

PL-TR-94-2114

SUPPORT OF ENVIRONMENTAL REQUIREMENTS FOR CLOUD ANALYSIS AND ARCHIVE (SERCAA): ALGORITHM DESCRIPTIONS

**Gary B. Gustafson
Ronald G. Isaacs
Robert P. d'Entremont
Jeanne M. Sparrow
Thomas M. Hamill
Christopher Grassotti
Douglas W. Johnson**

**Charles P. Sarkisian
Daniel C. Peduzzi
Brian T. Pearson
Vida D. Jakabhazy
James S. Belfiore
Anthony S. Lisa**

**Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139**

28 March 1994

Scientific Report No. 2

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

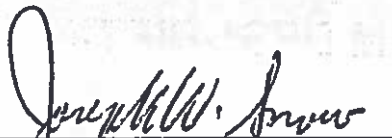
**Sponsored by the joint Department of Defense, Department
of Energy, and Environmental Protection Agency
Strategic Environmental Research and Development Program (SERDP)**



**PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010**

"This technical report has been reviewed and is approved for publication."


ALLAN J. BUSSEY
Contract Manager


JOSEPH W. SNOW
Chief, Satellite Meteorology Branch
Atmospheric Sciences Division


ROBERT A. McCLATCHEY
Director, Atmospheric Sciences Division

This report has been reviewed by the ESC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

Qualified requestors may obtain additional copies from the Defense Technical Information Center (DTIC). All others should apply to the National Technical Information Service (NTIS).

AER, INC. HAS RETAINED PATENT RIGHTS TO CERTAIN ASPECTS OF THE SERCAA ALGORITHMS UNDER FAR 52.227-11.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/TSI, 29 Randolph Road, Hanscom AFB, MA 01731-3010. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 28 March 1994	3. REPORT TYPE AND DATES COVERED Scientific Report #2	
4. TITLE AND SUBTITLE Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA): Algorithm Descriptions			5. FUNDING NUMBERS PE: 35160F PR DSPO TA GR WU AA CONTRACT F19628-92-C-0149	
6. AUTHOR(S) G.B. Gustafson, R.G. Isaacs, R.P. d'Entremont, J.M. Sparrow, T.M. Hamill, C. Grassotti, D.R. Johnson, C.P. Sarkisian, D.C. Peduzzi, B.T. Pearson, V.D. Jakabhazy, J.S. Belfiore, A.S. Lisa			8. PERFORMING ORGANIZATION REPORT NUMBERS	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Atmospheric and Environmental Research, Inc. 840 Memorial Drive Cambridge, MA 02139			10. SPONSORING / MONITORING AGENCY REPORT NUMBER PL-TR-94-2114	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Phillips Laboratory 29 Randolph Road Hanscom AFB, MA 01731-3010 Contract Manager: Allan Bussey/GPAS			11. SUPPLEMENTARY NOTES Sponsored by Joint Department of Defense, Department of Energy, and Environmental Protection Agency Strategic Environmental Research and Development Program (SERDP)	
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13 ABSTRACT (Maximum 200 words) <p>This report describes the SERCAA multiplatform, multisensor, multispectral cloud product analysis algorithms. These algorithms incorporate high-resolution sensor data from multiple military and civilian satellites, polar and geostationary, into a real-time cloud analysis model and apply multispectral cloud analysis techniques that improve the detection and specification of clouds, especially cirrus and low clouds. The SERCAA algorithms consist of a number of processes involved in integrating cloud analyses from multiple satellite platforms into a single cloud analysis product. The steps required to process the raw sensor data, collected from each of the satellite platforms, into each of the individual cloud analysis products include total cloud algorithms for DMSP, AVHRR, and geostationary platforms, cloud layer and type algorithms, and an analysis integration algorithm. The SERCAA data products for the CDFS II baseline include: total cloud cover fraction, number of cloud layers (up to four floating layers), cloud layer coverage fraction, cloud type, cloud height, and analysis confidence level.</p>				
14. SUBJECT TERMS clouds, satellite meteorology, remote sensing			15. NUMBER OF PAGES 108	
17. SECURITY CLASSIFICATION OF REPORT Unclassified			16. PRICE CODE	
18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified		19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified		20. LIMITATION OF ABSTRACT Unlimited

1. Name of the organization: [Faint text]

2. Address: [Faint text]

3. City: [Faint text]

4. State: [Faint text]

5. Zip: [Faint text]

6. Telephone: [Faint text]

7. Fax: [Faint text]

8. E-mail: [Faint text]

9. Website: [Faint text]

10. Other: [Faint text]

11. [Faint text]

12. [Faint text]

13. [Faint text]

14. [Faint text]

15. [Faint text]

16. [Faint text]

17. [Faint text]

18. [Faint text]

19. [Faint text]

TABLE OF CONTENTS

	<u>Page</u>
1. INTRODUCTION	1
2. DATA REQUIREMENTS AND SOURCES	3
2.1 Sensor Data	3
2.1.1 Visible Sensor Data	4
2.1.2 Infrared Sensor Data	4
2.1.3 Sensor Data Spatial Resolution	5
2.2 Supporting Databases	7
2.2.1 Surface Temperature	8
2.2.1.1 Predicted Clear Scene Brightness Temperature	8
2.2.2 Clear Scene Visible Channel Backgrounds	13
2.2.3 Upper Air Data	14
2.2.4 Snow and Ice Location	16
2.2.5 Geographic Data	16
2.2.6 Sun-Satellite Geometry	16
2.2.7 Earth Location	17
3. AVHRR CLOUD ANALYSIS ALGORITHM DESCRIPTION	18
3.1 Background Surface Filter Tests	21
3.1.1 Sun Glint Test	21
3.1.2 Desert Background Test	23
3.1.3 Snow/Ice Cover Background Test	25
3.2 Cloud Cover Tests	25
3.2.1 Solar Independent Cloud Tests	25
3.2.1.1 Cold Cloud Test	27
3.2.1.2 Cirrus Cloud Test	27
3.2.2 Day Condition Cloud Tests	29
3.2.2.1 Low Cloud and Fog Test	29
3.2.2.2 Precipitating Cloud Test	30
3.2.2.3 Daytime Thin Cirrus Cloud Test	30
3.2.2.4 Visible Brightness Ratio Test	31
3.2.2.5 Visible Brightness Test	32
3.2.3 Night Condition Cloud Tests	33
3.2.3.1 Fog, Low Stratus Test	33
3.2.3.2 Nighttime Thin Cirrus Cloud Test	34
3.3 Cloud Test Result Data Filter	34
3.4 Cloud Determination	36
3.5 Confidence Flag Determination	39
3.6 Output Product	41
4. DMSP CLOUD ANALYSIS ALGORITHM DESCRIPTION	43
4.1 Threshold Calculation	44
4.1.1 Infrared Channel Thresholds	44
4.1.2 Visible Channel Thresholds	46
4.2 Single Channel Test	46
4.2.1 Partial Cloud Amount Calculation	48
4.3 Bispectral Test	48
4.3.1 Partial Cloud Amount Calculation	52
4.4 Confidence Flag Determination	52
4.5 Output Product	53

TABLE OF CONTENTS (Continued)

		<u>Page</u>
5.	GEOSTATIONARY CLOUD ANALYSIS ALGORITHM DESCRIPTION	55
5.1	Temporal Difference Test	55
5.2	Dynamic Threshold Test	61
5.3	Spectral Discriminant Tests	62
	5.3.1 Solar Independent Spectral Tests	63
	5.3.2 Day Condition Spectral Tests	65
	5.3.3 Night Condition Spectral Tests	67
5.4	Confidence Flag Determination	68
5.5	Output Product	68
6.	CLOUD TYPING AND LAYERING ALGORITHM DESCRIPTION	70
6.1	Data Requirements and Inputs	71
6.2	Cloud Typing	72
	6.2.1 Height Stratification	73
	6.2.2 Top-Down Connectivity	76
6.3	Conversion from Sensor to Polar Stereographic Projection	79
6.4	Cloud Layering	80
6.5	Output Product	82
7.	ANALYSIS INTEGRATION ALGORITHM DESCRIPTION	83
7.1	Integration of Total Cloud Fraction	85
7.2	Integration of Layer Cloud Parameters	85
7.3	Cloud Type Assignment	88
7.4	Estimated Analysis Errors	89
7.5	Confidence Flag Index and Cloud Type Index	90
7.6	Optimum Interpolation	90
7.7	Cirrus and/or Low Cloud from NOAA/AVHRR and GOES/VAS...	91
8.	REFERENCES	93
	APPENDIX A Cloud Test Threshold Tables	95
	APPENDIX B Acronyms	99

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	SERCAA Multisource Cloud Analysis and Integration Procedure	1
2	Example Histogram of Comparison Between Satellite Brightness Temperature and Corresponding Surface Skin Temperature Value for 89,763 Clear Pixels Obtained from an AVHRR Scene from 1947 UTC on 6 June 1992. Vertical Lines Represent 2 Standard Deviations About the Mean.	9
3	Geostationary Clear Scene Visible Channel Background Generation Algorithm Functional Flow Diagram	15
4	Satellite-Earth-Solar Geometry	17
5	AVHRR Sun Glint Test Functional Flow Diagram	22
6	AVHRR Desert Background Test Functional Flow Diagram	24
7	AVHRR Snow/Ice Cover Background Test Functional Flow Diagram	26
8	Cloud Test Result Data Filter Examples. Each Group of Boxes Represents the Cloud Analysis Results for One Filter Window. A Black Box Signifies Cloud Has Been Detected; a White Box Means Clear.	35
9	AVHRR Cloud Classification Procedure	37
10	DMSP/OLS Cloud Analysis Algorithm Approach	43
11	IR Single Channel Test Dual Threshold Classification Approach	45
12	Single Channel Algorithm Functional Flow Diagram	47
13	Bispectral Classification Approach	49
14	Bispectral Algorithm Functional Flow Diagram	51
15	Geostationary Cloud Analysis Algorithm Functional Flow Diagram	56
16	Visible and Infrared Temporal Difference New Cloud Algorithm Conceptual Approach	57
17	Bispectral Temporal Difference Functional Flow Diagram	59
18	Dynamic Threshold Technique	61
19	Spectral Test Functional Flow Diagram	64
20	SERCAA Cloud Typing and Layering Algorithm	71

LIST OF FIGURES (Continued)

<u>Figure</u>		<u>Page</u>
21	Cloud Type Determination	73
22	Unsupervised Clustering Pixel Selection	75
23	Cloud Analysis Integration Functional Flow	86
24	Cloud Analysis Integration Example	87

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1 Current Sensor Channel Data Attributes Used for Algorithm Development	4
2 Required External Supporting Databases	7
3 Required Internal Supporting Databases	7
4 AFGWC Upper Air Database	14
5 AVHRR Sensor Channel Naming Conventions	18
6 AVHRR Cloud Tests	19
7 AVHRR Background Surface Filter Tests	20
8 Background Surface Filters for Cloud Tests	36
9 AVHRR Quanta Value Classification Assignments	40
10 AVHRR Cloud Analysis Algorithm MCF File Bit Assignments	41
11 DMSP Confidence Flag Criteria	52
12 DMSP Confidence Flag Assignment	53
13 DMSP Cloud Analysis Algorithm MCF File Bit Assignments	54
14 Geostationary Spectral Discriminants	63
15 Geostationary Cloud Analysis Algorithm MCF File Bit Assignments	68
16 Cloud Analysis Algorithm MCF File Bit Assignments	72
17 Pixel Attributes	80
18 Cloud Typing and Layering Output	82
19 Analysis Integration Processed Parameters	84
20 Grid Box Minimum Pixel Requirements	84
21 Cloud Types Processed During Analysis Integration	89
22 Estimated Analysis Errors	89
23 Acceptance Thresholds	90

1. INTRODUCTION

This document provides descriptions of a set of cloud analysis algorithms developed under Phase I of the Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA) research and development project. The project objective is to provide a global cloud analysis capability for use in determining the radiative and hydrological effects of clouds on climate and global change and in initializing operational cloud forecast models. To achieve this objective, high resolution sensor data from multiple military and civilian satellite platforms, both polar and geostationary, are integrated into a real-time cloud analysis product. Figure 1 illustrates the processing flow employed to analyze the multi-platform data to detect and classify cloud and to then integrate the separate analysis results.

