P5.4 OVER-OCEAN VALIDATION OF THE GLOBAL CONVECTIVE DIAGNOSTIC

David W. Martin
University of Wisconsin-Madison, Wisconsin
Richard A. Kohrs
University of Wisconsin-Madison, Wisconsin
Frederick R. Mosher
Aviation Weather Center, Kansas City, Missouri

1. INTRODUCTION

The Global Convective Diagnostic (GCD) is a binary index of deep, moist convection. It depends on a temperature threshold applied to the difference between two infrared (IR) images of a geostationary satellite. Mosher (2001) created the GCD to help the Aviation Weather Center meet the needs of the aviation community for current information on thunderstorms over remote areas.

We describe a test of the GCD over tropical open-ocean waters. The test treats images from GOES-12. As “ground truth” it uses radar observations of the Tropical Rainfall Measuring Mission (TRMM). In respect to ground truth, the test resembles one other. However, the present test targets particular convective systems, resolves individual pixels, and limits the difference between test and truth data to a few minutes.

The test addresses three questions. First, how well does the nominal GCD discriminate between cumulonimbus clouds (cbs) and other clouds? Second, does the GCD beat a “benchmark” index? Finally, can severe cbs (those dangerous to aircraft) be isolated by adjusting the GCD threshold?

2. METHODOLOGY

Operating at a pixel in a bi-spectral image received from an operational geostationary satellite, the GCD algorithm calculates the difference δT between blackbody temperatures at water-vapor and thermal-window wavelengths. Let T be equivalent blackbody temperature. Subscripts 3 and 4 refer, respectively, to bands 3 and 4 of the GOES-12 imager. Band 3 peaks near 6.7 μ (in a water-vapor absorption band); band 4, near 11 μ (in the main infrared window band). Then

$$\delta T \equiv T_3 - T_4 \quad (1)$$

The GCD is defined in terms of δT.

$$\text{If } \delta T < 1^{\circ}\text{C, GCD} = 1;\quad \text{else, GCD} = 0 \quad (2)$$

The value "one" represents deep convective cloud. Repeating this calculation for all pixels in the image pair, the algorithm constructs a simple map of deep convection.

Truth data for any test of the GCD must map deep convection. For our particular test, data must also cover tropical parts of the West Atlantic Ocean and the East Pacific Ocean and they must from time to time coincide with GOES-12 image data.

TRMM orbits the earth asynchronously at an inclination of 35° and with a period of about 90 minutes. The Precipitation Radar (PR), one of five instruments on the TRMM, provides at least one reflectivity value every 250 m from the surface to an altitude of at least 15 km (Kozu et al. 2001). Since August 2001 the PR has been scanning a swath 247 km wide with a nominal horizontal resolution of 5.0 km.

Of the TRMM Science Data and Information System products, Algorithm 2A23 (TRMM PR Rain Characteristics) best suits present requirements. The 2A23 algorithm (developed by the TRMM Science Team) reduces PR column reflectivities to one of two states, rain or no-rain. When rain is present (with a high degree of confidence), the algorithm calculates its height. This parameter, Height of Storm (HS), is given in meters above mean sea level. It does not necessarily imply rain reaching the surface (J. Kwaitkowski 2004, personal communication). Inter alia, Algorithm 2A23 provides (to the nearest second) scan time and (whether raining or not) the latitude and

longitude of the center of each integrated field of view along a scan.

Each orbit of the TRMM crosses the face of the earth viewed by GOES-12. Orbital geometries and scan modes practically assure the eventual intersection of a bundle of PR scans with a bundle of GOES-12 scans. However, GOES-12 scan time and scan coverage vary from day to day. Absent a consistent relation between the TRMM orbit (hence, PR scan) and the GOES-12 scan, each day's set of space-time coincidences must be found as it occurs.

Premised on geolocation errors less than 4 km and scan-time errors less than half a minute, a script was written to check SSEC Desktop Ingest images from the previous 24 hours for space-time coincidences between GOES-12 and TRMM. If a GOES image falls within the geographical confines of this study (5°N-30°N, 45°W-105°W), the script calculates the time of the first and last scan line. If a TRMM orbit coincides in space and time with the image, a subsection of the GOES-12 data is archived and the principal investigators are alerted.

To benchmark the performance of the GCD, we use a temperature threshold applied to the band-4 image. For the threshold we chose 215 K. In Jordan's (1958) mean West Indies sounding, this temperature corresponds to a pressure of 190 hPa and a height of 12.8 km. Others also have used a threshold at or near 215 K to isolate deep convection (Mapes and Houze 1993; also see references therein). Hereafter, this is the benchmark index, or BMI.

Using the Man-computer Interactive Data Access System (McIDAS; Lazzara et al. 1999), the script was run each day through the period 1 June 2003--4 July 2003. The script sent 54 alerts. Processing the alerts involved three main steps. First, the set was screened. Second, the remaining alerts were ranked. Finally, a domain was defined for each of the top-ranking alerts.

Screening identified space-time coincidences. To qualify as a space-time coincidence, an alert had to meet three conditions: overlapping scan data of dimension ca 100 km (or more), scan-time difference of not more than six to seven minutes and (within the space-time window) at least one active cumulonimbus cloud. This screening reduced the set of alerts from 54 to 10.

Ranking ordered the space-time coincidences. It took place in two steps. First, coincidences were scored according to a set of five convective parameters. In order of importance, the parameters are distance from land, scan lag, intensity, size, and organization. Here "scan lag" means the absolute difference in scan times across the center of the convective entity of interest.

The second step was intended to promote diversity among the top-ranked coincidences. It involved classification of the space-time coincidences into one of four synoptic types (cloud cluster, easterly wave, tropical storm, and squall line). From the top of the list downward, these types were allowed to appear in any order. However, none was allowed to repeat until all types had appeared.

A domain is that latitude-longitude box which best balances four attributes: overlapping scans, small lag, central cbs, and compact size. The domain for a case was determined using information in the corresponding alert.

From the ranked list we chose the top six space-time coincidences for further processing. As "cases," they are listed in Table 1. Domains are mapped in Figure 1.

In test mode, the script also was run on several days in May 2003. One of these days—15 May—yielded a space-time coincidence as well qualified as any of the test coincidences. As case 11, it was added to the six "production" cases.

Whatever its domain, each case was treated similarly. First, "product" images were constructed from the native images. In this step, δT was redefined as follows:

$$\text{If } \delta T \le 10^{\circ}\text{C}, \ \delta T' = \delta T; \quad \text{else, } \delta T' = 0 \qquad (3)$$

All temperature-difference images were generated from the redefined (rather than original) δT. In the second step, referenced always to the domain for a case, graphs were plotted (and statistics calculated) for both product and native images. Appendix A more fully describes the processing of a case.

Table 1. Cases.

