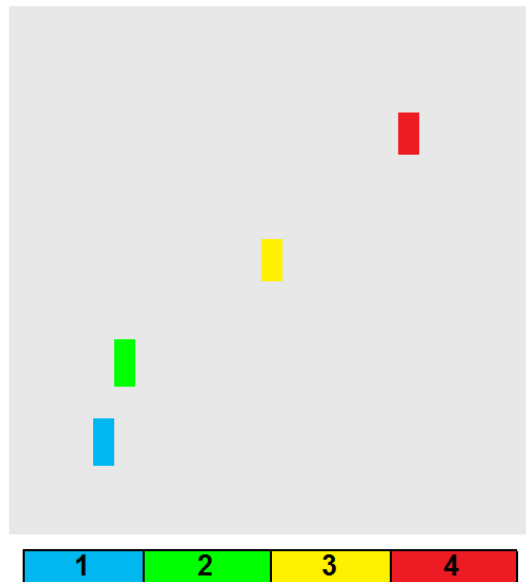


Figure 6. Continuing the example from Fig. X.3, the overlap regions have all been assigned unique ID numbers.

Next, the algorithm loops through the ID number overlap array and iteratively “spreads” each overlap region’s ID number left, right, up, and down across the entire space occupied by each given object through both T1 and T2. The final product from this step is a single collective array where the pixel space occupied by individual objects that overlapped their selves from T1 to T2 is accordingly assigned an integer ID number that is unique to a given object (Fig. 7).

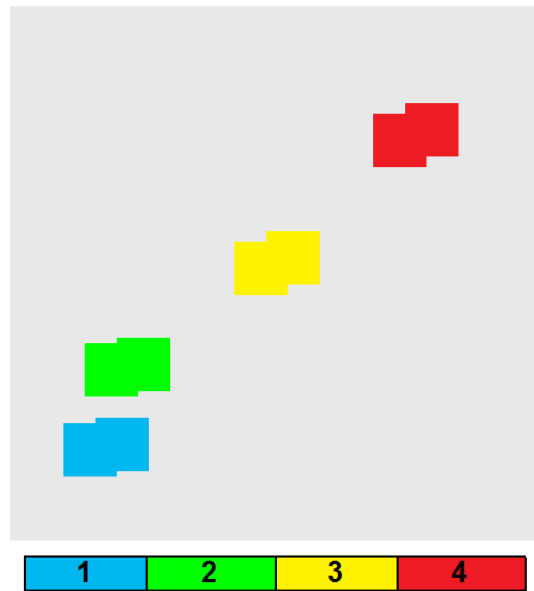


Figure 7. Illustration showing the result after “spreading” ID numbers from the overlap regions to the rest of the space occupied by objects both at T1 and at T2.

The final step in the object tracking algorithm is to separate the recently created T1/T2 collective array into its original T1 and T2 components, now that unique integer ID numbers have been assigned to each object (Fig. 8). From there, cloud top characteristic trends and interest fields can be derived consistently for whole, individual objects.

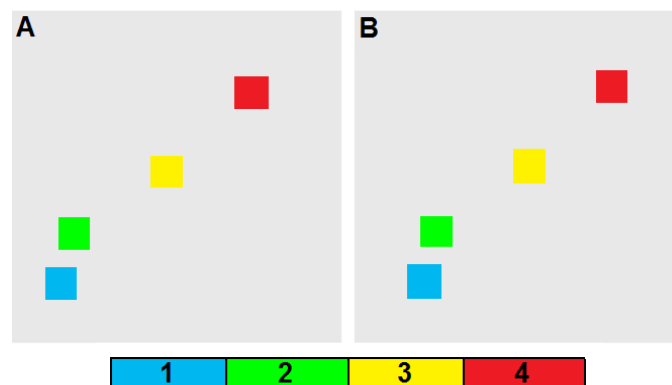


Figure 8. Final output from the object tracking algorithm. Each object has been assigned a unique ID number that remains consistent from T1 (a) to T2 (b).

Spectral Tests

In order to perform the spectral tests, all pixels within each object are sorted from coldest to warmest using the 11.2 μm spectral channel. Then the 25% coldest pixels are averaged to come away with an average brightness temperature for the particular channel. Using the same pixels used to perform the averaging for the 11.2 μm channel, the average brightness temperature is found for each of the other spectral channels as defined in Table 1.