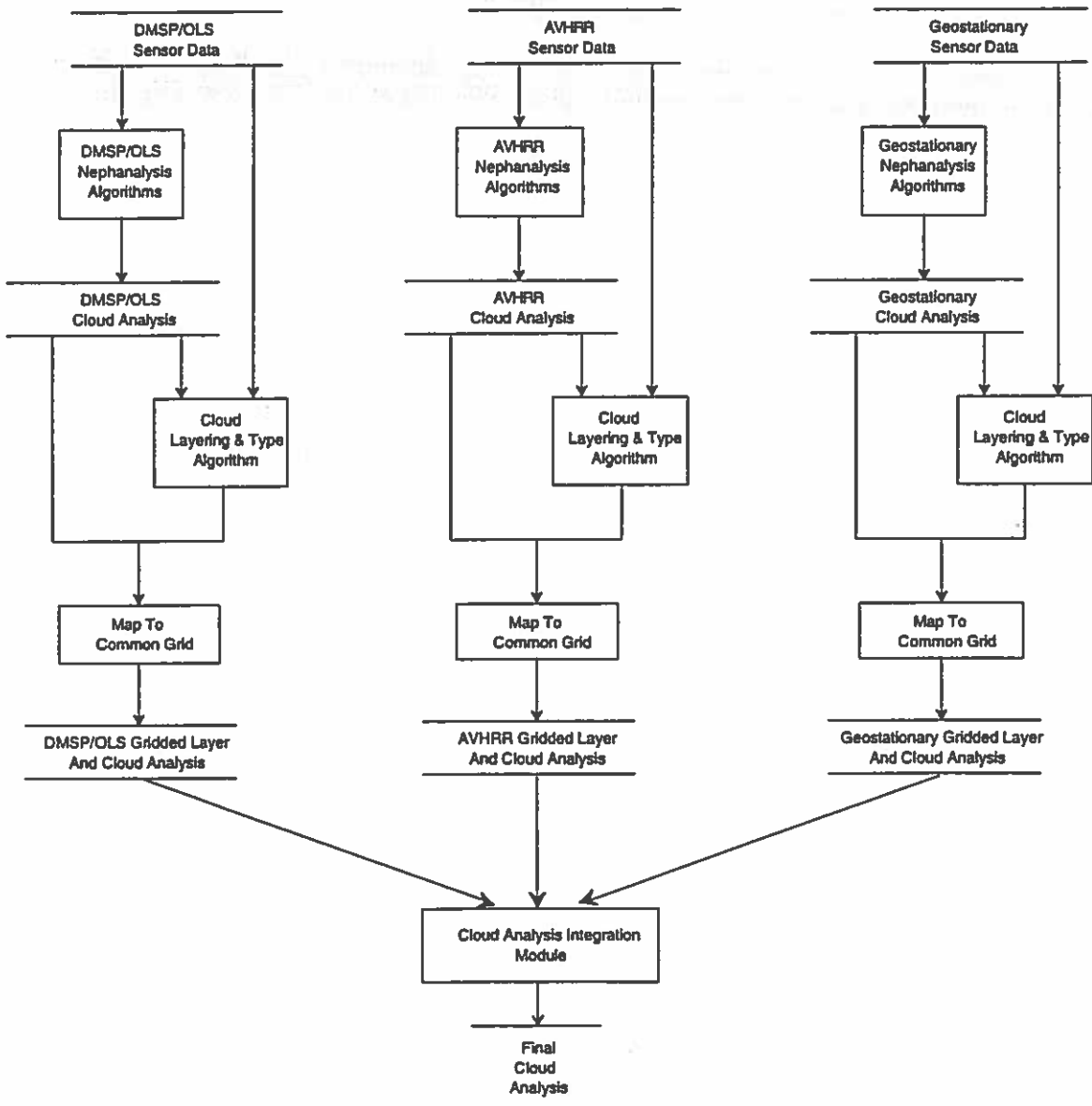


Figure 1. SERCAA Multisource Cloud Analysis and Integration Procedure

The purpose of this document is to provide descriptions of the algorithms employed in the SERCAA research and development project. Extensive scientific background does not accompany these descriptions. Rather, this document is intended to describe the procedures performed by these algorithms. The algorithm descriptions contained within this document are organized into six sections, as follows:

- Data Requirements and Sources
- AVHRR Cloud Analysis Algorithm Description
- DMSP Cloud Analysis Algorithm Description
- Geostationary Cloud Analysis Algorithm Description
- Cloud Typing and Layering Algorithm Description
- Analysis Integration Algorithm Description

Each section may be treated as a stand-alone document but when viewed as a whole an improved level of understanding will be obtained of the SERCAA project.



2. DATA REQUIREMENTS AND SOURCES

SERCAA nephanalysis algorithms are designed to operate on satellite sensor data from the DMSP/OLS, NOAA/AVHRR, GOES/VAS, METEOSAT/VISSR and GMS/VISSR. Specific satellites used during the algorithm development and testing process were: DMSP F10 and F11, NOAA 11 and 12, GOES 7, METEOSAT 3 and 4, and GMS 4. It is recognized that the constellation of operational satellites that will be in place at the time of operational implementation of these algorithms will likely change from that under which they were developed and tested. In preparing for this likelihood, it should be noted that the analysis algorithms can be adapted to accommodate changes or additions to the satellite sensor suites listed above, including those pertaining to spatial resolution of the sensor data and sensor channel band selection. Anticipated changes in imaging sensor characteristics related to the introduction of GOES I, NOAA K, DMSP 5D-3 and Feng-Yun will require modifications to the algorithms through the addition of new cloud tests and/or supporting databases but are not expected to require re-engineering of the analysis approach. The general satellite data processing philosophy guiding SERCAA algorithm design was to operate at the highest spatial resolution at which both visible and infrared data are globally available and to use the full range of spectral information available from the sensors (i.e., use all data bits from each sensor channel). Supporting data are assumed to be available at the Air Force Global Weather Central (AFGWC) from models or analysis programs external to the SERCAA algorithms except as noted in Section 2.2.

2.1 SENSOR DATA

Visible and infrared satellite imaging sensor data are required by the SERCAA cloud algorithms to provide reflectance and brightness temperature information for cloud detection and layer classification. Currently there is no provision for processing of data from collocated IR or microwave sounding instruments with the exception of GOES VAS data collected in the imaging mode. Table 1 provides a list of data sources and attributes for all sensor platforms available to SERCAA during the algorithm development and testing process. This should be considered the baseline data set required by the SERCAA algorithms. As stated above, it is anticipated that some sensor data characteristics will change prior to the operational implementation of the algorithms due to the launch of new or modified systems. SERCAA data requirements listed in Table 1 should be considered flexible both to accommodate data missing or data denied situations and to provide an upgrade path for any improvements that may occur in satellite and supporting data quality or resolution prior to or during the operational implementation.

To minimize spatial distortion, sensor data are maintained and analyzed in original scan projection just as they are received from the satellite. Data are required to be Earth-located, calibrated, and subjected to data quality checks prior to being processed by the SERCAA algorithms. Any missing or bad data are required to be flagged as such before processing, since the SERCAA algorithms perform no data quality control checks. Earth location, viewing geometry, and solar geometry information are considered supporting databases and are addressed separately in Section 2.2.

Table 1. Current Sensor Channel Data Attributes Used for Algorithm Development

Satellite	Sensor	Channel (μm)	Data Format	Resolution ¹ (km)	Bits per Pixel ²	Pixels per Scan Line
DMSP	OLS	0.40-1.10	counts	2.7	6	1464
		10.5-12.6	EBBT	2.7	8	1464
NOAA	AVHRR	0.58-0.68	percent albedo	4.0	10	409
		0.72-1.10	percent albedo	4.0	10	409
		3.55-3.93	EBBT	4.0	10	409
		10.3-11.3	EBBT	4.0	10	409
		11.5-12.5	EBBT	4.0	10	409
GOES	VAS	0.55-0.75	counts	0.86	6	15288
		3.71-4.18	EBBT	13.8	10	1911
		10.5-12.6	EBBT	6.9	10	3822 ³
		12.5-12.8	EBBT	13.8	10	1911
METEOSAT	VISSR	0.55-0.75	counts	2.5	8	5000
		10.5-12.6	EBBT	5.0	8	2500
GMS	VISSR	0.5-0.75	counts	1.25	6	10000
		10.5-12.5	EBBT	5.0	8	2500

¹Sensor resolution at satellite subpoint that will provide global coverage.

²AVHRR radiance data are transmitted at 10-bit resolution, however, the SERCAA development system could only accommodate 8-bit brightness temperature data (although the full 10-bit resolution is used in the radiance to brightness temperature transformation).

³GOES long wave infrared data are over sampled in the across-track direction by a factor of 2.

2.1.1 Visible Sensor Data

The majority of SERCAA cloud analysis algorithms process visible reflectance data in a relative or band-differencing sense. As such, visible sensor data are not required to be absolutely calibrated. The primary reason for this is that visible counts have different physical meanings for DMSP, NOAA, and geostationary satellite sensors. For OLS, visible counts are proportional to upwelling reflected solar energy (Heacock, 1985). For AVHRR, visible counts are linearly proportional to percent albedo, defined as the albedo that would be observed from a diffuse, isotropic reflector at an incident solar zenith angle of 0° (Kidwell, 1988). Geostationary satellites (GOES, METEOSAT, and GMS) all use a version of the Visible Infrared Spin Scan Radiometer (VISSR) to measure visible data. Visible data from this instrument are proportional to the square root of reflected upwelling energy (Gibson, 1984; MEP, 1989; MSC, 1989). Thus, to more precisely characterize the physical meaning of each data source the following naming conventions are adopted for this report: references to AVHRR visible data will be termed "albedo," while all other sources of visible data will be referred to as visible "counts". All SERCAA nephanalysis algorithms described in this document expect visible data to conform to these conventions.

2.1.2 Infrared Sensor Data

In contrast to visible sensor data, infrared radiance measurements from all platforms are required to be absolutely calibrated and subsequently converted to equivalent blackbody brightness temperature (EBBT). Throughout this document the term "brightness temperature" is treated as synonymous with EBBT. The calibration operation is performed differently for the individual sensors but generally requires a linear calibration function to convert IR counts to radiance. Conversion to EBBT is then performed by first making an assumption of blackbody emission from the radiating

surface and then inverting the Planck function over the bandpass-weighted spectral range of each infrared sensor channel. The exception to this convention is the DMSP OLS sensor which performs calibration and data conversion on-board the satellite and then transmits an 8-bit IR count that is directly proportional to brightness temperature. Note that in instances where the blackbody assumption is not correct (i.e., surface emissivity less than 1.0) then this convention will produce brightness temperatures which can be significantly different than the physical temperature of the surface. This phenomena is recognized and addressed separately by the individual analysis algorithms (see Sections 3 through 5), it is often useful in discriminating different types of cloud and background surfaces. It should be noted that other factors, in addition to surface emissivity, can also cause satellite derived brightness temperatures to differ from actual temperature of the radiating surface (e.g., atmospheric attenuation). These conditions are addressed in Section 2.2.1.1.

Calibration procedures for AVHRR channel 4 and 5 data include, in addition to the linear calibration function, a correction term to account for a slight non-linearity in their calibrations (Planet, 1988). IR channel calibration for GOES, METEOSAT, and GMS use a straight linear relationship and is performed according to the procedures in Gibson (1984), MEP (1989), and MSC (1989) respectively for each satellite. During SERCAA, for all sensors except OLS, IR counts were converted to radiance using the full 8-bit or, in the case of AVHRR, 10-bit data resolution (see "Bits per Pixel" in Table 1). For convenience, derived infrared brightness temperature data were maintained in the SERCAA database as 8-bit quantities with a resolution of 0.5 K over the range of 200.0 to 327.5 K.

The algorithms described in this document expect infrared sensor data to be available in the form of brightness temperatures. While the procedures described above produced satisfactory results during the limited real-data testing performed during the algorithm development process, they should be treated as guidelines only. Operational methods for calibrating, storing and accessing IR brightness temperature data should be considered an implementation issue with the goal of maximizing IR data quality and resolution. Recall that data calibration and data quality checking are to be performed prior to execution of the cloud algorithms.

2.1.3 Sensor Data Spatial Resolution

As stated above, SERCAA algorithms operate on satellite sensor data at the highest available spatial resolution that provides global coverage (see Table 1). Note that for polar satellites the visible and infrared channel IFOVs are the same size while for geostationary satellites the visible channel resolution is some discrete factor higher than the IR. However all cloud analysis algorithms, including geostationary, require that the visible and infrared channel data be processed at the same resolution. To accommodate this requirement for geostationary satellites, visible data are subsampled to match the IR resolution. Since all geostationary satellites used by SERCAA employ a version of the VISSR instrument, the sampling process follows the same general procedure for each. The process is most complex for GOES which uses the more advanced VISSR Atmospheric Sounder (VAS) sensor. To illustrate the sampling process a detailed description is provided here for that satellite. The process is generalized for the other platforms.