Case #	Date	GOES scan-time start	GOES scan-time coincidnce	Lag*	Comment
1	1 Jun	06:45	06:54	1	Gulf Stream/Hatteras; squall line
2	5 Jun	09:32	09:36	1	coastal Campeche Bay; cluster
4	12 Jun	19:40	19:45	2	Caribbean Sea; easterly wave
5	12 Jun	01:15	01:27	7	TX/OK; string of strong cells
7	12 Jun	22:15	22:23	0	Caribbean Sea; easterly wave
10	30 Jun	14:15	14:24	4	Gulf coast; Tropical Storm Bill
11	15 May	20:45	20:56	0	East Pacific; cloud cluster

*minutes

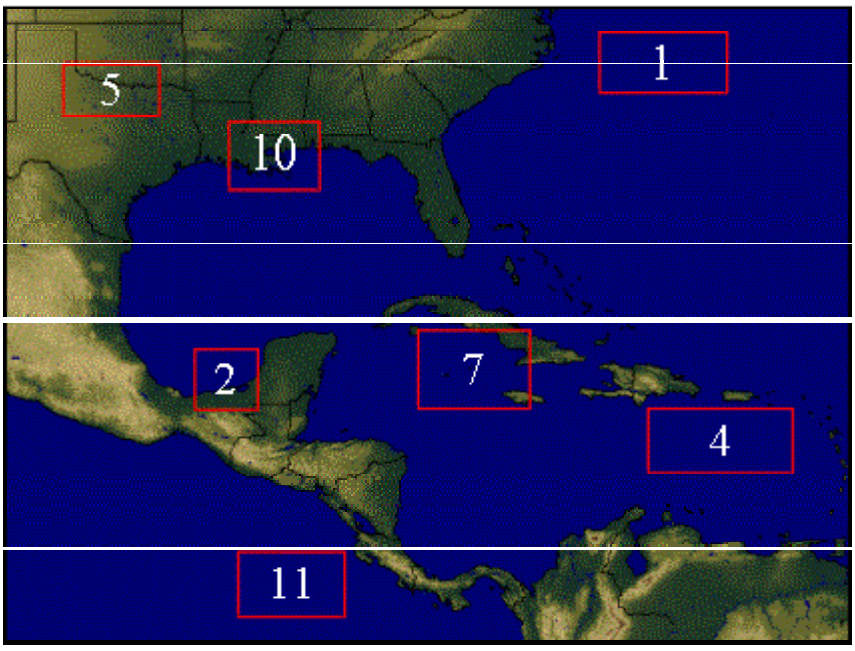

Figure 1. Domains for each of the seven cases.

3. RESULTS

Results are presented first for one case and then (in summary) for all cases. "All cases" results include a synthetic case.

3.1 *Case 11*

At 20:56 UTC on 15 May 2003 GOES-12 scanned a large cloud cluster along the Inter-tropical Convergence Zone in the East Pacific Ocean. The band-1 (visible) image (Figure 2a) suggests overshooting tops within the canopy of cirrus. Some interior structure appears in the band-4 (window infrared) image (Figure 2b), with brighter spots tending to correspond with overshooting tops inferred from the visible image. Except in lacking its contrast, the band-3 (water vapor) image (Figure 2c) closely resembles the window-infrared image.

In the temperature-difference image ($\delta T'$; Figure 2d), the cluster emerges from its background. Zonally oriented lines appear. The fine lines reflect noise in the band-3 and band-4 sensors. The coarse lines suggest linear arrangements of convective cells. One of these lines tends to coincide with the overshooting tops noted in the band-1 image (Figure 2a).

Enlarging the $\delta T'$ image highlights sensor noise. Noise notwithstanding, Figure 3 shows a large, ragged feature roughly centered on the domain. It affirms the presence of a line of convective cells near the center of the feature.

At the nominal threshold (+1°C; Figure 4a) the GCD presents a jagged object adorned with perforations and stripes. With each successive step towards the negative end of the δT range (Figures 4b through 4g), this object shrinks, fractures, and corrodes.

At the GCD threshold of –1°C (Figure 4g) three clusters of stars remain. The stars in each cluster tend to group into lines. The strongest of these lines coincides with the cell line in the $\delta T'$ image (Figure 2d; also see Figure 3); hence, with the visible image's line of overshooting tops (Figure 2a).

HS (Figure 5a) bears a fractal character. The largest of the echoes contains a number of cells, including four or five aligned zonally near the center of the domain. This echo lies almost entirely within the boundary of the $\delta T'$ feature. Its line of cells nearly coincides with the feature's line of cells

Especially near the center of the domain, the HS tower image (Figure 5b) underscores the cellular nature of the largest echo and the presence of a line of cells. In positions and spacing, the radar cells tend to match both the overshooting tops inferred from the band-1 GOES image (Figure 2a) and the cell line in the $\delta T'$ image (Figures 2d and 3).

BMI (Figure 6; shown only for the PR scan swath) suggests Rorschach blots. In position and shape (if not size), the biggest of the blots resembles the threshold-one object (Figure 5a).

The joint distribution of band temperature is shown in Figure 7. The points form a long, slightly crooked finger pointing from the warm corner of the diagram toward the cold corner. Especially at its warm end, the axis of the finger slopes a little more than one. Thus, the tip of the finger would lie along and to the right of a 1:1 (zero-intercept) diagonal.

The bivariate frequency distribution of HS and 34DF, a scaled version of $\delta T'$ ($34DF = -20\delta T'$) is shown in Figure 8. Minus 20 on the 34DF scale (y-axis) corresponds to the nominal GCD threshold (plus one). Ideally, in this plot all tall-echo points (HS exceeding 9000 to 10,000 m) would lie above the –20 line; all others, below. The cloud of points suggests a mushroom. Except for values above roughly 11,000 m, the cap of the mushroom tilts upward with HS. Tilt in the cap indicates power in the GCD to discriminate deep convection. The position of the cap relative to the –20 line suggests more power at a somewhat higher threshold—up to a 34DF value of zero.

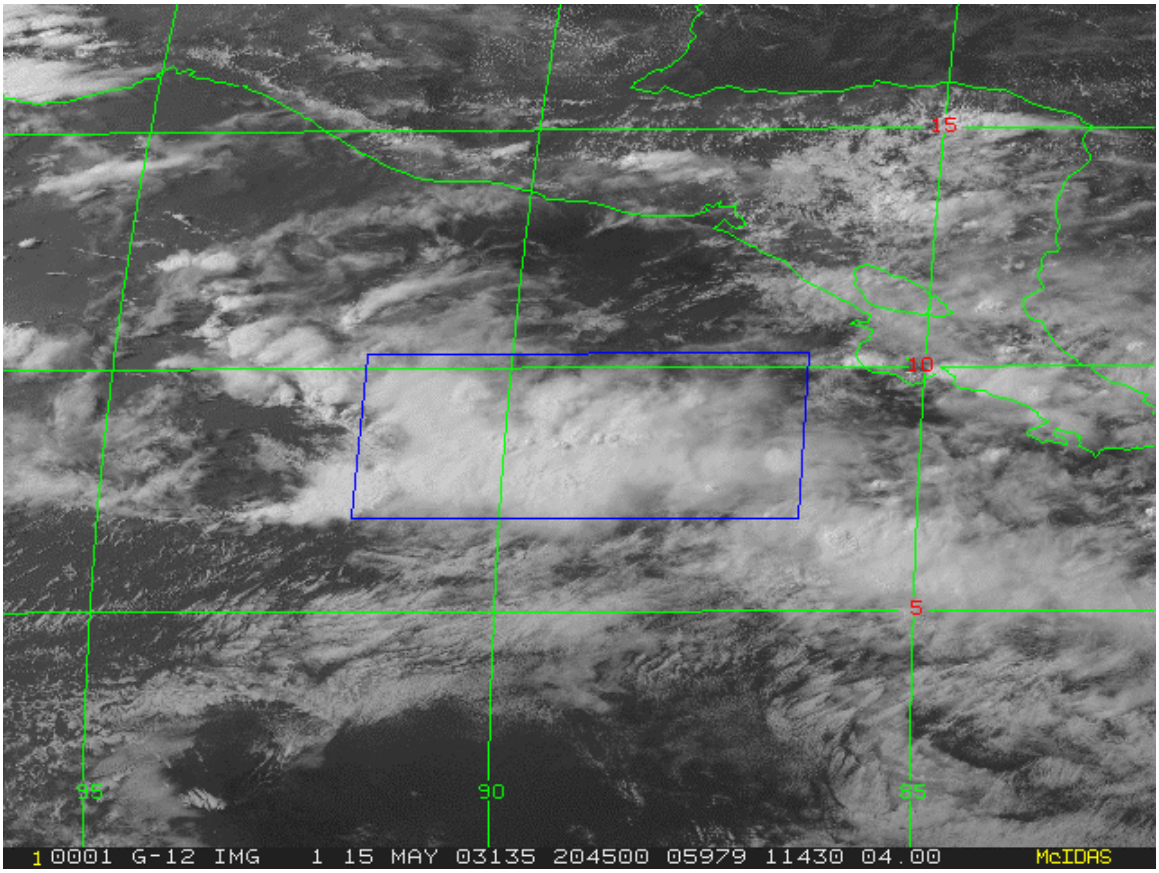

Figure 2a. Case 11 visible (band 1).