For each object, the spectral tests defined in Table 2 are performed (including temporal tests using the previous image). If the resultant value from each test is within the critical value, then that test receives a 1, and all tests, which the critical value was not achieved, then a value of 0 is recorded.

Determination of Likelihood for CI

The spectral tests performed above are summed, and empirical results have shown that if 7 or greater of the tests have triggered, then there is a high likelihood for convective initiation. The object is then flagged for CI in a separate array, which contains the summed values of each of the spectral tests binary results. For example, if 5 of the spectral tests were within the critical value, then the resultant summed value of the binary spectral test is 5, which is less than the 7 or greater required to be likely for CI.

3.4.2 Physical/Mathematical Description

Interest Field Development

Meciklaksi and Bedka (2006) outlined several spectral threshold interest fields, which are used within the current GOES satellite (table 3). This interest fields provide information into the growth characteristics of the cloud. Knowing this information will provide lower false alarms because knowing the vertical location of the cloud within the atmosphere will remove the erroneous cooling rates from the 11 micron channel. Studies have shown that the spectral channels within the current GOES satellite provide such information and allows for effective monitoring of growing cumulus clouds. Having more spectral information within GOES-R will facilitate the use of more cloud property information, which will further reduce false alarms.

CI interest field	Critical value
10.7 μ m T (one score)	$^{\circ}$ C
10.7 μ m T time trend (two scores)	$< -4 \text{ }^{\circ}\text{C (15 min)}^{-1}, \Delta T (30 \text{ min})^{-1} < \Delta T (15 \text{ min})^{-1}$
Timing of 10.7 μ m T drop below 0 $^{\circ}$ C (one score)	Within prior 30 min
6.5-10.7 μ m difference (one score)	-35 $^{\circ}$ C to -10 $^{\circ}$ C
13.3-10.7 μ m difference (one score)	-25 $^{\circ}$ C to -5 $^{\circ}$ C
12.0-10.7 μ m difference (GOES-11)	-3 $^{\circ}$ C to 0 $^{\circ}$ C
6.5-10.7 μ m time trend (one score)	$> 3^{\circ}\text{C (15 min)}^{-1}$
13.3-10.7 μ m time trend (one score)	$> 3^{\circ}\text{C (15 min)}^{-1}$
12.0-10.7 μ m time trend (GOES-11)	$> 2^{\circ}\text{C (15 min)}^{-1}$

Table 4. Current Operational GOES infrared interest fields used within the current CI algorithm.

The 6.7-10.7 micron spectral difference provides information on cloud top height location relative to the tropopause (Mecikalski and Bedka 2006). Typically the difference is negative because the surface temperature is warmer than the upper-troposphere where the water vapor channel weighting function peaks. A positive difference corresponds to clouds at or above the tropopause (Ackerman 1996; Schmetz et al. 1997). This information can identify clouds which are immature or which are growing into the midlevel of the atmosphere. The temporal trend of this interest field allows for the growth of the cloud with respect to the tropopause to be monitored over time. Essentially, this field allows for the determination of how fast the cloud is moving through the troposphere. This has implications of ensuring that the 10.7 micron temporal cooling rate is accurate and minimize large cooling rates originating from clouds growing rapidly within the boundary layer below the capping inversion. This situation will not be shown within the 6.7 micron spectral channel because the weighting functions peak higher in the tropopause.

The 12.0-10.7 micron spectral difference, known as the “split window” technique, is typically used for identifying the presence of cirrus, volcanic ash, and deep convective clouds. Inoue (1987) has found that near-zero 12.0-10.7 micron spectral differences provide a means to identify areas convective rainfall. This is an enhancement to the Griffith et al (1978) method which uses a $< -20^{\circ}\text{C}$ 10.7 micron brightness temperature. The purpose for this interest field is to highlight areas, which are evolving into a convection rainfall cloud. The temporal trend of the spectral channel allows a more effective approach to monitoring this transition.