VAS instrument components include eight visible channel detectors linearity aligned in the north-south direction that are sampled simultaneously and digitized as 6-bit words to provide imagery with a nominal resolution of 0.86 km at nadir. Six thermal

detectors of two different sizes, used in pairs, sense infrared radiation in 12 spectral channels. The IR detector pairs are offset north-south in the optical plane, one mirror step for a single small detector pair and two mirror steps for two large detector pairs. Spectral selection is achieved through selection of either the small detector pair or the large detector pairs in combination with one of twelve filters placed in the optical path. The current GOES 7 transmission schedule provides full visible and thermal IR coverage of the northern hemisphere every half hour, plus an additional IR channel that alternates between 3.9 and 12.6 μm on the half hour and hour respectively (small detectors are used for thermal IR channel, large detectors for the additional IR channels). Also a 6.7 μm image overrides the 3.9 μm transmission at 0030, 0630, 1230, and 1830 UTC. Note that for the southern hemisphere full coverage is obtained every 3 hours; only partial coverage is available half hourly due to conflicts with sounding operations.

The ratio of the visible sampling rate to IR sampling rate for GOES 7 is 4 to 1, resulting in a raw visible image 14568 lines by 15288 samples and a raw thermal IR image composed of 1821 lines and 3822 samples. The visible detectors have a linear dimension of 83.6 μrad and the small IR detector has a linear dimension of 192 μrad , thus consecutive IR samples overlap by about 56%. Similarly, the linear dimension of the large IR detector is 384 μrad resulting in consecutive large IR sample overlap by about 78%.

The chosen spatial resolution for SERCAA GOES imagery is 3.45 km at nadir. To achieve 3.45 km spacing for the visible channel, data are sampled every fourth element along a scan line and every fourth scan line. Generation of 3.45 km image resolution for IR data depends on detector geometry as discussed above. For small detector configurations, 3.45 km data are digitally produced by selecting one for one all over sampled elements along a scan line followed by a one time replication of the line. Large detector configurations are built similarly; all over sampled elements are selected along a scan line and replicated once followed by replication of the line a total of three times.

Note that with the launch of GOES I (expected to be designated GOES 8 when operational) in April 1994, the issue of obtaining co-registered data from all sensor channels will become simplified. Visible sensor resolution will be 1 km square at nadir and IR resolution will be 4 km square at nadir for 3.9, 11, and 12 μm channels (Koenig, 1989). Thus subsampling of visible data to one in four will achieve an image resolution of 4 km for all four channels. Changes in the GOES 8 transmission schedule also have important implications for SERCAA since, in addition to visible data, all three IR channels (plus a 6.7 μm image at 8 km resolution) will be available with each transmission.

For METEOSAT, only two visible sensors and a single IR sensor are used. The visible sensors have a nadir resolution of 2.5 km and are offset in the N-S direction such that each produces an image of 2500 lines by 5000 samples with each line providing non-overlap coverage between successive lines from the other sensor. The IR sensor resolution is 5 km square at nadir and produces an image of 2500 lines by 2500 samples. Visible and IR data are co-registered by sampling every other pixel from only one of the visible detectors.

The GMS VISSR consists of two redundant sets of four visible detectors and one IR detector. Similar to GOES and METEOSAT the visible detectors are offset in the N-S direction, each providing 2500 non-overlapping lines by 10000 samples at 1.25 km nadir resolution. IR sensor resolution is 5 km and image size is 2500 lines by 2500 samples. Image co-registration is performed by subsampling visible data at four to one.

2.2 SUPPORTING DATABASES

In addition to satellite sensor data, supporting data are required by the SERCAA cloud algorithms to provide information on the terrestrial background (e.g., clear-scene reflectance and brightness temperature) and atmosphere plus positional and Earth location reference points. Supporting data come from two sources: databases created and maintained external to SERCAA and data that are generated as by-products of either the SERCAA algorithms or the satellite data ingest function. Table 2 provides a listing of required external supporting data and the spatial resolution at which they were available during SERCAA. Table 3 summarizes the number and type of the other required support databases. Currently, the external databases are either generated and maintained at AFGWC or are expected to be available there in the near future. All AFGWC databases are maintained as regular gridded fields superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole mesh grid spacing of exactly 381 km at 60° latitude. Nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid (e.g., 1/8th mesh has 8 x 8 cells per whole mesh box, 1/16th mesh has 16 x 16, etc.). Complete information on the AFGWC polar grid system is provided by Hoke et al. (1981). Descriptions of each database are provided in the following sections.

Table 2. Required External Supporting Databases

Data Type	Resolution (km)	Grid Mesh ¹
Surface Temperature	47	8
Upper Air Data	381	1
Snow and Ice Location	47	8
Geographic Type	6	64
Terrain Height	24	16

¹All external databases are maintained in the AFGWC standard polar stereographic grid projection based on a whole mesh grid spacing of 381 km at 60° latitude. Grid mesh designation is 1 - whole mesh, 8 - 1/8 mesh, etc.

Table 3. Required Internal Supporting Databases

Data Type	Refresh Interval ¹ / Number of Data Sets per Satellite ²				
	AVHRR	DMSP	GOES	METEOSAT	GMS
Visible Background Count	C/2	C/2	B/24	B/24	B/24
IR-Skin Temperature Statistics	T/60	T/60	NA	NA	NA
Earth Location	S/1	S/1	S/1	S/1	S/1
Sun-Satellite Geometry	S/1	S/1	S/1	S/1	S/1
Refresh Interval	Number of Data Sets per Satellite				
C - continuous	2 - ascending / descending orbit				
B - biweekly rotating	60 - ascending / descending orbit, land-water-desert background, 10 days				
S - single or bit / scan	24 - (maximum number) each time visible data are ingested through a day				
T - 10 days rotating	1 - per satellite data set				

¹Refresh Interval indicates frequency of update or period of record for the specified database.

²Number of Data Sets describes the number of separate data sets required for each satellite.

2.2.1 Surface Temperature

Global surface skin temperature data obtained from the AFGWC Surface Temperature Model (SFCTMP) are required by the SERCAA algorithms to help characterize clear-scene brightness temperatures. SFCTMP operationally produces eighth-mesh databases of analyzed shelter and skin temperature plus 3-hour and 4.5-hour forecasts. For ocean surfaces both skin and shelter temperatures are set equal to a single water temperature value. Global analyses of water temperature are obtained from the Navy every 12 hours through the Shared Processing Network. Over land, a new surface temperature analysis is performed every three hours. Conventional shelter temperature observations are blended with a first guess composed of the previous 3-hour forecast, HIRAS global spectral model surface temperature products, OLS infrared brightness temperatures at points determined to be clear by the RTNEPH, and surface temperatures derived from SSM/I measurements. The skin temperature analysis is calculated by modifying the 3-hour skin temperature forecast upward or downward by the same amount the 3-hour forecast of shelter temperature differs from the new shelter temperature analysis. Forecasts of skin and shelter temperature are made using a simplified version of the Oregon State University planetary boundary layer model. A detailed description of the AFGWC SFCTMP model is provided by Kopp et al. (1994).

2.2.1.1 Predicted Clear Scene Brightness Temperature

While all satellite data sources used in SERCAA have different channel and observing characteristics, they all have in common at least one long wave thermal infrared channel. Accordingly, all SERCAA cloud algorithms include a single channel infrared threshold technique and as such require estimates of clear-column satellite brightness temperatures to discriminate cloud-free from cloud-contaminated radiative signatures. The Geostationary Cloud Analysis algorithm uses SFCTMP skin temperatures directly to estimate clear column temperatures, however, the OLS and AVHRR polar orbiting satellite algorithms require predicted clear scene satellite brightness temperatures computed from a dynamic correction to AFGWC skin temperature values. The correction is used to account for the combined effect of multiple error sources that can occur when using the SFCTMP model to predict a satellite derived clear scene brightness temperature. Of particular concern are modeled skin temperatures that are not representative of bandpass-weighted satellite infrared brightness temperatures, differing spatial resolutions between the modeled and satellite derived data, satellite sensor calibration errors, and the effect of IR atmospheric attenuation. Accurate modeling of individual errors, let alone their combined effect, is problematic even with the resources available at a center like AFGWC. The selected SERCAA approach uses a single correction factor to account for all error sources collectively.

The procedure to predict the brightness temperature that would be observed by a satellite from the cloud-free terrestrial background at a given location and time is the same for both polar orbiter nephanalysis algorithms. While the actual calculation of the predicted temperature is performed as part of the respective satellite cloud analysis algorithms described in Sections 3 and 4, a description of the process is provided here because: 1) the process is identical for both algorithms, 2) it requires use of the AFGWC skin temperature database (Table 2), and 3) it is closely tied to the generation, maintenance, and use of the IR-Skin Temperature Statistics internal database (Table 3) also described in this section.

Before the cloud algorithms can calculate a predicted clear-scene temperature it is necessary to first compile a ten-day record of the deviation of the AFGWC SFCTMP

modeled temperature (T_{skin}) from the satellite derived temperature (T_{sat}) for locations previously classified as cloud-free. This record is used to characterize the natural variability of the difference between the two temperature values when no cloud is present so that future measurements can be tested to see if they fall within the expected range for clear conditions.

Temperature difference information is maintained as an internal database of statistics summarizing the distribution of the temperature differences stratified by location, satellite, time of day, and surface type: land, water, or desert (IR-Skin Temperature Statistics in Table 3). Statistics are accumulated over a large area recommended to be no smaller than that defined by a polar grid with a resolution of two whole mesh grid boxes (i.e., 32×32 per hemisphere). Thus, for each area a separate database entry is required for each of the ten days, for each polar satellite, its ascending and descending orbits, and the three possible surface types identified in the geographic supporting database (see Section 2.2.5).

Daily statistics for clear regions are developed as a by-product of the cloud analysis algorithms. As polar satellite data are received and analyzed through the appropriate nephanalysis algorithm, two passes through the algorithm are performed: 1) cloud detection and 2) cloud clearing. In cloud detection mode, algorithm thresholds are set to provide an optimal analysis with no preference toward over or under analysis of cloud. In cloud clearing mode cloud thresholds are set with a bias toward over analysis to insure identification of all cloud-contaminated pixels. Once clouds have been identified, they are removed from further processing and the difference between the sensor channel brightness temperature and the corresponding surface skin temperature is calculated for all remaining pixels and used to generate a frequency distribution. Figure 2 illustrates a sample temperature difference distribution developed from 89,763 cloud-free pixels observed during a NOAA 11 afternoon ascending pass over land surfaces in the east central United States.

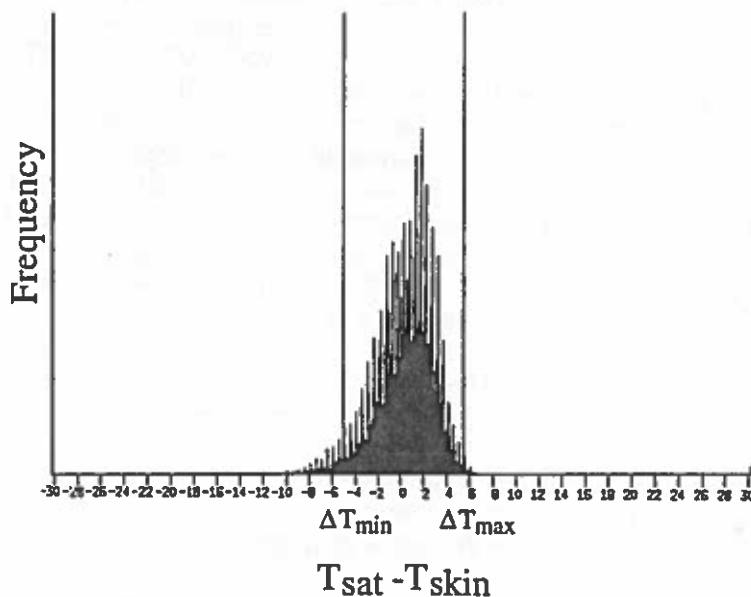


Figure 2. Example Histogram of Comparison Between Satellite Brightness Temperature and Corresponding Surface Skin Temperature Value for 89,763 Clear Pixels Obtained from an AVHRR Scene from 1947 UTC on 6 June 1992. Vertical Lines Represent 2 Standard Deviations About the Mean.

If $n = 1, 2, \dots, N_i$ is the number of clear pixels from the cloud clearing analysis for some day $i = 1, 2, \dots, 10$ then for each pixel n , the temperature difference (ΔT_n) is defined as:

$$\Delta T_n = T_{\text{sat}_n} - T_{\text{skin}_n} \quad (1)$$

where T_{sat_n} is the satellite brightness temperature at pixel n and T_{skin_n} is a time interpolated skin temperature value, derived from the Surface Temperature database, that corresponds to the time and location of pixel n . T_{skin_n} is defined by first locating the 1/8th mesh grid box closest to the latitude and longitude of pixel n (note the AFGWC polar grid convention uses the upper left corner to define the location of a grid box, see Hoke et al., 1981). The two AFGWC surface temperature database entries with valid times that bracket the time of the satellite observation are located and the respective skin temperature values at the specified 1/8th grid point are linearly interpolated to the time of the satellite observation:

$$T_{\text{skin}_n} = \frac{T_{\text{skin}_{t_2}} - T_{\text{skin}_{t_1}}}{t_2 - t_1} (t_{\text{sat}} - t_1) + T_{\text{skin}_{t_1}} \quad (2)$$

where t_{sat} is the satellite observation time, t_1 and t_2 are the valid times of the bracketing surface temperature database entries (i.e., $t_1 \leq t_{\text{sat}} \leq t_2$), and $T_{\text{skin}_{t_1}}$ and $T_{\text{skin}_{t_2}}$ are the respective skin temperature values for the specified grid box valid at times t_1 and t_2 .