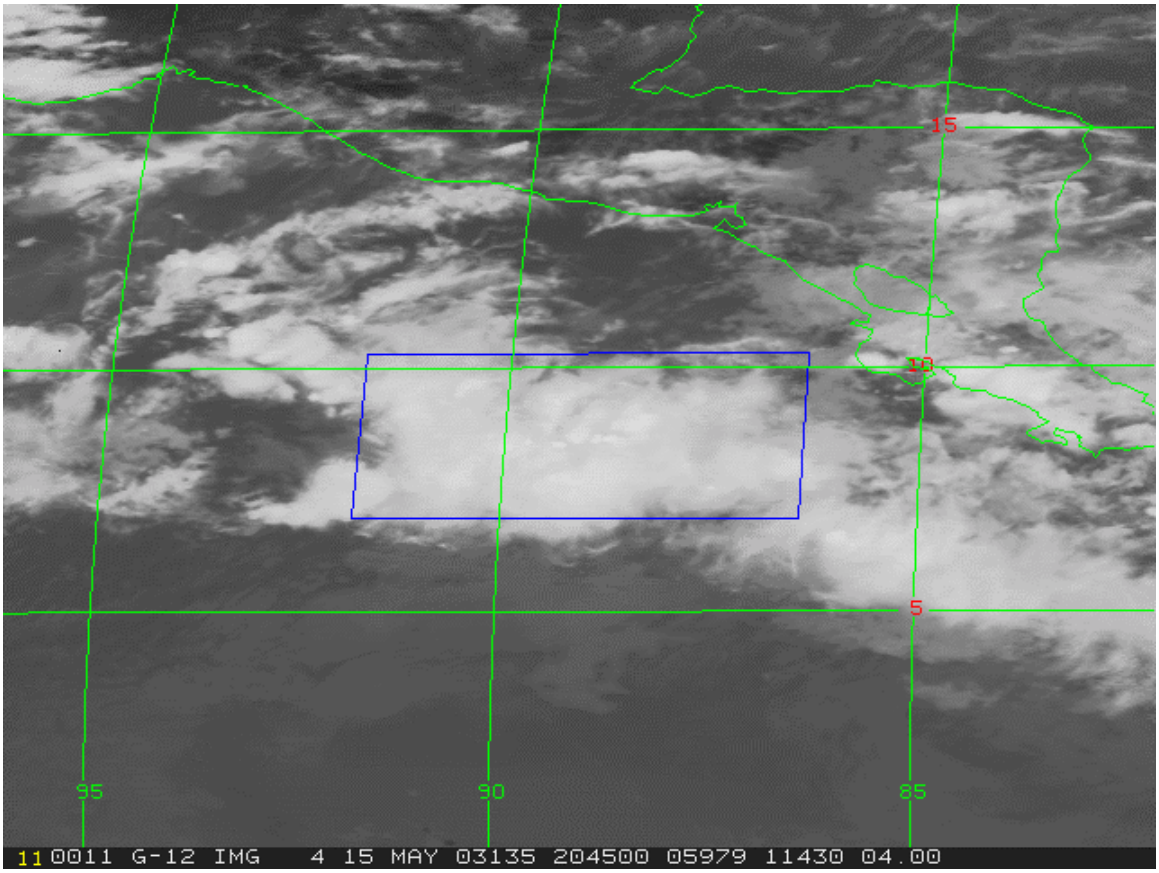

Figure 2b. Case 11 thermal infrared (band 4).

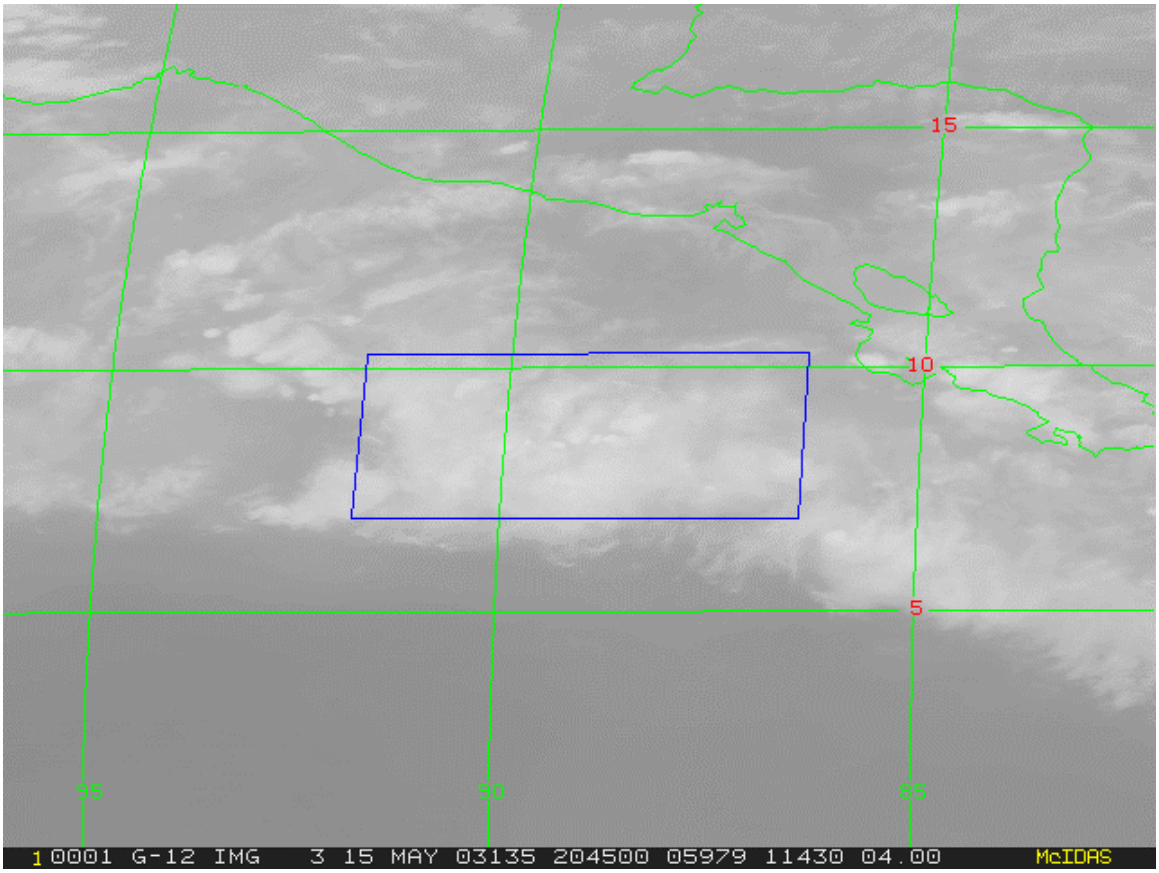

Figure 2c. Case 11 water vapor (band 3).