The 13.3-10.7 micron spectral difference does provide information for growing cumulus clouds since Mecikalski and Bedka (2006) found that this spectral difference has different characteristics from mature cumulus to pre-ci cumulus, similar to the 6.5-10.7 micron spectral difference. Mecikalski et al. (2008) found using a principal component analysis found that the 13.3-10.7 micron channel is one of the most important interest fields. It is hypothesized that this spectral difference was found to be important because of the 8 km spatial resolution of the 13.3 micron channel on the current GOES. If the 13.3 micron channel saturates, then it is very likely that the storm would initiate because of the poor spatial resolution. Since the 13.3 micron channel will have a spatial resolution of 2 km

on GOES-R, it is uncertain how important this channel will be due to the relatively few studies performed on this spectral channel.

Mecikalski et al. (2010) contains a detailed explanation of the best uses for the infrared fields for pre-convective clouds. Further, Siewert et al. (2009) discusses how to use Meteosat Second Generation (MSG SEVIRI) data for convective initiation purposes over South Africa using a different tracking methodology. However, Siewert et al. (2009) does demonstrate the importance of using multiple spectral tests within a convective initiation algorithm. Mecikalski et al. (2010) examined all the possible spectral tests and divided them into three physical categories: 1) cloud depth, 2) cloud top glaciation and 3) updraft strength. From these three physical categories, tests were performed to determine which spectral tests are redundant and which ones contain the most information. Thus, the result from the study was to determine the average value from all of the pre-convective events and also examine which particular spectral tests are important. For the GOES-R CI algorithm, this study supplied the spectral interest fields (or spectral tests) to be used. The main components of using the additional spectral channels on MSG are to exploit the three categories. The results from this study was taken from the average value and taking a standard deviation from the average to develop the critical values. Some changes were required to account for the change from 15-minute temporal resolution to 5-minute temporal resolution since the original study was performed using 15-minute data.

The Mecikalski and Bedka (2006) algorithm is currently in operation using the spectral channels on the current GOES series. The current algorithm has a high probability of detection (upwards of 90% when all interest field thresholds are met) but high False Alarm Rate (Mecikalski et al. 2008). The high false alarm rate is caused from a pixel based tracking and verification and an object based verification will give more accurate probability of detection and false alarm rate, as we have found from our current validation. GOES-R will allow for the addition of other spectral interest fields in order to constrain false alarm rates.

Determination of Likelihood for CI

Empirical results have shown that 7 or greater of the 12 spectral tests give the optimal statistics. Using 213 cases over Europe during the summer of 2007, tests were performed ranging over all sums of spectral tests. Table 4 contains the accuracy for each of the summation of spectral tests. Accuracy is defined as the sum of hits plus correct negatives divided by the total of all four values within the validation contingency table. Notice that 7 or greater spectral tests is where the accuracy is maximized.

# of spectral tests triggered	Accuracy
1 or greater	57.7%
2 or greater	60.56%
3 or greater	61.97%
4 or greater	67.6%
5 or greater	69.95%
6 or greater	76.06%
7 or greater	80.75%
8 or greater	73.24%
9 or greater	53.05%
10 or greater	48.83%
11 or greater	43.2%
12 or greater	42.25%

Table 5. Comparison of the number of spectral tests triggered to impact on the accuracy.

3.4.3 Algorithm Output

The final output of this algorithm is a binary field of whether a particular object will be likely to initiate (achieving a 35 dBZ radar reflectivity). All pixels within the object will be highlighted with a 2 km pixel resolution. Quality flags, product quality information and metadata are also included as output.

Quality Flags

Quality flags for the product are defined as follows:

Bit 1: 0=good, >0 = bad

Bit 2: 0=good Level 1B data, 1=bad Level 1B data

Bit 3: 0=clear, 1=cloudy

Bit 4: 0= LZA <=65 degrees, 1 = LZA > 65 degrees

Bit 5: 1=missing data

Product Quality Information

Byte	Bit	Flag	Source	Value
0	0	Satellite zenith angle block-out zone	L1B	1=zenith angle>65° or lat>66°; 0=OK
	1	Cloud Type Algorithm Input	Cloud Type	1=bad data; 0=OK
	2	Level 1B data	L1B	1=bad data; 0=OK
	3	Pre-Convective Cloud Object Flag	CI	1=Cloud Object; 0=No Cloud Object
	4	CI Yes/No	CI	1=No CI Likely; 0=CI Likely
	5	Not Used		
	6	Not used		
	7	Not used		
1	0-7	Number of CI Interest Fields Triggered	CI	Number of CI Interest Field Triggered within Object ranging from 0 to 12

Table 6. Product Quality Information

Metadata

The metadata needed for the CI algorithm is as follows:

Number of Objects Time 1
Number of Objects Time 2
Number of Tracked Objects
Average number of passed spectral tests over all objects
Average value of each spectral test calculated over all objects
Average number of pixels within all objects at Time 1
Average number of pixels within all objects at Time 2
Average number of pixels within tracked objects

Table 7. Metadata information.