During SERCAA testing a problem with the calculation of T_{skin} was discovered along coastlines (i.e., land/water boundaries). Due to the large difference in spatial resolution between the satellite and Surface Temperature data (see Tables 1 and 2), the T_{skin} calculated by taking the nearest 1/8th mesh grid point to a given pixel was often representative of a geographic type other than that corresponding to the pixel (e.g., if the pixel was located over water the skin temperature would be representative of a nearby land temperature or vice versa). When there is large thermal contrast between the adjacent land and water points, the incorrect skin temperature values can result in a false cloud signal in the cloud analysis algorithms. The AFGWC SFCTMP model uses a separate 1/8th mesh geographic database to identify land and water grid points and then assigns a representative temperature to the entire grid box based on that geographic type classification. To correct this problem a technique was developed that exploits the high resolution, 1/64th mesh, Geographic Type database developed for SERCAA (refer to Section 2.2.5). The 1/64th mesh data generally have sufficient resolution to accurately delineate land and water pixels in the satellite imagery. By comparing the corresponding geographic types from the two geographic databases it is possible to establish whether the 1/8th mesh temperatures used to calculate T_{skin} is representative of the same geographic type (i.e., land or water) as the satellite pixel. If they are the same then T_{skin} is calculated as described above. If they are different, a search is performed on the 1/8th mesh geographic database over the 3 x 3 array of 1/8th mesh grid points surrounding the grid point closest to the satellite pixel location. If a geographic type match with the satellite pixel is found within the 3 x 3 array then the skin temperatures associated with the matching 1/8th mesh grid point are used to define $T_{\text{skin}_{t_1}}$ and $T_{\text{skin}_{t_2}}$ in Eq. 2. If no match is found, then the skin temperatures for each grid point in the 3 x 3 array are linearly interpolated to the time of satellite observation using Eq. 2, and the lowest interpolated value within the array is selected as T_{skin_n} . This lowest temperature is chosen to minimize the risk of misclassifying clear pixels as cloudy in the cloud analysis algorithm.

The general Gaussian shape of the distribution in Fig. 2 is typical, therefore it was decided that the range of ΔT_i values found for each day i could be represented using the

limits defined by two standard deviations taken about the mean of the temperature difference distribution (labeled ΔT_{\min} and ΔT_{\max} in Fig. 2). Thus, the IR-Skin Temperature Statistics database contains an historical record of the ΔT values corresponding to the 2σ limits for each location, satellite, orbit, and background combination. For day i the mean difference:

$$\overline{\Delta T}_i = \frac{1}{N_i} \sum_{n=1}^{N_i} \Delta T_n \quad (3)$$

and the standard deviation:

$$\sigma_i = \left(\frac{\sum_{n=1}^{N_i} (\Delta T_n - \overline{\Delta T}_i)^2}{N_i - 1} \right)^{\frac{1}{2}} \quad (4)$$

are computed and used to define the 2σ limits (ΔT_{\min_i} and ΔT_{\max_i}) that represent the extremes of the ΔT distribution:

$$\Delta T_{\min_i} = \overline{\Delta T}_i - 2\sigma_i \quad (5)$$

and

$$\Delta T_{\max_i} = \overline{\Delta T}_i + 2\sigma_i. \quad (6)$$

Once the IR-Skin Temperature Statistics have been accumulated for the previous 10 days they are used by the OLS and AVHRR cloud detection algorithms to help predict the brightness temperature (T_{pred}) that would be measured from the satellite in the absence of cloud for the current time, location and background type. The procedure is applied to the thermal infrared channel data from each sensor (AVHRR channel 4 and OLS-T). First, the satellite data being analyzed (e.g., quarter orbit) are segmented into a series of small analysis regions. The size of each region is determined by the relative spatial scales of the AFGWC surface temperature database and the satellite sensor data. During SERCAA it was set empirically for both OLS and AVHRR at 16×16 pixels. However, this number should be considered a minimum size since testing has indicated that, in practice, it can be increased if necessary to improve computational efficiency. The critical factor in determining the size of the analysis region is whether the spatial variability of the background temperature resolved by the satellite data is captured in the modeled skin temperature database (i.e., is the magnitude of the temperature variation over the analysis region approximately the same in the skin temperature database as in the satellite IR channel data).

After the data are segmented the next step is to compute a separate correction factor for each analysis region. Recall that the correction will be added to the $1/8^{\text{th}}$ mesh surface skin temperatures to predict the local (i.e., 16×16) clear scene brightness temperatures. One pixel from the analysis region that is considered most likely to be cloud-free is selected and used to establish a reference satellite-skin temperature difference value (ΔT_{ref}) for the entire local region. To minimize the likelihood of cloud contamination the reference pixel is taken as the warmest pixel in the analysis region. To avoid using a warm anomaly to establish the reference value, the warmest 1% of all pixels in the analysis region are first removed before the reference pixel is selected (e.g., for a 16×16 region, pixels with the two highest brightness temperatures are excluded). Thus ΔT_{ref} is defined as:

$$\Delta T_{\text{ref}} = T_{\text{ref}} - T_{\text{skin}} \quad (7)$$

where T_{ref} is the brightness temperature of the reference pixel and T_{skin} is the time interpolated skin temperature corresponding to the time and location of the reference pixel (calculated in the same way as T_{skin_n} in Eq. 2).

If ΔT_{ref} can be established to be cloud-free it is used as the surface skin temperature correction for its respective analysis region. However, the critical step affecting cloud analysis accuracy is testing of the reference pixel for possible cloud contamination. If the reference pixel does contain cloud then the predicted clear-scene brightness temperature, T_{pred} , will be representative of the cloud brightness temperature and not that of the terrestrial background. Testing of the reference pixel is accomplished by comparing the magnitude of ΔT_{ref} against the range of expected clear-scene temperature differences established from the ten-day IR-Skin Temperature Statistics. If ΔT_{ref} falls within the expected range, then the reference pixel is assumed to be cloud-free, otherwise it is assumed to be cloud-contaminated and a default value is used for ΔT_{ref} . To establish the expected range of clear scene values a time and frequency-weighted average of the historical clear scene ΔT limits is used:

$$\Delta T_{\text{min}} = a \frac{\sum_{i=1}^{10} t_i \Delta T_{\text{min}_i}}{\sum_{i=1}^{10} t_i} + b \frac{\sum_{i=1}^{10} N_i \Delta T_{\text{min}_i}}{\sum_{i=1}^{10} N_i} \quad (8)$$

and

$$\Delta T_{\text{max}} = a \frac{\sum_{i=1}^{10} t_i \Delta T_{\text{max}_i}}{\sum_{i=1}^{10} t_i} + b \frac{\sum_{i=1}^{10} N_i \Delta T_{\text{max}_i}}{\sum_{i=1}^{10} N_i} \quad (9)$$

where a and b are empirically defined coefficients for the temporal and frequency average terms, respectively: the sum of $a + b$ must equal 1.0, currently $a = 0.9$ and $b = 0.1$. The time weighting factor t_i is defined to give greatest weight to the most recent day and decreases as the clear-scene data age. To avoid the use of anomalous data in the time-frequency averaging process, a minimum sample size is required. For data from a given day to be included in the ten-day average, the number of clear scene data points in the distribution, N_i , must exceed 5000. Any days for which N_i is less than 5000 are excluded from the averaging process and data from the next oldest day are added to the series to maintain the ten-day total. The value of t_i is assigned to 1 for the oldest day in the series and increases in value by the difference in Julian date from the date of the oldest day. For example, if the Julian date for the first day in the series (i.e., $i=1$, $t_1=1$) were, say, 140, and two subsequent days had sample sizes, N_i , of less than 5000 then the Julian date of the tenth day in the series would be 152 and the time weight, t_{10} , would be 12 (152-140).

Thus to calculate a predicted brightness temperature corresponding to any pixel n within the analysis region ΔT_{ref} is first tested for cloud contamination. If:

$$\Delta T_{\text{min}} \leq \Delta T_{\text{ref}} \leq \Delta T_{\text{max}} \quad , \quad (10)$$

then ΔT_{ref} is assumed to be cloud-free and is added to the time interpolated skin temperature corresponding to that pixel as the correction factor:

$$T_{\text{pred}_n} = T_{\text{skin}_n} + \Delta T_{\text{ref}} , \quad (11)$$

otherwise ΔT_{ref} is assumed to be cloud-contaminated and a default correction based on the mean of the time-frequency weighted average ΔT limits is used to calculate the predicted clear-scene brightness temperature:

$$T_{\text{pred}_n} = T_{\text{skin}_n} + \frac{1}{2}(\Delta T_{\text{min}} + \Delta T_{\text{max}}) . \quad (12)$$

As stated above, the predicted clear scene brightness temperature is calculated in exactly the same way by both the OLS and AVHRR cloud analysis algorithms. However, once established it is used differently by each algorithm in the cloud detection process (see Sections 3 and 4 for descriptions of how the predicted temperature is used in the AVHRR and OLS algorithms, respectively).

2.2.2 Clear Scene Visible Channel Backgrounds

Visible channel clear scene information is required by all cloud analysis algorithms to provide a reference background for visible reflectance tests. This information is generated as a by-product of the analysis algorithms and is maintained in the Visible Background Count (VBC) databases (Table 3). Since the appearance of a terrestrial background at visible wavelengths can change significantly with time and sun-satellite geometry, it was decided to generate visible background databases dynamically using satellite data from locations classified as cloud-free by the analysis algorithms. However, since the scan characteristics of polar and geostationary satellites differ significantly, the visible sensor data are analyzed differently by the respective cloud analysis algorithms and, consequently, the visible background supporting databases reflect those differences. As each polar satellite pass is processed, the clear-scene data, representative of the current background type and sun-satellite geometry, are used to continuously update the background field. Geostationary satellites view the same region of the Earth on each scan, therefore visible background information is computed and maintained in a set of rotating Visible Background Count files, one for each scan time.

The Visible Background Count databases for OLS and AVHRR polar satellites are generated using the background brightness technique developed for the RTNEPH. This technique requires a separate global background database for each satellite and its ascending and descending orbits, as long as both parts of the orbit have usable visible data. The databases are updated continuously for non-water surfaces (based on the Geographic Type database) as new data are processed. Oceans and other water backgrounds are assumed to be uniform with a low reflectance and a default background count is used. Regions of sun glint are handled differently by each cloud analysis algorithm (refer to Sections 3.1.1 and 4.3). At AFGWC, background brightness data are maintained at 1/8th mesh grid resolution. For each 1/8th mesh grid box, all pixels classified as clear by the analysis algorithm run in a cloud clearing processing mode (see Sections 3 and 4 for a description of the cloud clearing mode of the AVHRR and OLS algorithms), and that meet a set of acceptability criteria, are accumulated and a mean count value is calculated. A weighted average of the mean visible count and the corresponding background database value is used to update the database. The process is designed to be conservative to minimize contamination by cloud, snow, or sun glint. A complete description of this process is provided by Kiess and Cox (1988) including acceptability criteria and details of the database update procedure.

Visible background information used by the geostationary cloud analysis algorithm is derived by exploiting the fixed scan characteristics of the geostationary satellites. Note that for all geostationary satellites the subpoint on the Earth remains nearly fixed, thus it is possible to compare multiple visible images obtained from the same satellite to determine pixel-by-pixel how the visible counts change over time. Given this characteristic the assumption is made that, at each pixel location, the minimum visible count observed over an extended period of time will be representative of the clear background. Inherent in this assumption is the idea that over the observation period each pixel in the satellite field of view will be cloud-free at least once and that clouds are brighter than the underlying terrestrial surface. Based on these assumptions, the geostationary Visible Background Count databases are produced by storing the minimum visible count observed over the previous 14-day period for each pixel in the satellite field of view. To account for surface anisotropy and differences in solar illumination that vary with time of day, a separate database is maintained for each satellite scan time for which there is usable visible data. Database files have the same spatial characteristics as the original satellite imagery (i.e., there is a one-to-one correspondence between lines and elements in the background database and the satellite imagery). Images from subsequent days are required to be co-registered so that there is a one-to-one match between pixels in each image. Since image co-registration is done as part of the cloud analysis algorithm (see Section 5), no additional processing is required to accomplish this step (recall that Visible Background Count databases are generated as by-products of the analysis algorithms). Figure 3 illustrates the procedure for generating the geostationary clear scene visible background database. The database is updated daily using the data from the previous 14 days in a rotating file. Note that generation of this database does not require any a priori information on cloud cover or surface geographic type.