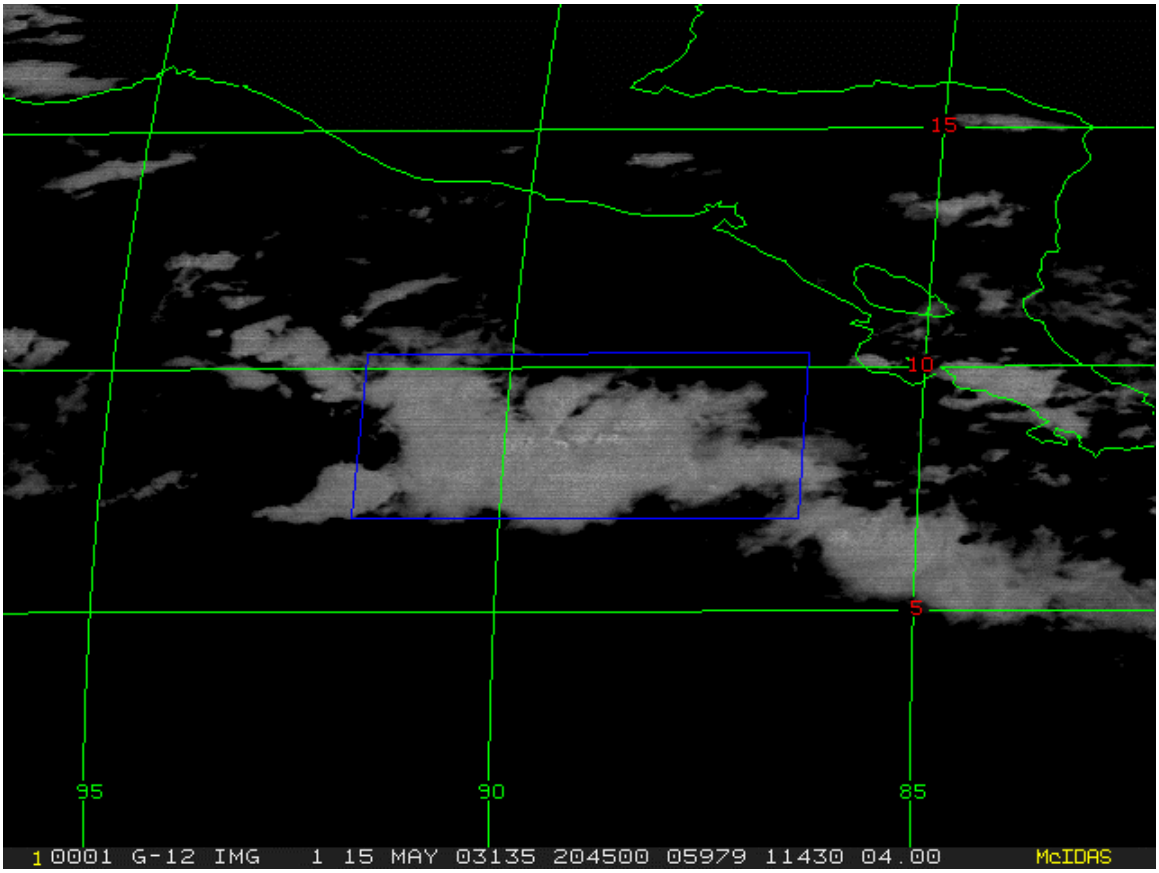

Figure 2d. Case 11 δT' (band 3 minus band 4).

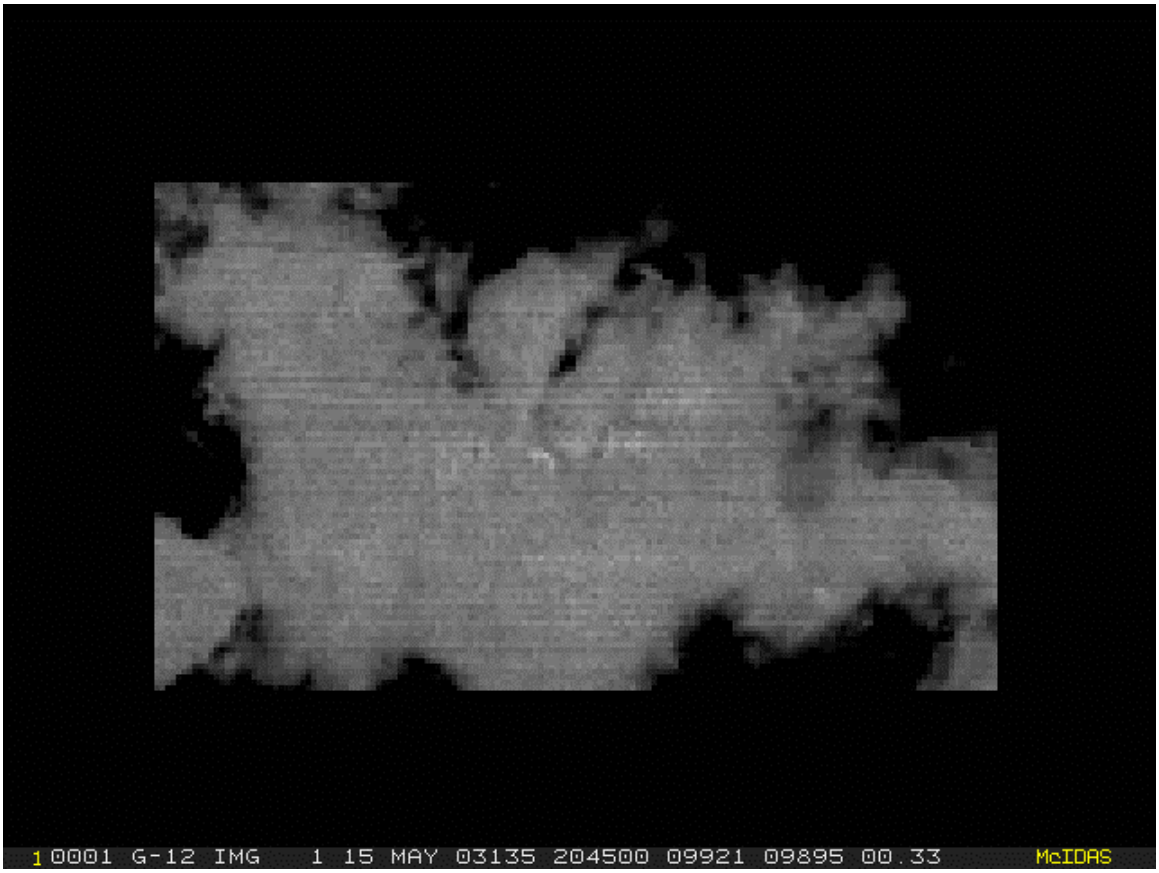

Figure 3. Case 11 domain δT' in rectilinear projection.