4 Test Data Sets and Outputs

4.1 Simulated/Proxy Input Data Sets

Proxy datasets used for this study is the SEVIRI dataset. Current research is currently underway using the spectral channels on SEVIRI and creating a robust dataset of threshold values using those available spectral channels. The object tracking has been tested using RAMS simulated radiances, and also is currently being tested on SEVIRI 5-minute rapid scan.

4.2 Output from Proxy Input Data Sets

The output from the proxy data sets will resemble the planned output for GOES-R. However, since SEVIRI data has a horizontal spatial resolution of 3 km at nadir, the output from SEVIRI examples will be larger. Also, when statistics are derived using SEVIRI data, it is important to note the increase in horizontal spatial resolution of SEVIRI may cause some areas of small-scale convective initiation to be missed.

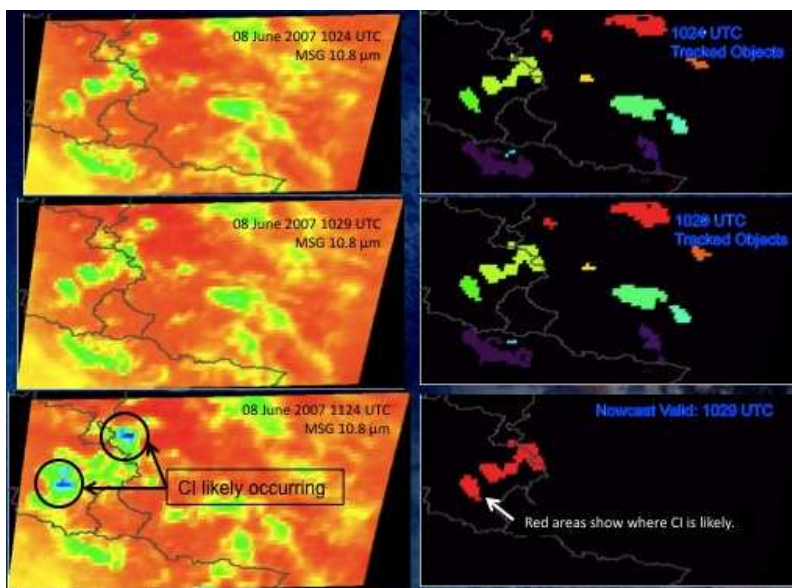


Figure 9. CI Algorithm output examples from 5 minute MSG SEVIRI data from 08 June 2007. The top two images are at time 1 (1024 UTC), the middle two images are time 2 (1029 UTC). For the top two rows, the left side is the 10.7 micron channel image and the images on the right are the defined objects. The bottom row on the right is algorithm output valid at 1029 UTC and the bottom image on the left is the actual IR image in the future at 1124 UTC.

4.2.1 Precision and Accuracy Estimates

The validation complete has shown that the algorithm exceeds the requirement of greater than 70% accuracy. Accuracy is defined as the sum of hits plus correct negatives divided by the sum of all events. Two different validation datasets were used to examine algorithm performance.

For the first set of cases, was using 216 events from the MSG SEVIRI instrument from the summer 2007 from different convective days from southern Germany. The contingency table from the results is listed in Table 8.

HITS	FALSE ALARMS
107	20
MISSES	CORRECT NEGATIVES

16	41
----	----

Table 8. Contingency table from validation using MSG SEVIRI.

The second validation dataset was from a case day from the RAMS simulated ABI channels which allows for a test using the high spatial and temporal resolution data. The case day is 28 June 2005 and was a convective day over the front range in eastern Colorado. Table 9 is a contingency table from the results of this day.

HITS	FALSE ALARMS
21	9
MISSES	CORRECT NEGATIVES
7	37

Table 9. Contingency table from validation using RAMS simulated datasets.