2.2.3 Upper Air Data

Upper air data are provided to the SERCAA algorithms by the AFGWC HIRAS global spectral model. Model output consists of gridded fields of temperature, geopotential height, and specific humidity at 10 pressure levels plus the geopotential height of the tropopause. Table 4 provides a description of the upper air database maintained at AFGWC. Update frequency is once every six hours.

Table 4. AFGWC Upper Air Database

Level (mb)	Temperature (K)	Geopotential Height (M*10)	Specific Humidity (g/kg)
Sfc			✓
1000	✓	✓	✓
850	✓	✓	✓
700	✓	✓	✓
500	✓	✓	✓
400	✓	✓	✓
300	✓	✓	✓
250	✓	✓	
200	✓	✓	
150	✓	✓	
100	✓	✓	

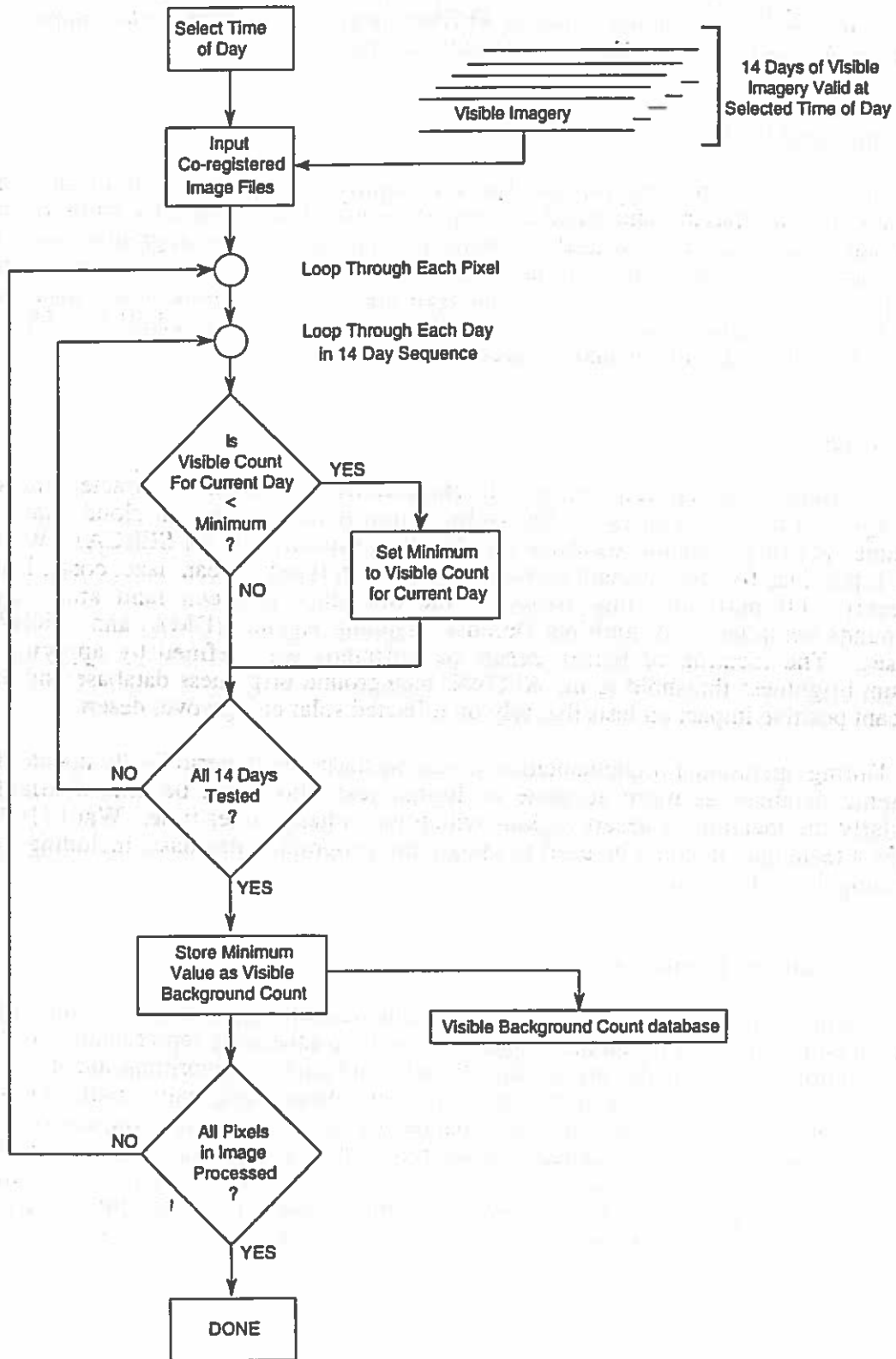


Figure 3. Geostationary Clear Scene Visible Channel Background Generation Algorithm Functional Flow Diagram

It is anticipated that the Navy NOGAPS global forecast model will replace HIRAS as the operational model in use at AFGWC beginning in 1995. Navy upper air data will be obtained over the Shared Processing Network.

2.2.4 Snow and Ice Location

Snow and ice location information are required by all three cloud analysis algorithms to help discriminate cloud spectral signatures from those of a snow or ice background. The AFGWC snow analysis model provides an 1/8th mesh gridded analysis of ice location and snow depth, with an update cycle of every 24 hours. It should be emphasized that the accuracy of all cloud analysis algorithms is critically dependent on this database. Any improvements in accuracy of the snow model (e.g., addition of SSM/I data) will directly benefit cloud analysis accuracy.

2.2.5 Geographic Data

Geographic data are required to help characterize the radiative characteristics of the background terrestrial surface. This information is required by all cloud analysis algorithms. A new geographic database was developed specifically for SERCAA (Ward, 1992). It provides five background surface classification types: ocean, lake, coast, land, and desert. Of particular importance is the boundary between land and water backgrounds compiled from multiple Defense Mapping Agency (DMA) and AFGWC databases. The location of barren desert backgrounds was defined by applying a minimum brightness threshold to the AFGWC background brightness database and has significant positive impact on tests that rely on reflected solar energy over desert.

During operational implementation it may be desirable to periodically update the geographic database as more accurate or higher resolution data become available, particularly the location of desert regions which may change over time. Ward (1992) provides a technique that can be used to update the geographic database, including that for defining desert locations.

2.2.6 Sun-Satellite Geometry

Satellite zenith, solar zenith, and sun-satellite azimuth angle data are required for each pixel being analyzed for cloud. Figure 4 provides a schematic representation of the angle definitions. These angles are required by all cloud analysis algorithms and are used to help characterize sun glint, visible reflectance, and atmospheric path length. During SERCAA, these data were maintained in separate files generated during the sensor data ingest operation. To minimize database storage requirements, angle data were archived at reduced spatial resolution and retrieved using a linear run-time interpolation routine. However, the only requirement on maintenance of this database is a capability to retrieve the angle data for each pixel on demand, thus the SERCAA convention need not be followed in the operational implementation.

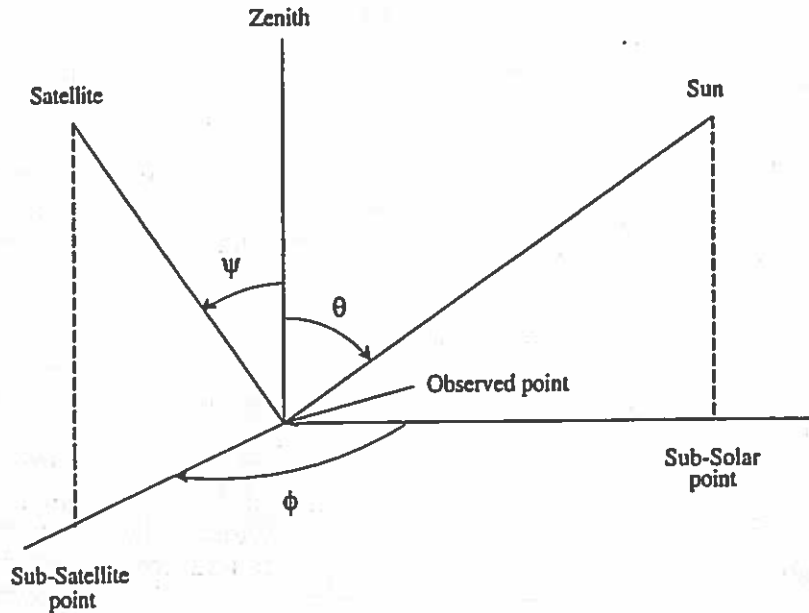


Figure 4. Satellite-Earth-Solar Geometry (after Taylor and Stowe, 1984)

ψ - satellite zenith angle

θ - solar zenith angle

ϕ - sun-satellite azimuth angle

2.2.7 Earth Location

Following the same convention used for sun-satellite geometry information, Earth location data were maintained in separate files that contain latitude/longitude data for each scene. The files were produced automatically during sensor data ingest and maintained in the same spatial projection as the sensor data but at a degraded resolution. A linear run-time interpolation scheme was used to access the latitude/longitude files and compute the Earth location for any pixel within a scene. As is the case with sun-satellite geometry data, the only requirement that affects the operational implementation of the SERCAA algorithms is to provide Earth location information on demand for each analysis location. This requirement holds regardless of whether the data have been pre-registered to a standard grid or are left in scan projection.

3. AVHRR CLOUD ANALYSIS ALGORITHM DESCRIPTION

The SERCAA cloud analysis algorithm for the NOAA Advanced Very High Resolution Radiometer (AVHRR) employs a decision tree type structure to analyze the five channel sensor data to identify specific features, or characteristics, of the cloud scene. The decision tree provides the basis for a multispectral classification of scene attributes depending on such factors as scene illumination, background surface type and spectral information content. The algorithm uses multispectral signatures to identify and characterize clear and cloudy regions of the scene. Scene analysis is performed on a pixel-by-pixel basis.

The decision tree consists of a series of cloud tests that separately identify individual cloud and cloud-free background attributes or signatures in the multispectral satellite data. Each cloud test is based on a specific spectral signature that exploits the information content of radiance measurements from one or more sensor channels. Table 5 provides the naming conventions, or designations, for each of the five AVHRR sensor channels discussed in this section. Recall from Section 2.1 that visible sensor data are converted to percent albedo and infrared data to equivalent blackbody brightness temperatures before they are used in the cloud algorithm. In addition to sensor data, the cloud tests also require clear scene brightness temperature and albedo information to characterize the terrestrial background. Characterization of the clear-scene background is supported by the Visible Background Count, the AFGWC Surface Temperature, and the IR-Skin Temperature Statistics databases described in Section 2.2.

Table 5. AVHRR Sensor Channel Naming Convention

AVHRR Channel	Description	Designation
Channel 1	0.58 - 0.68 μm percent albedo	A ₁
Channel 2	0.72 - 1.10 μm percent albedo	A ₂
Channel 3	3.55 - 3.93 μm brightness temperature	T ₃
Channel 4	10.3 - 11.3 μm brightness temperature	T ₄
Channel 5	11.5 - 12.5 μm brightness temperature	T ₅

The cloud algorithm tests evaluate the spectral information content of the sensor data by analyzing data from one or more of the AVHRR sensor channels along with the supporting data. Table 6 summarizes the nine separate cloud detection tests that populate the cloud analysis algorithm decision tree. Three additional tests are summarized in Table 7 that were developed to identify problematic background surface conditions that can cause the cloud tests to classify the clear-scene as cloud. Threshold values used by these tests are tabulated in Table A-1 in Appendix A.

The algorithm is structured to run each of the tests and store the intermediate results internally. Since individual tests are sensitive to selected cloud or background characteristics, it requires the combined results of all tests to produce the final cloud analysis. As successive tests are run more information on the total cloud environment is built up. Note that some tests require the results of other tests to make a cloud or background determination. A final cloud/no-cloud decision is made by jointly evaluating the intermediate results of all applicable cloud and background surface tests. Each of the individual cloud and background surface tests, as well as the procedure used to make a final cloud decision based on the results of these tests, are discussed in detail in the sections that follow.