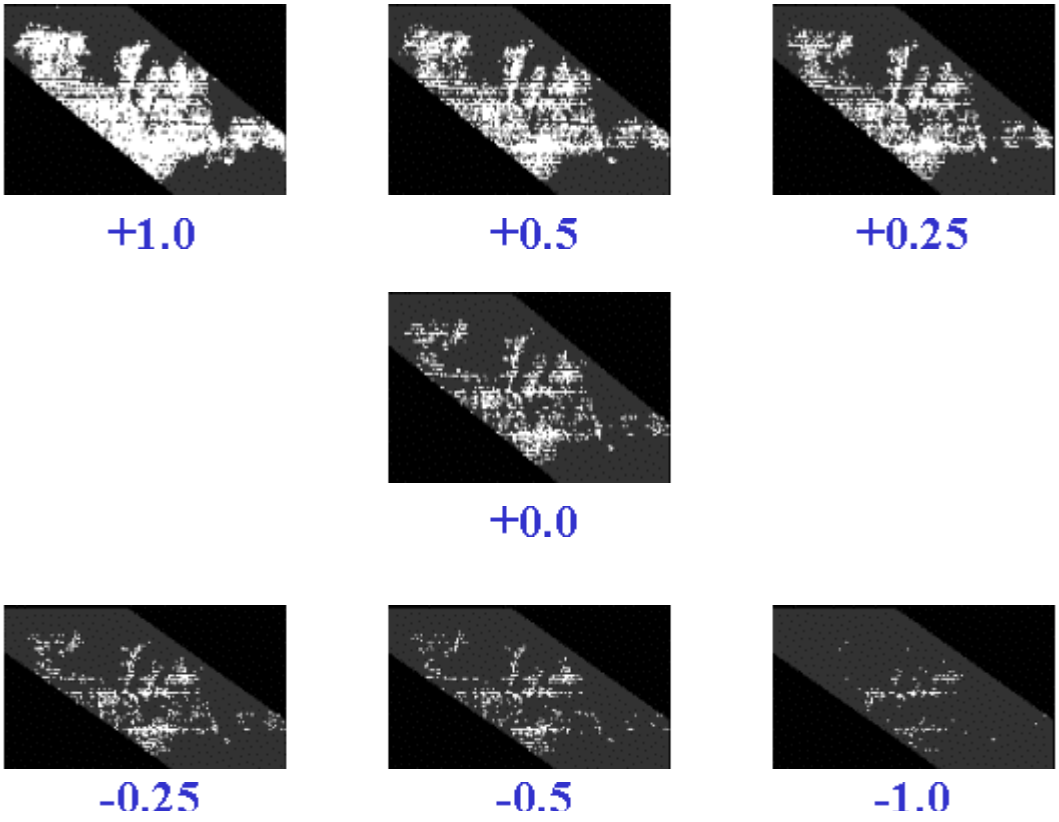

Figure 4 (a-g). Case 11 GCD image sequence for a range of thresholds, all in °C. (a) +1; (b) +0.5; (c) +0.25; (d) 0; (e) –0.25; (f) –0.5; and (g) –1.

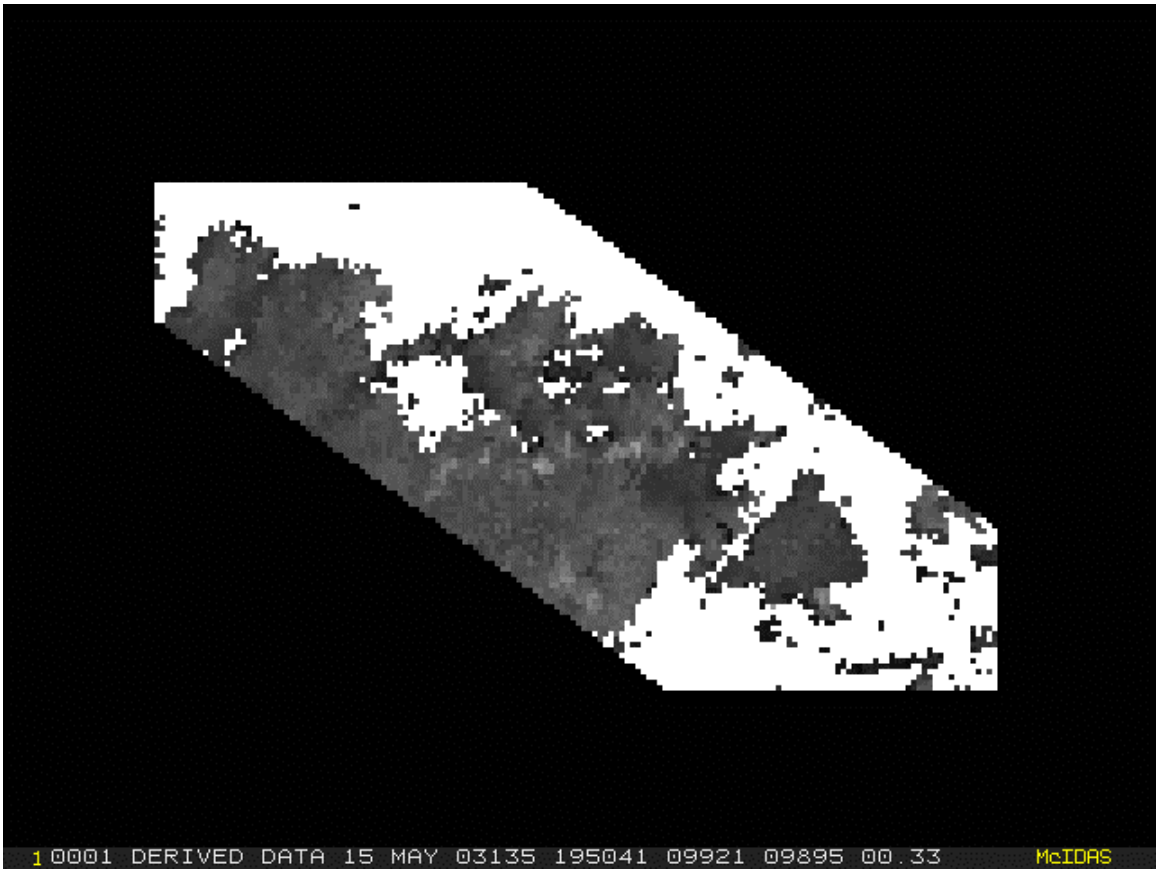

Figure 5a. Case 11 Precipitation Radar image - Height of Storm (HS). White indicates no rain. Where rain is present, the height of the rain column is scaled as zero (black) to 20,000 m (light grey).

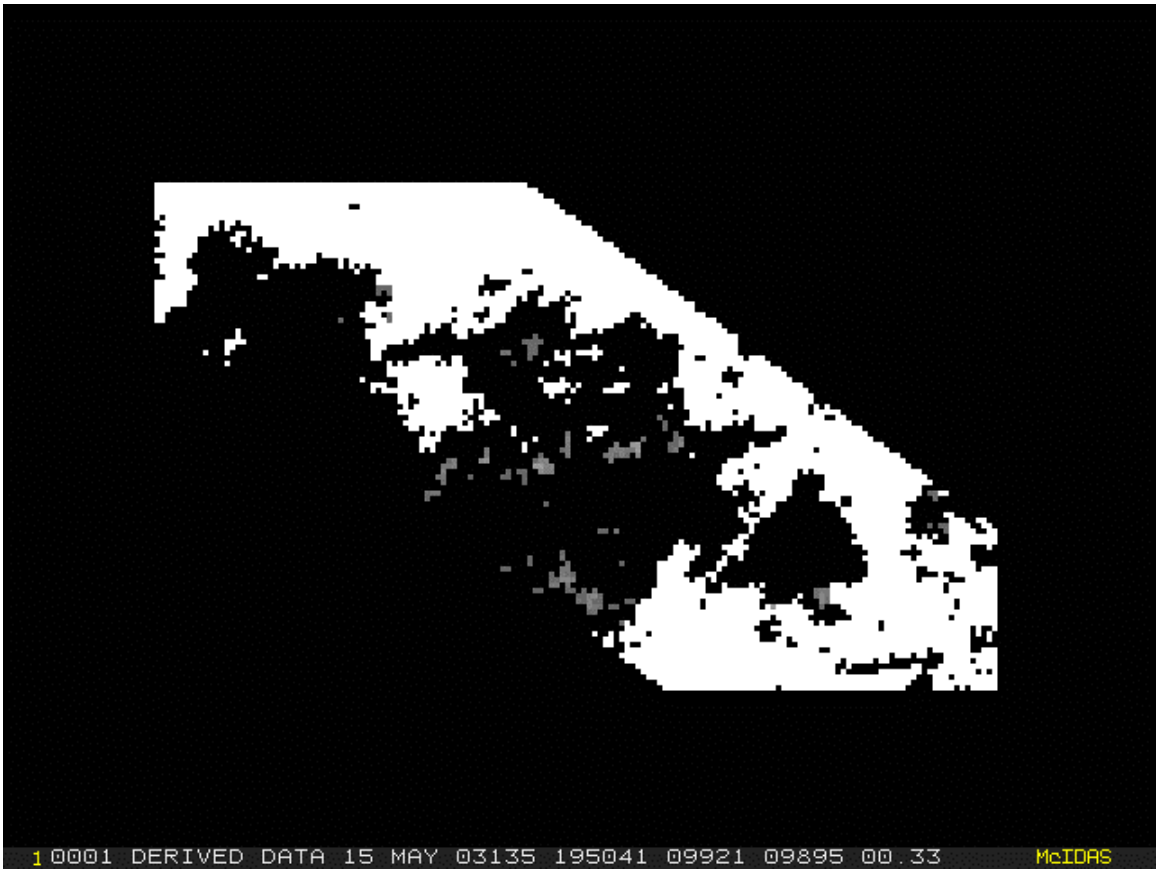

Figure 5b. Case 11 Precipitation Radar image - Tower (HS ≥ 10,000 m).