4.2.2 Error Budget

The F&PS requirement for the convective initiation algorithm is an accuracy of 70% or greater. From the two different independent datasets, the accuracy was 80.75% for the MSG dataset and 78.9% for the RAMS simulated dataset. This exceeds the requirement.

5 Practical Considerations

5.1 Numerical Considerations

Since the object tracking requires the use of the use of the cloud-typing algorithm, it is required to be processed first. Once that is received, using the current time image and previous image in time, object tracking can be processed, along with the interest field calculations. The object tracking code may be slowed down when there is a large number of defined objects, but this can be mitigated by code enhancements such as parallelize the code or other efficiencies such as subsetting the domain.

5.2 Programming and Procedural Considerations

The CI object tracking as well as the temporal spectral interest fields requires the use of the current time and the previous image in time. Other than the requirement for the current image and previous image in time, the CI algorithm is an object-based algorithm, which provides an end result of a pixel-by-pixel image.

5.3 Quality Assessment and Diagnostics

The quality flags will contain product quality by passing in information from the quality flags from the input data. Also, the metadata and product quality information output will give assessments to determine how the algorithm is doing with respect to identifying cloud objects and tracking those objects. This is in addition to ensuring that the spectral tests are working sufficiently.

5.4 Exception Handling

The quality control flags for CI algorithm will be checked and inherited from the flagged Level 1b sensor input data, including bad sensor input data, missing sensor input data and validity of each channel used; and will also be checked and inherited from the ABI cloud mask at each pixel.

The CI Algorithm also expects the Level 1b processing to flag any pixels with missing geolocation or viewing geometry information.

5.5 Algorithm Validation

The validation of the CI algorithm should be object-based. Since we are concerned with whether a particular cloud will convectively initiate, object-based approach to validating the algorithm is the best approach since it will give accurate statistical information.

The validation strategy will include a full contingency-based statistical analysis. Since validation of this product is a dichotomous forecast, 2x2 contingency table validation can easily be performed (Table 7). Not only will probability of detection (POD) and false alarm ratio (FAR), but also probability of false detection (false alarm rate), threat score (critical success index), Heidke skill score, and bias score. Knowing and publishing all these variables will allow for a complete validation, which is important for the purpose of providing a robust algorithm.

Following the objects that have been flagged for convective initiation using the method outlined in section 3.1.3 easily performs object-based validation. At each satellite time, each pixel within an object will be corrected for parallax and search the nearest WSR-88D lowest elevation angle. Within each defined object, the radar image at the same within the object is searched for the occurrence of a 35 dBZ anywhere within the defined object.

In order to develop statistics within the 2x2 contingency table, we must also examine all areas of the first occurrence of a 35 dBZ radar echo. If an area of 35 dBZ radar reflectivity is found and no object pixels are found within that area, then this must be documented. We must also explore areas where CI was not forecasted and no 35 dBZ radar echo was found in order to fill the 2x2 contingency table. Following this procedure and performing a 2x2 contingency table will allow for the statistics to ensure the CI algorithm is robust.

CI Forecasted? Yes	CI Forecasted? Yes
CI Occurred? Yes	CI Occurred? No
CI Forecasted? No	CI Forecasted? No
CI Occurred? Yes	CI Occurred? No

Table 10. Dichotomous forecast verification 2x2 contingency table.

An issue with validating the algorithm is gaining access to radar data over Europe. For the validation study, we were able to gain access to radar data over southern Germany during the summer of 2007. We continue to attempt to gain access radar data to ensure algorithm validation.

To mitigate some of the issues with gaining access to radar data over Europe, the Cooperative Institute for Research in the Atmosphere (CIRA) at Colorado State University has performed several RAMS model simulations over several convective days and provided ABI simulated datasets to run the CI algorithm. Model derived simulated radar reflectivity has been used to validate the algorithm for the day that was provided. CIRA is providing more case days to validate.

6 ASSUMPTIONS AND LIMITATIONS

The following sections describe the current limitations and assumptions in the current version of the CI algorithm.

6.1 Performance

The following assumptions have been made in developing and estimating the performance of the CI algorithm. The following list contains the current assumptions and proposed mitigation strategies.