Table 6. AVHRR Cloud Tests

Test	Day Application	Night Application	Cloud Test Name
$T_3 - T_4 > \text{THRESH}_{\text{icf}}^1$	✓		Low Cloud and Fog
$T_3 - T_4 > \text{THRESH}_{\text{precip}(1)}$ and $T_{\text{pred}} - T_4 > \text{THRESH}_{\text{precip}(2)}$ and $A_2 * \sec(\theta) > \text{THRESH}_{\text{precip}(3)}$	✓		Precipitating Cloud
$T_4 - T_5 > \text{THRESH}_{(T_4, \Psi)}$ and if AFGWC Snow Analysis Model identifies snow then $T_{\text{pred}} - T_4 > \text{THRESH}_{\text{ci}}$ and Over Water: $A_2 * \sec(\theta) < \text{THRESH}_{\text{dc}_w}$ or Over Land: $A_1 * \sec(\theta) < \text{THRESH}_{\text{dc}_l}$	✓		Daytime Thin Cirrus Cloud
$\text{THRESH}_{\text{ratio_lo_dry}} < \frac{A_2}{A_1} < \text{THRESH}_{\text{ratio_up_dry}}$ High Humidity Areas: $\text{THRESH}_{\text{ratio_lo_wet}} < \frac{A_2}{A_1} < \text{THRESH}_{\text{ratio_up_wet}}$	✓		Visible Brightness Ratio
Over Land: $A_1 * \sec(\theta) - A_{\text{sfc}} > \text{THRESH}_{\text{land}}$ Over Water: $A_2 * \sec(\theta) > \text{THRESH}_{\text{water}}$	✓		Visible Brightness
$T_{\text{pred}} - T_4 > \text{THRESH}_{\text{cold}}^1$	✓	✓	Cold Cloud
$T_4 - T_5 > \text{THRESH}_{(T_4, \Psi)}$ and if AFGWC Snow Analysis Model identifies snow then $T_{\text{pred}} - T_4 > \text{THRESH}_{\text{ci}}$	✓	✓	Cirrus Cloud
$T_4 - T_3 > \text{THRESH}_{\text{fls}}$		✓	Fog, Low Stratus
$T_3 - T_5 > \text{THRESH}_{\text{tci}}$ High Humidity Areas: $T_3 - T_4 > \text{THRESH}_{\text{tci}}$		✓	Nighttime Thin Cirrus Cloud

θ = solar zenith angle

$\text{THRESH}_{(T_4, \Psi)}$ = threshold as a function of channel 4 brightness temperature (T_4) and satellite zenith angle (Ψ)

¹ Separate thresholds maintained as a function of background surface type.

Table 7. AVHRR Background Surface Filter Tests

Test	Background Surface Filter Test Name
<p>Background Surface Type Must Be Water</p> <p>and</p> $ \psi - \theta < \text{THRESH}_{\text{zenith}}$ <p>and</p> $\text{THRESH}_{\text{loazimuth}} < \phi < \text{THRESH}_{\text{upazimuth}}$ <p>and</p> <p>Visible Brightness Test detects cloud</p> <p>or</p> <p>Visible Brightness Ratio Test detects cloud</p> <p>and</p> $T_3 > T_4 + \text{THRESH}_{\text{glint}(1)}$ <p>and</p> $T_3 > \text{THRESH}_{\text{glint}(2)}$ <p>and</p> <p>Cold Cloud Test does not detect cloud</p> <p>and</p> <p>Cirrus Cloud Test does not detect cloud</p>	<p>Sun Glint</p>
$\text{THRESH}_{\text{desert_lo_ratio}} < \frac{A_2}{A_1} < \text{THRESH}_{\text{desert_up_ratio}}$ <p>and</p> $A_2 < \text{THRESH}_{\text{desert}}$ <p>and</p> $T_3 > \text{THRESH}_{\text{temp_desert}(1)}$ <p>and</p> $T_{\text{air}} - T_4 < \text{THRESH}_{\text{temp_desert}(2)}$ <p>and</p> $\text{THRESH}_{\text{desert_lo_diff}} < T_3 - T_4 < \text{THRESH}_{\text{desert_up_diff}}$	<p>Desert Background</p>
$T_4 < \text{THRESH}_{\text{snow}(1)}$ <p>and</p> $ T_{\text{pred}} - T_4 < \text{THRESH}_{\text{snow}(2)}$ <p>and</p> <p>Over Water:</p> $A_2 > \text{THRESH}_{\text{snow_water}}$ <p>Over Land:</p> $A_1 > \text{THRESH}_{\text{snow_land}}$ <p>and</p> $ T_3 - T_4 < \text{THRESH}_{\text{snow}(3)}$	<p>Snow/Ice Cover Background</p>

ψ = satellite zenith angle
 θ = solar zenith angle
 ϕ = sun-satellite azimuth angle

3.1 BACKGROUND SURFACE FILTER TESTS

Background surface filter tests are used to identify problematic surface backgrounds that have spectral signatures similar to cloud. Results of these tests are used to either modify affected cloud tests or eliminate channels from the analysis process.

3.1.1 Sun Glint Test

The Sun Glint Test is used to detect specular reflection off of water surfaces which could be mistakenly identified as cloud by tests that rely on reflected solar radiation. Sun glint is a potential problem over any water surfaces that can be resolved by the satellite, however, in practice the Sun Glint Test is only applied over permanent water surfaces large enough to be captured in the 1/64th mesh Geographic Type database. A series of conditions involving the background, solar/satellite geometry, and spectral signatures must be met to detect glint.

The first set of tests determine if the background surface type and solar/satellite geometry will support sun glint. These tests are:

- Background surface type must be water ,
- and* • $|\psi - \theta| < \text{THRESH}_{\text{zenith}}$,
- and* • $\text{THRESH}_{\text{loazimuth}} < \phi < \text{THRESH}_{\text{upazimuth}}$.

where $\text{THRESH}_{\text{upazimuth}}$ and $\text{THRESH}_{\text{loazimuth}}$ are empirically derived threshold values. $\text{THRESH}_{\text{zenith}}$ defines the magnitude by which the solar zenith angle may depart from the satellite zenith angle to support sun glint. Figure 4 provides an illustration of the solar/satellite geometry definitions used by the above tests. Background surface type information is provided by the Geographic Type database described in Section 2.2.5.

The second set of tests examines the spectral signature of any pixels that passed the background surface type and solar/satellite geometry tests. These tests are:

The albedo must be high in the visible channels:

- Visible Brightness Test detects cloud (Section 3.2.2.5) ,
- or* • Visible Brightness Ratio Test detects cloud (Section 3.2.2.4) .

Channel 3 must be nearly saturated:

- $T_3 > T_4 + \text{THRESH}_{\text{glint}(1)}$,
- and* • $T_3 > \text{THRESH}_{\text{glint}(2)}$,

where $\text{THRESH}_{\text{glint}(1)}$ and $\text{THRESH}_{\text{glint}(2)}$ are empirically derived threshold values.

The IR brightness temperature must be relatively high in the infrared channels (i.e., not indicative of a low liquid water cloud):

- Cold Cloud Test does not detect cloud (Section 3.2.1.1) ,
- and* • Cirrus Cloud Test does not detect cloud (Section 3.2.1.2) ,

A schematic illustration of the AVHRR Sun Glint Test is provided in Fig. 5.

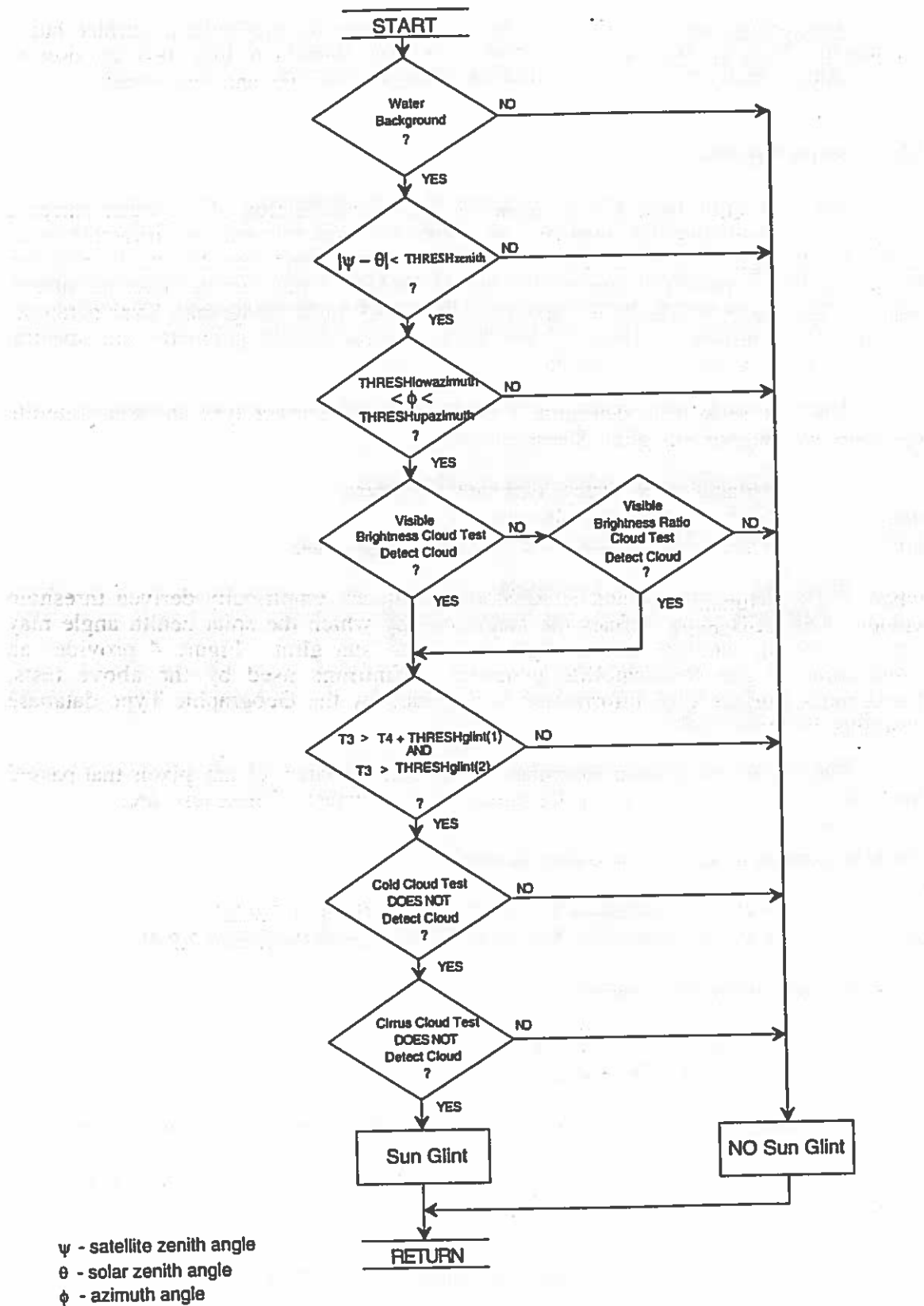


Figure 5. AVHRR Sun Glint Test Functional Flow Diagram

3.1.2 Desert Background Test

The Desert Background Test is used to identify clear scene desert backgrounds through the examination of multispectral daytime AVHRR data. In this application, the term desert is used to indicate any highly reflective, non-vegetated surface; it does not necessarily follow the geographer's definition based on annual precipitation. Results of this dynamic desert test are used to augment desert information contained in the Geographic Type supporting database. Also, in addition to their run-time use in specifying cloud-free desert pixels in the AVHRR algorithm, desert flags are potentially useful to the SERCAA DMSP and geostationary cloud analysis algorithms as a high resolution source for identifying bright sandy backgrounds.

A series of five AVHRR spectral conditions must be met in order to classify a pixel as cloud-free desert. The first is a modified version of the Visible Brightness Ratio cloud test (Section 3.2.2.4):

$$\bullet \text{ THRESH}_{\text{desert_lo_ratio}} < A_2/A_1 < \text{THRESH}_{\text{desert_up_ratio}} ,$$

where $\text{THRESH}_{\text{desert_lo_ratio}}$ and $\text{THRESH}_{\text{desert_up_ratio}}$ are thresholds that bound the range of the near-IR to visible channel ratio for clear-scene desert backgrounds. Note these thresholds are more limiting than the cloud detection ratio thresholds since highly reflective land surfaces generally do not exhibit as much variability as clouds in near-IR and visible sensor channel measurements.

The second test is an absolute check on the channel 2 albedo testing for a (potentially) clear background. The test is defined as:

$$\bullet A_2 < \text{THRESH}_{\text{desert}} ,$$

and is employed to ensure the measured albedo is not large enough to be a cloud signature since desert surfaces are generally not as bright as cloud in channel 2.

The third test checks to determine if the channel 3 brightness temperature is near saturation. Clear non-vegetated surfaces exhibit a strong solar component in channel 3 resulting in a large T_3 brightness temperature. The test is defined as:

$$\bullet T_3 > \text{THRESH}_{\text{temp_desert}(1)} ,$$

where $\text{THRESH}_{\text{temp_desert}(1)}$ is a desert detection threshold.