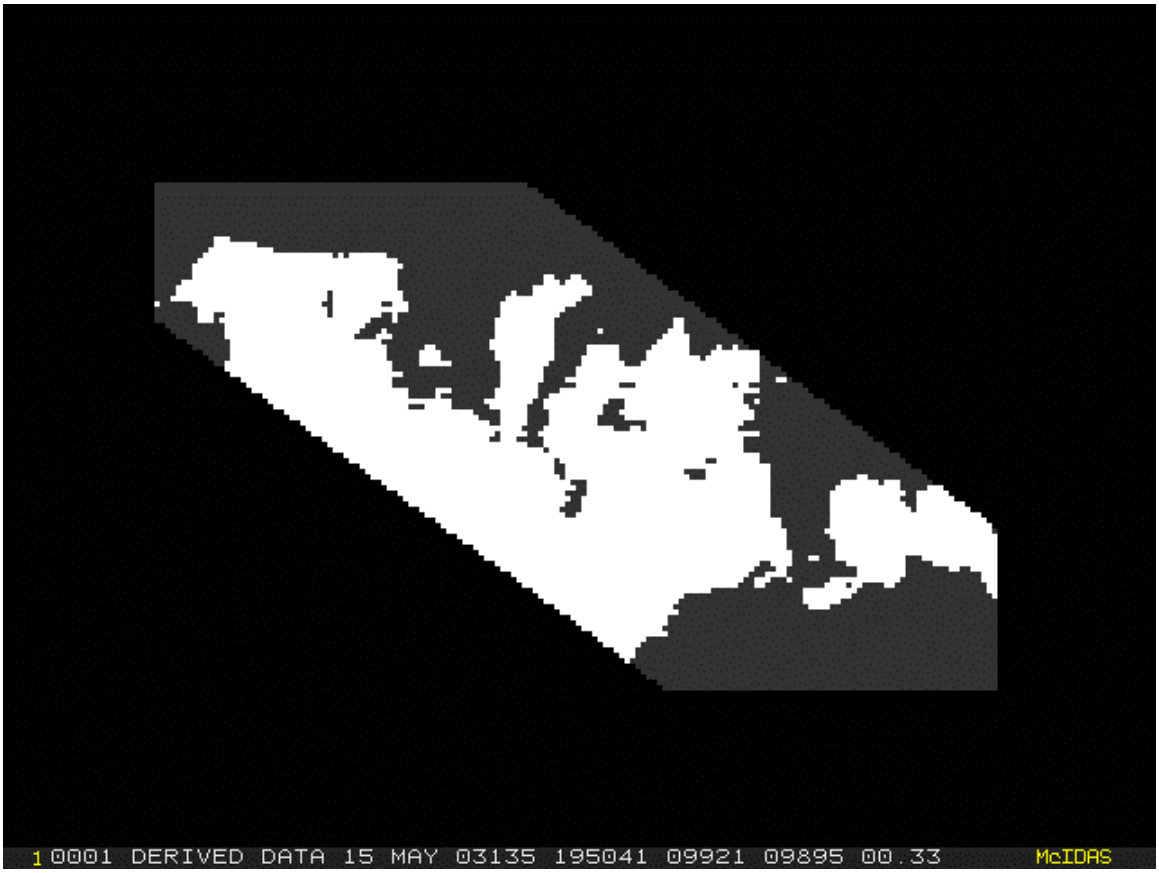

Figure 6. Case 11 BMI image.

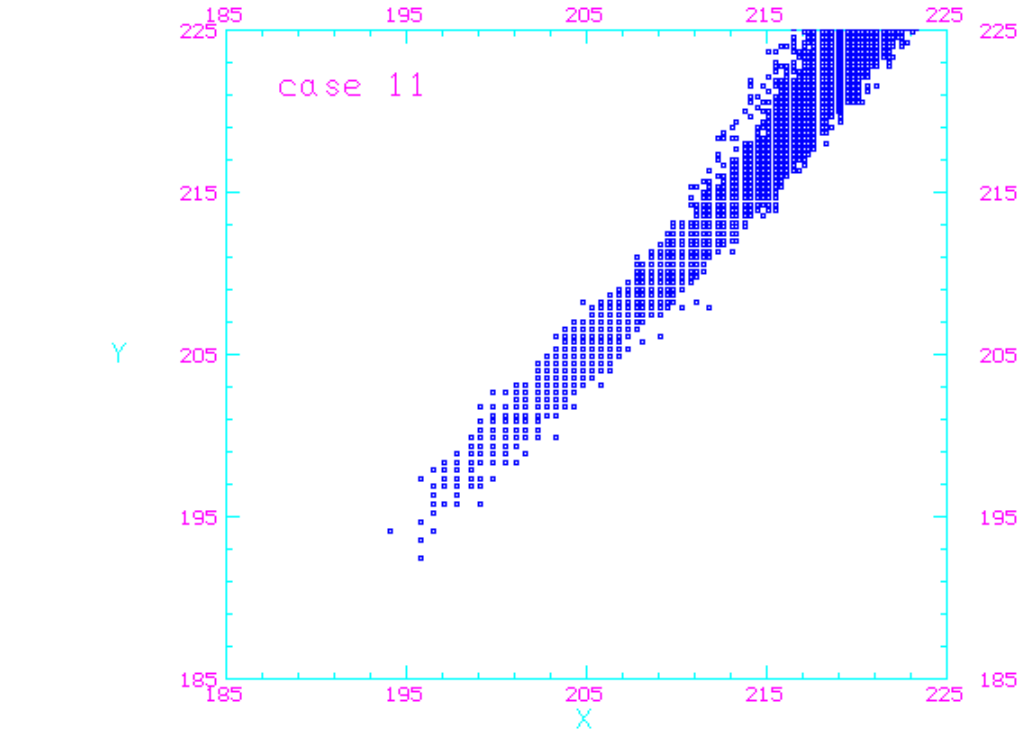

Figure 7. Case 11 bi-variate frequency distribution of temperature for water vapor band (x) and thermal infrared band (y).