1. We assume that the WSR-88D radar reflectivity values are accurate. The WSR-88D radar available from NCDC is quality controlled and thus the risk is mitigated.
2. For correction of the satellite parallax issue, we assume that the data used for cloud-top pressure from GOES sounder is accurate. This has been validated in the literature and for our purpose of radar based validation the error is minimal.
3. It is assumed that in general, clouds, which are growing vertically over time, are also growing horizontally. However, in the instance where that is not the case, there could be missed CI events in the event the object sizes are small and are moving fast.

6.2 Assumed Sensor Performance

It is assumed that the sensor will meet its current specifications. However, the algorithms will be dependent on the following instrumental characteristics.

- Errors in navigation will affect the ability of the temporal overlap tracking technique to identify overlap regions.
- 5-minute temporal resolution is required as a minimum for effective object tracking.
- Sensor accuracy is important since the spectral threshold tests require accurate measurement especially when using 5-minute data. The changes within cloud top temperatures will be on the order of the sensor accuracy over 5 minutes in most cases.

6.3 Pre-Planned Product Improvements

Currently examining a methodology used within current GOES (adapted from Zinner et al. 2008) to track objects using a motion field. This would allow for higher temporal resolution of satellite

data for the CI algorithm to effectively operate. This would require the input of the GOES-R AWG Derived motion vectors within the CI algorithm. Research is on-going within this area testing the algorithm performance using the AWG wind product.

7. REFERENCES

Ackerman, S.A., 1996: Global Satellite Observations of Negative Brightness Temperature Differences between 11 and 6.7 μm . *J. Atmos. Sci.*, **53**, 2803–2812

Carvalho, L.M.V., and C. Jones, 2001: A Satellite Method to Identify Structural Properties of Mesoscale Convective Systems Based on the Maximum Spatial Correlation Tracking Technique (MASCOTTE). *J. Appl. Meteor.*, **40**, 1683–1701.

Inoue, T., 1987: An instantaneous delineation of convective rainfall area using split window data of NOAA-7 AVHRR. *J. Meteor. Soc. Japan*, **65**, 469-481.

Machado, L.A.T., and H. Laurent, 2004: The Convective System Area Expansion over Amazonia and Its Relationships with Convective System Life Duration and High-Level Wind Divergence. *Mon. Wea. Rev.*, **132**, 714–725.

MacKenzie, W.M., C. Siewert, J.R. Meciklaksi, J.R. Walker, and E.W. McCaul, 2008: Enhancements to nowcasting convective initiation within the 0-1 hour timeframe. In Preparation for *Wea. Forecasting*.

Mecikalski, J.R., and K.M. Bedka, 2006: Forecasting Convective Initiation by Monitoring the Evolution of Moving Cumulus in Daytime GOES Imagery. *Mon. Wea. Rev.*, **134**, 49–78.

Mecikalski, J.R., K.M. Bedka, S.J. Paech, L.A. Litten, 2008: A Statistical Evaluation of GOES Cloud-top Properties for Nowcasting Convective Initiation. *Mon. Wea. Rev.*, In Press.

Mecikalski, J.R., W. M. MacKenzie, Jr., M. Koenig, and S. Muller, 2010: Cloud-top properties of growing cumulus prior to convective initiation as measured by meteosat second generation. Part 1: Infrared Fields. *J. Appl. Meteor.*, **49**, 521–534.

Roberts, R.D., and S. Rutledge, 2003: Nowcasting Storm Initiation and Growth Using GOES-8 and WSR-88D Data. *Wea. Forecasting*, **18**, 562–584.

Schmetz, T.J., S.A. Tjemkes, M. Gube, and L. van de Berg, 1997: Monitoring deep convection and convective overshooting with METEOSAT. *Adv. Space Res.*, **19**, 433-441.

Siewert, C. W., M. Koenig, and J. R. Mecikalski, 2009: Application of Meteosat second generation data towards improving the nowcasting of convective initiation. *Meteor. App.*, doi:10.1002/met.176

Vila, D.A., L.A.T. Machado, H. Laurent, and I. Velasco, 2008: Forecast and Tracking the Evolution of Cloud Clusters (ForTraCC) Using Satellite Infrared Imagery: Methodology and Validation. *Wea. Forecasting*, **23**, 233–245.

Zinner, T., H. Mannstein, and A. Tafferner, 2008: Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. *Meteorol. Atmos. Phys.*, doi:10.1007/s00703-008-0290-y.