The fourth test is used to check T_4 against the surface ambient air temperature:

$$\bullet T_{\text{air}} - T_4 < \text{THRESH}_{\text{temp_desert}(2)} ,$$

where $\text{THRESH}_{\text{temp_desert}(2)}$ is a desert detection threshold and where T_{air} is determined using AFGWC Upper Air database (refer to Section 2.2.3) in conjunction with the Terrain Height database. The upper air temperature profile is interpolated to the actual terrain height at the pixel location. The assumption here is that, over desert, a satellite observed daytime clear-column thermal IR brightness temperature will be close to or exceed the ambient air temperature.

The final desert criteria requires that the brightness temperature difference between channels 3 and 4 be within a specified range. The test is defined as:

$$\bullet \text{THRESH}_{\text{desert_lo_diff}} < T_3 - T_4 < \text{THRESH}_{\text{desert_up_diff}} ,$$

where $THRESH_{desert_lo_diff}$ and $THRESH_{desert_up_diff}$ are threshold values that specify the range of expected channel differences. This is to ensure that low clouds do not get classified falsely as desert.

All five of the above tests must pass in order for a pixel to be considered clear desert background. A schematic illustration of the AVHRR Desert Background Test is provided in Fig. 6.

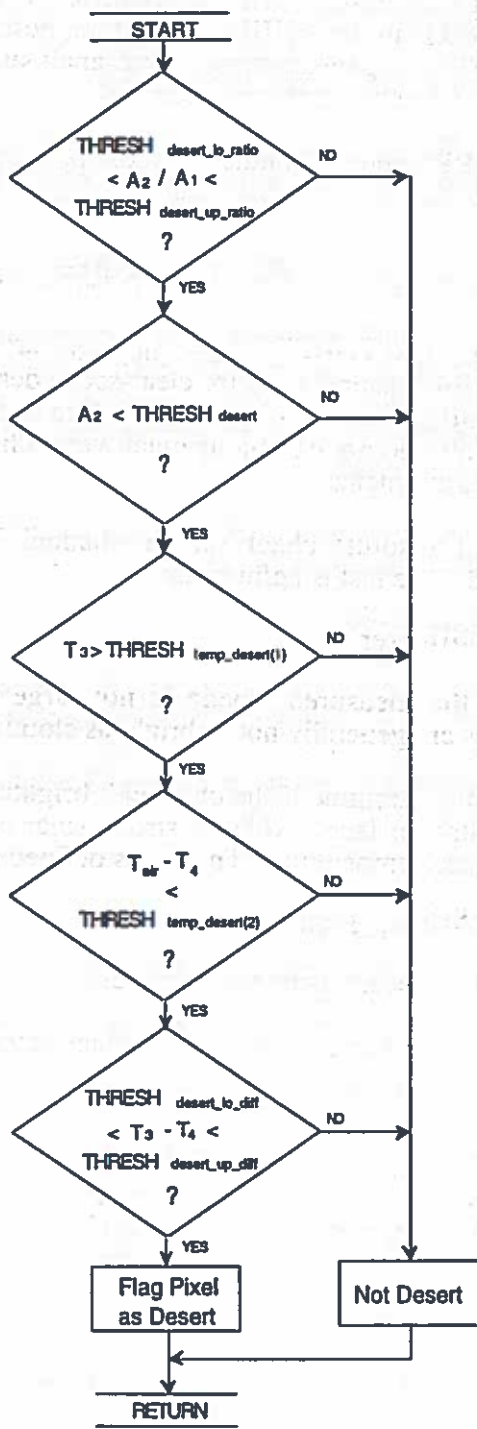


Figure 6. AVHRR Desert Background Test Functional Flow Diagram

3.1.3 Snow/Ice Cover Background Test

The Snow/Ice Cover Background Test is used to discriminate snow and ice backgrounds from cloud features. Results from this test both augment the Snow and Ice Location database and provide a technique for discriminating cloud over snow and ice backgrounds. The test uses visible and infrared channel data to first identify pixels with characteristics consistent with snow, but not necessarily separate from cloud, and then uses a multispectral discriminant to separate snow from cloud. The tests are defined as:

- $T_4 < \text{THRESH}_{\text{snow}(1)}$,
- and* • $|T_{\text{pred}} - T_4| < \text{THRESH}_{\text{snow}(2)}$,
- and* • $A_2 > \text{THRESH}_{\text{snow_water}}$ (Over Water)
- or* • $A_1 > \text{THRESH}_{\text{snow_land}}$ (Over Land) ,
- and* • $|T_3 - T_4| < \text{THRESH}_{\text{snow}(3)}$,

where T_{pred} is the predicted clear scene brightness temperature calculated through the procedure described in Section 2.2.1, $\text{THRESH}_{\text{snow_water}}$ and $\text{THRESH}_{\text{snow_land}}$ are thresholds over water and land background surface types respectively, and where $\text{THRESH}_{\text{snow}(1)}$, $\text{THRESH}_{\text{snow}(2)}$, and $\text{THRESH}_{\text{snow}(3)}$ are separate snow/ice detection thresholds.

Note that if this spectral snow test evaluates as true, the pixel is unambiguously classified as cloud-free. A schematic illustration of the AVHRR Snow/Ice Cover Background Test is provided in Fig. 7.

3.2 CLOUD COVER TESTS

Cloud tests are divided into three groups: 1) those that rely on reflected solar radiation (daytime cloud tests); 2) those that are only applicable in the absence of direct sunlight (nighttime cloud tests); and 3) those that are equally applicable without any regard to the amount of solar illumination on the scene (solar independent cloud tests). A solar zenith angle threshold is used to determine which tests are applicable to a particular situation.

The cloud cover tests can be run in two different modes: 1) cloud detection, and 2) cloud clearing. When run in cloud detection mode the algorithm is designed to provide an optimal cloud analysis with no bias toward over or under analysis. The cloud clearing mode is used to remove all cloud-contaminated pixels from the satellite imagery to support generation of clear scene statistics as described in Section 2.2.2. When run in cloud clearing mode the algorithm has a definite bias toward over analysis. The two modes differ only in the magnitude of the threshold values used in each test. Threshold values for both modes are provided in Table A-1 of Appendix A.

3.2.1 Solar Independent Cloud Tests

There are two cloud tests that are executed independent of scene solar illumination (i.e., applicable both day and night). The first is a single channel LWIR threshold test designed to detect mid-level clouds and the second uses the split window LWIR channels to detect cirrus cloud.

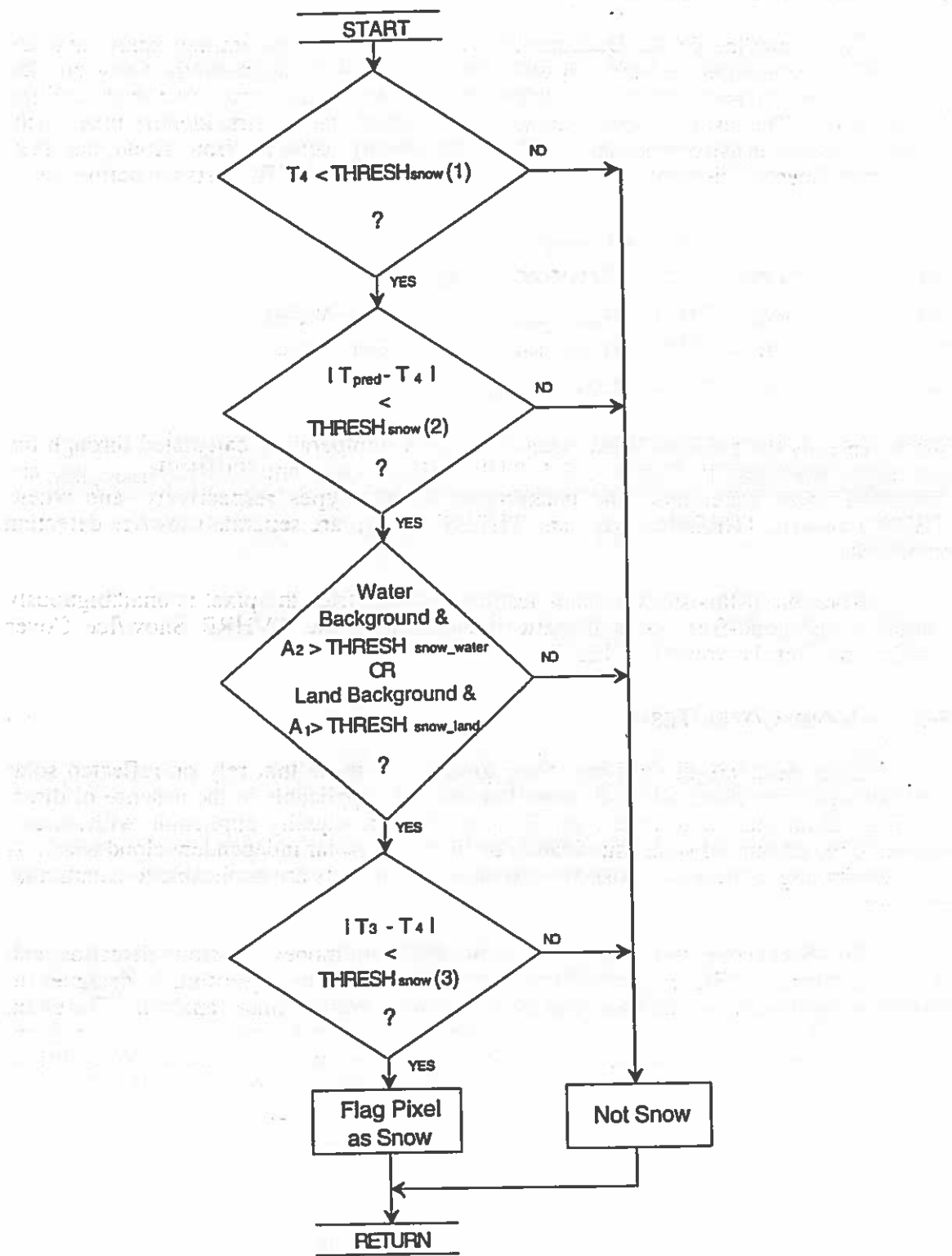


Figure 7. AVHRR Snow/Ice Cover Background Test Functional Flow Diagram

3.2.1.1 Cold Cloud Test

The Cold Cloud Test is a single LWIR channel threshold test designed to discriminate the thermal signature of obvious mid-level clouds from the terrestrial background signature. The background radiative temperature (identified in this report as the clear scene brightness temperature) is predicted by applying a correction to the surface skin temperature supplied by the AFGWC Surface Temperature database using the procedure described in Section 2.2.1.1.

A cloud decision is made by comparing the T_4 value to the predicted clear scene brightness temperature (T_{pred}). The test requires that the T_4 value be significantly lower than the predicted clear scene brightness temperature in order for the pixel to be classified as cloud-filled. If the T_4 value is less than the predicted clear scene brightness temperature by an amount greater than a preset threshold ($THRESH_{cold}$), the pixel is classified as cloudy. The magnitude of the threshold varies as a function of the background surface type to account for the differences in the expected accuracy of T_{pred} over the different backgrounds. Separate thresholds are maintained for water, land, coast, desert, and snow backgrounds. Snow background information is supplied by the AFGWC Snow and Ice Location database (Section 2.2.4) while water, land, coast and desert backgrounds are identified by the Geographic Type database (Section 2.2.5). The Cold Cloud Test is defined as:

$$\bullet T_{pred} - T_4 > THRESH_{cold} ,$$

where $THRESH_{cold}$ is the surface-dependent background threshold value.

3.2.1.2 Cirrus Cloud Test

Split window LWIR $T_4 - T_5$ brightness temperature differences exhibit a small but persistent cirrus cloud signature. There are three radiative effects that combine to account for the split-LWIR cirrus signatures. First, ice particle emissivity is lower at $11.8 \mu\text{m}$ than at $10.7 \mu\text{m}$. Second, atmospheric water vapor attenuation is stronger at the longer $11.8 \mu\text{m}$ wavelengths. Third, there is a slightly stronger Planck dependence on temperature at the shorter $10.7 \mu\text{m}$ wavelengths, resulting in a higher $10.7 \mu\text{m}$ brightness temperature for what are essentially mixed fields of view that occur with transmissive cirrus. Each of these factors contribute to cirrus brightness temperatures that are consistently higher at $10.7 \mu\text{m}$ than at $11.8 \mu\text{m}$. However in the absence of cloud, water vapor attenuation can, by itself, cause a positive $T_4 - T_5$ difference that could be mistaken for a cloud signature. Thus when using a split-LWIR technique to detect cirrus it is necessary to first eliminate cases where the channel difference is caused by clear-scene atmospheric moisture. To accomplish this, the cloud detection threshold is defined as a function of atmospheric water vapor and path length through the atmosphere. During SERCAA a predefined threshold table developed by Saunders and Kriebel (1988) for use over the North Atlantic was applied globally. While this provided reasonable results over most parts of the world, a recommended alternative approach which requires real-time calculation of a water vapor dependent threshold is also provided here.