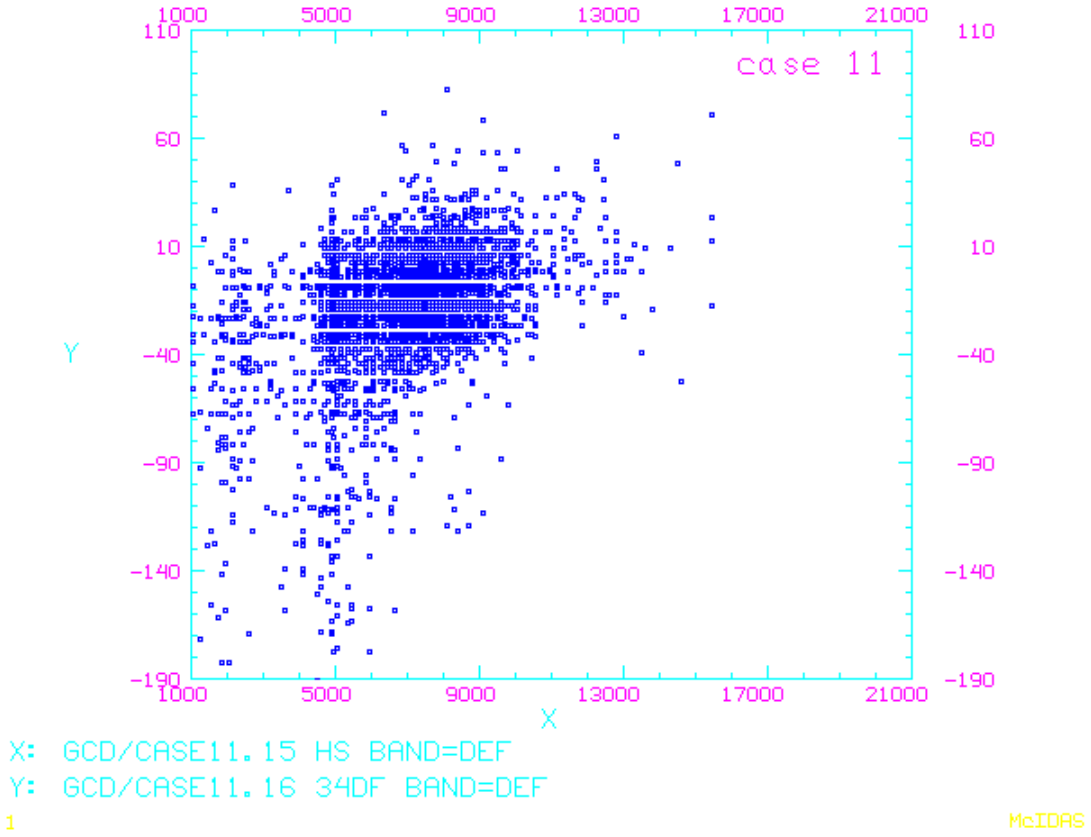

Figure 8. As for Figure 7, except HS (x) and scaled δT' (34DF; y axis).

As percents, Table 2 gives statistics for case 11. (For definitions of the terms, see Appendix B.) The GCD over-forecasts (bias > 100 and FAR ≅ 90). PODy drops drastically as the δT threshold decreases; PODn rises somewhat. Skill indices indicate discrimination power, especially at lower thresholds. Except for PODy, by every measure the GCD beats the BMI.

Table 2. Statistics for Case 11*.

	GCD							BMI
	1	0.5	0.25	0	-0.25	-0.5	-1	
Bias	2236	1531	1118	683	478	311	100	2792
PODy	90	78	67	53	40	29	15	98
PODn	59	72	80	88	91	94	98	48
FAR	95	94	93	92	91	90	84	96
CSI	4	5	5	7	7	7	8	3
Heidke	4	6	8	11	11	12	14	3

3.2 *All Cases*

For two thresholds of δT, plus one and zero, Table 3 presents statistics for all cases. At the nominal threshold (δT = 1°C) two cases (4 and 7) yielded less than 20 YY pixels; at the zero threshold, three cases (4, 7 and 10). Small samples in the YY quadrant of the two-by-two contingency table decrease confidence in the performance statistics.

This caution notwithstanding, at the nominal threshold (1°C) performance ranges widely from case to case. In all cases the false alarm ratio runs to 0.5 or higher. Except for case 2, good detection of one category of pixel means poor detection of the other.

By most measures two cases stand out. In aggregate, case 10 demonstrates no skill; case 5, skill at the 40-to-50% level. In terms of CSI and Heidke, case 11 ranks fourth from the top.

At the zero-threshold level performance ranges less widely. The false alarm ratio drops as low as 0.24. For all cases, again, good detection of one category of pixel means poor detection of the other. Cases 5 and 10 still stand out. However, the spread between the two cases shrinks.

Decreased spread results mainly from poorer performance in case 5. In this case alone, dropping the threshold diminishes performance. As noted in Table 1, case 5 was a Great Plains squall line. Large cells, each sheared to the east, were aligned roughly north/south. Overshooting tops in the visible image and enhanced Vs in the infrared image indicate young and vigorous towers within the cells. Within the domain the PR measured towers to 17 km. Towers occupied 1200 pixels. For west, central, and southeast Oklahoma, Storm Data (NOAA 2003) lists several reports of hail and damaging winds in the one-hour window centered on the case 5 coincidence.

Case 10 was a tropical storm (Bill) crossing the Louisiana coast. Two bands emanated from a single head. Shadows in the visible image and coincident cold spots indicated (at best) narrow, shallow overshooting tops. Within the domain, again the PR measured towers to 17 km. However, towers occupied fewer than 30 pixels. According to Storm Data, several parishes in southeastern Louisiana reported flash flooding in the six hours after the coincidence. None reported severe convective weather in the hour window centered on the case-5 coincidence.

If they are assumed to be independent, the seven cases can be combined. In the calculation of the combined, or synthetic, statistics a pixel of one case was assumed to carry the same weight as a pixel of any other case. Table 4 presents results for the "synthetic" case.

Extremes in the performance measures tend to fall either at high or low thresholds. For example, bias is lowest and PODn highest for small values of δT. On the other hand, PODy peaks for a δT of one. The aggregate effect (as measured by CSI and Heidke scores) favors the nominal (plus-one) threshold over others. Minus-one fares worst. At a δT of plus one skill scores run about 10%.

Viewed this way, BMI does surprisingly well. Although its bias exceeds the GCD bias at all thresholds, the BMI PODy beats the GCD PODy at all thresholds. BMI and GCD (plus one) CSI and Heidke scores are about equal.

The BMI performance for the synthetic case prompted a check of individual cases. With CSI and Heidke scores of 42% and 51%, respectively, Case 5 again stands out. To assess the contribution of this case to the synthetic case, statistics were recalculated for the set less case 5. The exclusion of the Plains squall line reduced BMI skill scores by about half. For the nominal GCD threshold, dropping case 5 also halves scores. However, by about 50% it also raises scores at the low δT end of the GCD range of thresholds. For the "short" synthetic case the GCD beats the BMI at all δTs less than 0.5°C.

Table 3. Summary of case statistics. Numbers in parentheses give sample sizes.

δT	Statistic	Case						
		1 (10666)	2 (5745)	4 (10039)	5 (8034)	7 (10188)	10 (8046)	11 (8604)
1	YY	80	211	6	776	19	24	144
	YN	3634	1635	573	675	480	2578	3435
	NN	6948	3846	9457	6166	9673	5442	5009
	NY	4	53	3	417	16	2	16
	Bias	4421	699	6433	121	1425	10007	2236
	PODy	95	79	66	65	54	92	90
	PODn	65	70	94	90	95	67	59
	FAR	97	88	98	46	96	99	95
	CSI	2	11	1	41	3	0	4
	Heidke	3	13	2	51	7	1	4
0	YY	49	142	6	172	9	14	85
	YN	1544	1044	136	55	86	981	1008
	NN	9038	4437	9894	6786	10067	7039	7436
	NY	35	122	3	1021	26	12	75
	Bias	1896	449	1577	19	271	3826	683
	PODy	58	53	66	14	25	53	53
	PODn	85	80	98	99	99	87	88
	FAR	96	88	95	24	90	98	92
	CSI	3	10	4	13	7	1	7
	Heidke	4	13	8	20	13	2	11

Table 4. Statistics for the synthetic case*.

	GCD							BMI
	1	0.5	0.25	0	-0.25	-0.5	-1	
Bias	806	584	452	301	234	163	70	867
PODy	71	51	39	27	20	16	7	77
PODn	78	84	88	92	94	96	98	77
FAR	91	91	91	91	91	90	90	91
CSI	9	8	8	7	6	6	4	9
Heidke	11	11	10	10	8	9	6	11

*Pixels number 61322.

4. CONCLUSION

We describe a test of GCD performance over the tropical and sub-tropical northwestern Atlantic and northeastern Pacific Oceans. Drawing on GOES-12 image data, the test targeted individual convective systems but operated on pixels rather than convective entities. Height of Storm (HS) measurements from the Precipitation Radar (PR) on the Tropical Rainfall Measuring Mission served as ground truth data.

For a six-week period beginning in mid-May 2003, passes of the PR were matched in time and space to images from GOES-12. In eleven of the matches the time-space window of coincidence framed all or part of a convective system. Systems ranged from cloud clusters through easterly waves to squall lines.

The subset of these systems so far analyzed yields the following results. First, as expected, the nominal GCD erred on the side of predicting too many (rather than too few) deep-convective pixels. Performance tended to improve when the threshold temperature difference (infrared minus water vapor) was set to roughly zero (rather than plus one).

Second, with the exception of one case, near a "zero" threshold the GCD quite consistently outperformed a single-threshold (215 K), infrared-only index. The exception involved a Great Plains squall line.

Finally, the question of whether severe cumulonimbus clouds could be identified more consistently at lower GCD thresholds remains open. Case 5—the Plains squall line—suggests a certain weakness in the GCD's capacity to pinpoint this class of cumulonimbus clouds.

Future work should consider the effect of raising the altitude threshold for a tower. Additionally, through the use of flags in the 2A23 data set, the PR Height of Storm towers could be classified as convective or stratiform. Finally, to resolve the ambiguity in Height of Storm treatment of echoes, either of two passive instruments on the TRMM, Lightning Imaging Sensor or the Visible and Infrared Radiometer System, could help to screen out elevated echoes. We note with regret NASA's decision to decommission TRMM.

ACKNOWLEDGEMENT

Radar data used in this study were acquired as part of the Tropical Rainfall Measuring Mission. Height of Storm and related algorithms were developed by the TRMM Science Team. TRMM is an international project jointly sponsored by the Japan National Space Development Agency (NASDA) and the US National Aeronautics and Space Administration (NASA) Office of Earth Sciences.

REFERENCES

Jordan, C. L., 1958: Mean soundings for the West Indies area. *J. Meteor.,* **58**, 91-97.

Kozu, T., T. Kawanishi, H. Kuroiwa, M. Kojima et al., 2001: Development of Precipitation Radar onboard the Tropical Rainfall Measuring Mission (TRMM) Satellite. *IEEE Trans. Geosci. Remote Sensing,* **39** (No. 1), 102-116.

Lazzara, M. A., J. M. Benson, R. J. Fox, D. J. Laitsch, J. P. Rueden, D. A. Santek, D. M. Wade, T. M. Whittaker, and J. T. Young, 1999: The Man computer Interactive Data Access System: 25 years of interactive processing. *Bull. Amer. Meteor. Soc.,* **80**, 271-284.

Mapes, B. E. and R. A. Houze, 1993: Cloud clusters and superclusters over the Oceanic Warm Pool. *Mon. Wea. Rev.*, **121**, 1398–1416.

Mahoney, J. L., B. G. Brown, J. E. Hart, and J. Henderson, 2004: A statistical evaluation of the Collaborative Convective Forecast Product (CCFP) and the Convective Significant Weather Advisories (C-SIGMET): 1 April - 31 October, 2001 and 1 April - 31 July, 2002. *NOAA Technical Memorandum*, Forecast Systems Laboratory, Boulder, CO.

Mosher, F. R., 2001: A satellite diagnostic of global convection. In *Preprints, 11th Conference on Satellite Meteorology and Oceanography (15-19 October 2001; Madison, WI)*, American Meteorological Society, Boston, MA, 416-419.

NOAA, 2003: Storm Data. National Environmental Satellite and Information Service, National Climatic Data Center, Asheville, NC, **45** (No. 6).

Appendix A. Case processing.

Figure A1 diagrams the main steps in processing a case.

TSDIS Orbit Viewer (a visualization tool from the Goddard DAAC: http://daac.gsfc.nasa.gov/) was used to strip a text-formatted array of HS data from a (PR) 2A23 file. McIDAS applications converted the text data into McIDAS area format. Using a nearest-neighbor scheme, the HS data were remapped into a rectilinear projection of resolution 4 km by 4 km. From the remapped HS data we created a "tower" map showing echoes at least 10 km tall.

GOES-12 (band-3 and band-4) images were corrected for parallax. Following the AWS GCD, the parallax correction assumed a 10-km cloud. However, the present study calculated a single correction for each case, placing the 10-km cloud at the center of the domain.

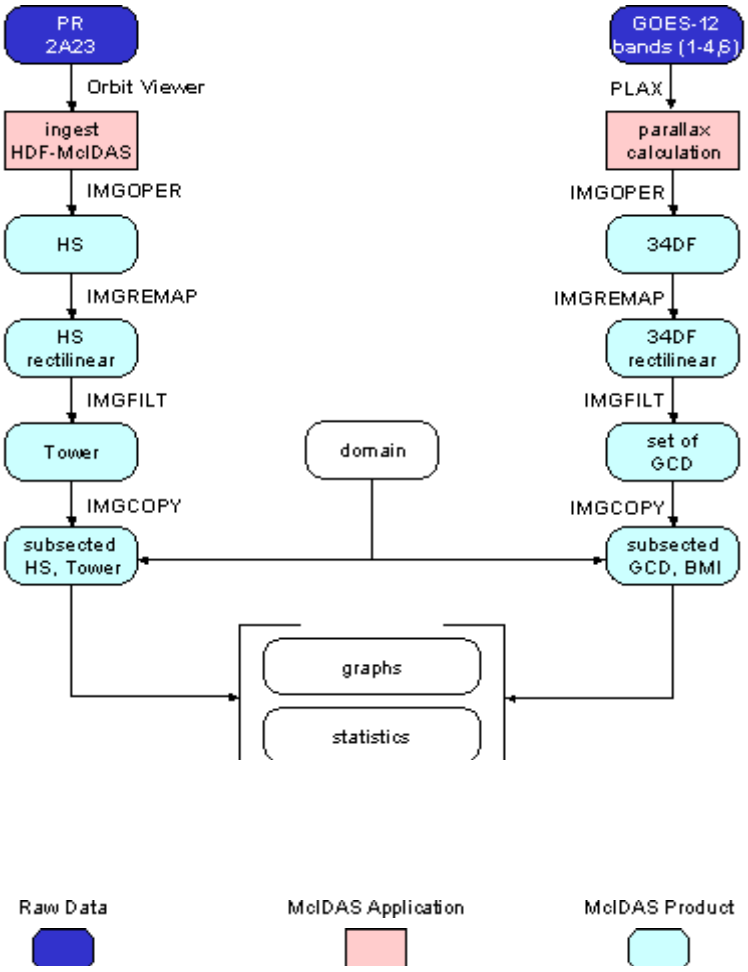

Figure A1. Schematic for the McIDAS processing of a case. Capitalized words (e.g., IMGOPER) between boxes refer to McIDAS commands. The left side deals with Precipitation Radar data; the right side, with GOES-12 data. The domain is specified outside the context of McIDAS

.

Pixel-by-pixel, corrected images were subtracted (band 3 minus band 4). Multiplied by −20 (and truncated at ±200), the resulting "34DF" image served as input to a set of GCD images.

Finally, from each of the rectilinear images (PR as well as GOES), arrays matching the domain were extracted. These "subsected" images served as inputs for most graphical and all statistical analysis.

Appendix B. Statistics.

Validation statistics are listed in Table B1. This set was drawn from standard (forecast) verification metrics (e.g., Mahoney et al. 2004). All are expressed as percents.

Table B1. Validation statistics.

Statistic	Variable	Range		Target
		low	high	
"yes" bias; also known as forecast bias	bias	0	∞	100
probability of detecting "yes" observations	PODy	0	100	100
probability of detecting "no" observations	PODn	0	100	100
false alarm rate	FAR	0	100	0
critical success index	CSI	0	100	100
Heidke skill score	Heidke	0	100	100