Saunders and Kriebel showed that expected clear sky $T_4 - T_5$ differences can be estimated as a function of the channel 4 brightness temperature (as a surrogate for water vapor loading) and satellite zenith angle (to account for atmospheric path length). They developed a look-up table of threshold values compiled for a range of T_4 temperatures and satellite zenith angles that is the basis for the SERCAA Cirrus Cloud Test (Table A-2a of Appendix A). Channel difference values that exceed the appropriate

threshold value are assumed to be larger than would occur under cloud-free conditions. Thus the cirrus test is defined as:

$$\bullet T_4 - T_5 > \text{THRESH}(T_4, \psi) ,$$

where $\text{THRESH}(T_4, \psi)$ is the cloud detection threshold obtained through interpolation from the table and ψ is the satellite zenith angle.

During SERCAA, real-data tests have shown that the Cirrus Cloud Test performs accurately and robustly for the majority of climatological situations. However, the test sometimes has difficulty accurately discriminating cirrus cloud from snow and ice backgrounds. To compensate for this, an additional requirement is placed on the cloud test when the background is classified as snow or ice covered in the Snow and Ice Location database. Based on the assumption that channel 4 brightness temperatures measured from cirrus clouds are colder than the terrestrial background, the T_4 brightness temperature is required to be lower than the predicted clear scene brightness temperature (T_{pred}) by an amount greater than a cloud detection threshold ($\text{THRESH}_{\text{ci}}$). This test is defined as:

$$\bullet T_{\text{pred}} - T_4 > \text{THRESH}_{\text{ci}} .$$

If the background is snow or ice and both criteria are met, then the pixel is classified as cloud-filled. If the background is not classified as snow or ice, then only the split LWIR test is required.

In the alternative approach, the split-LWIR cirrus detection threshold is calculated as a function of atmospheric water vapor directly as opposed to the channel 4 brightness temperature. Due to time constraints during SERCAA this approach has not been tested but, based on theory, implementation on a trial basis is recommended. Similar to the Saunders and Kriebel approach, an AVHRR pixel is classified as cloud-filled if the measured $T_4 - T_5$ brightness temperature difference exceeds a water vapor dependent threshold. The test is defined as:

$$\bullet T_4 - T_5 > \text{THRESH}(Q, \psi) ,$$

where Q is the precipitable water, defined as the mass of atmospheric water vapor in a vertical column of unit cross sectional area; and ψ is the satellite zenith angle, again used to characterize atmospheric path length. The threshold $\text{THRESH}(Q, \psi)$ is computed a priori using a radiative transfer band model (d'Entremont et al., 1990) in the AVHRR spectral bands and then tabulated as a function of Q and ψ for use by the SERCAA cloud detection algorithms. $\text{THRESH}(Q, \psi)$ is the expected clear scene $T_4 - T_5$ difference as computed by the radiative transfer model. Table A-2b of Appendix A provides a look-up table of these expected clear scene $T_4 - T_5$ differences based on the band model calculations. Note that these thresholds are theoretical and need to be tested using real data.

The precipitable water parameter required by this approach can be derived from the AFGWC Upper Air database (Section 2.2.3) since specific humidity q (g/kg) is a database parameter. Precipitable water Q is defined as:

$$Q = -\frac{1}{g} \int_{p_{\text{sk}}}^0 q(p) dp , \quad (13)$$

where g is the acceleration of gravity, p_{sfc} is the surface pressure, and $q(p)$ is the specific humidity profile obtained from the Upper Air database.

3.2.2 Day Condition Cloud Tests

There are five cloud tests used during day conditions that rely, at least partially, on reflected solar radiation for a cloud signature. Solar zenith angle information is used to define when day conditions exist. In general, the AVHRR cloud analysis algorithm defines day conditions as existing when the solar zenith angle is less than or equal to 90° . However as noted in the following sections some tests are restricted to higher solar zenith angle conditions.

3.2.2.1 Low Cloud and Fog Test

The Low Cloud and Fog Test relies on the different radiative characteristics of water droplet clouds at channel 3 and 4 wavelengths. During daylight conditions, the measured channel 3 radiance is a combination of both emitted and reflected energy. At the longer channel 4 wavelength there is only an emitted component. The result of this characteristic is that the brightness temperatures calculated from channel 3 radiance measurements that contain cloud are larger than those for channel 4, while for other surfaces they are roughly the same. The cloud test is applied by comparing the $T_3 - T_4$ brightness temperature difference. The test assumes that a liquid water cloud will reflect enough solar energy at $3.7 \mu\text{m}$ to make the channel 3 brightness temperature, T_3 , significantly higher than T_4 . If the T_3 brightness temperature is greater than the T_4 value by an amount greater than a cloud detection threshold, the pixel is classified as cloud-filled. The Low Cloud and Fog Test is defined as:

$$\bullet T_3 - T_4 > \text{THRESH}_{lcf} ,$$

where THRESH_{lcf} is a background surface-dependent cloud detection threshold. The magnitude of the thresholds were established empirically as a function of the background surface type. Separate thresholds are maintained for desert, non-desert and potential sun glint backgrounds.

The Low Cloud and Fog Test is extremely sensitive to desert surfaces since they are also reflective at $3.7 \mu\text{m}$. Desert background surfaces are required to be identified before the test is applied so that a different cloud detection threshold, designed for use over desert, can be used. Desert background is identified using both the Geographic Type database (Section 2.2.5) and the Desert Background Test (Section 3.1.2). When the solar zenith angle is greater than 80° the geographic database is used as the sole method to identify desert backgrounds since the Desert Background Test is not applied when the solar zenith angle exceeds 80° .

Like cloud, sun glint also produces a strong positive channel $T_3 - T_4$ brightness temperature difference. Due to the similar spectral signatures of cloud and sun glint, potential sun glint areas are identified prior to testing for cloud contamination. A larger threshold is applied over potential sun glint regions compared to the threshold used for non-glnt regions. Potential sun glint areas are defined by the same background and solar/satellite geometry tests used in the Sun Glint Test (Section 3.1.1). However, results from the Sun Glint Test cannot be used here since the spectral criteria applied in addition to the geometric requirements are not sensitive enough to detect all levels of glint

sufficient to corrupt the channel 3 data and produce a false low cloud signature. Thus for this test, only the broader geometric criteria are used:

- Background surface type must be water ,
- and
- $|\psi - \theta| < \text{THRESH}_{\text{zenith}}$,
- and
- $\text{THRESH}_{\text{loazimuth}} < \phi < \text{THRESH}_{\text{upazimuth}}$.

3.2.2.2 Precipitating Cloud Test

The Precipitating Cloud Test is predominantly a cumulonimbus test that exploits the reflective nature of thick ice clouds at $3.7 \mu\text{m}$. Typically, ice particle clouds such as thin cirrus are not good reflectors of MWIR radiation. However when the ice clouds are optically thick, such as for towering cumulonimbus, they reflect more strongly than their cirrus counterparts. Since solar $3.7 \mu\text{m}$ radiance is so high, the daytime MWIR brightness temperature of thick ice clouds is also high, being a combination of thermal emission and solar reflection. In fact it is much higher than the true physical temperature of the cloud, which is more accurately represented by T_4 . Thus the MWIR - LWIR brightness temperature difference $T_3 - T_4$ is very large. This is tested as:

- $T_3 - T_4 > \text{THRESH}_{\text{precip}(1)}$,

where $\text{THRESH}_{\text{precip}(1)}$ is a cloud detection threshold.

A high MWIR - LWIR brightness temperature difference is not in itself uniquely indicative of high, cold, precipitating ice clouds. Recall that such a spectral signature is also indicative of low water droplet clouds for essentially the same reasons. Thus, two other checks must also be performed in conjunction with the above test to discriminate cumulonimbus clouds from low liquid water clouds:

- $T_{\text{pred}} - T_4 > \text{THRESH}_{\text{precip}(2)}$,
- and
- $A_2 * \sec(\theta) > \text{THRESH}_{\text{precip}(3)}$,

where θ is the solar zenith angle, and where $\text{THRESH}_{\text{precip}(2)}$ and $\text{THRESH}_{\text{precip}(3)}$ are precipitating cloud detection thresholds for each of the tests.

The $T_{\text{pred}} - T_4$ test eliminates any low clouds that pass the $T_3 - T_4$ test by ensuring that the true physical cloud top temperature is significantly lower than the predicted clear scene brightness temperature. The final near-IR channel test (A_2) eliminates ice clouds that are not as optically thick, and hence not as bright, as precipitating clouds. This test discriminates between cirrostratus (which generally does not pass this test) and cumulonimbus. All three tests must evaluate as true in order for the pixel to be classified cloudy.

3.2.2.3 Daytime Thin Cirrus Cloud Test

The Daytime Thin Cirrus Cloud Test stratifies the results of the solar independent Cirrus Cloud Test (Section 3.2.1.2) into thin cirrus and thick cirrus through the use of visible channel data. Recall the Cirrus Cloud Test identifies cloud through the $T_4 - T_5$ difference to find detect cirrus and small water droplet clouds while the Daytime Thin Cirrus test only identifies thin ice clouds. The Daytime Thin Cirrus Cloud Test requires

that normalized visible channel albedo values be less than a cloud detection threshold to be classified as thin cirrus cloud.

To review, the Cirrus Cloud Test requires the following conditions to be met:

$$\bullet T_4 - T_5 > \text{THRESH}(T_4, \psi),$$

where $\text{THRESH}(T_4, \psi)$ is the cloud detection threshold obtained through interpolation from Table A-2a in Appendix A and ψ is the satellite zenith angle.

If the Snow and Ice Location database (Section 2.2.4) identifies the surface background as being snow or ice covered then the pixel is subjected to an additional test to ensure that the signature detected by the $T_4 - T_5$ difference test was not the underlying snow or ice background rather than cirrus cloud:

$$\bullet T_{\text{pred}} - T_4 > \text{THRESH}_{\text{ci}},$$

where T_{pred} is the clear scene brightness temperature (Section 2.2.1.1) and $\text{THRESH}_{\text{ci}}$ is the cirrus cloud detection threshold. Detailed descriptions of the above tests are provided in Section 3.2.1.2.

In addition to the tests listed above the Daytime Thin Cirrus Cloud Test uses visible or near-IR albedo to discriminate thin cirrus. The criterion used is dependent on the background surface type:

$$\begin{array}{ll} \text{If water} & \bullet A_2 * \sec(\theta) < \text{THRESH}_{\text{dci}_w}, \quad (\text{Over Water}) \\ \text{else if land} & \bullet A_1 * \sec(\theta) < \text{THRESH}_{\text{dci}_l}, \quad (\text{Over Land}) \end{array}$$

where θ is the solar zenith angle, and where $\text{THRESH}_{\text{dci}_w}$ and $\text{THRESH}_{\text{dci}_l}$ are the daytime thin cirrus cloud detection threshold values over water and land respectively.

3.2.2.4 Visible Brightness Ratio Test

The Visible Brightness Ratio Test compares the relative magnitudes of channel 1 and 2 albedo data using a channel ratio. The ratio test makes use of the fact that for clouds, the spectral signature in channels 1 and 2 are very close to each other, while for land and water surfaces they differ significantly. The test is applied by computing the ratio of the channel 2 albedo (A_2) to the channel 1 value (A_1). No normalization for anisotropic effects is needed since they cancel in the ratio operation. Clear land surfaces tend to have a ratio greater than 1.0 and water surfaces will be less than 1.0. However, clouds mask the terrestrial signatures resulting in a channel ratio approximately equal to 1.0. Thus, the cloud test is applied by testing the A_1/A_2 channel ratio against upper and lower limit cloud thresholds. If the channel ratio falls within these limits then the data are classified as cloudy.

When making a final cloud decision, the results of the Ratio test are only used in the absence of sun glint, desert, or snow/ice background conditions, all of which can produce a false cloud signal. Sun glint is identified by the Sun Glint Test (Section 3.1.1). Desert background surfaces are identified using the Geographic Type database (Section 2.2.5) and the Desert Background Test (Section 3.1.2). Snow/ice background surfaces are identified from the Snow and Ice Location database (Section 2.2.4) and the spectral Snow/Ice Cover Background Test (Section 3.1.3). It is also important to note that the

results of the Visible Brightness Ratio Test are not used over pixels that are classified as coast in the Geographic Type database. Empirical study has shown that mixed land and water fields of view may produce a channel ratio signature similar to that of cloud. For this reason, cloud results over pixels that have been identified as coast are not used when making a final cloud decision.

Empirical tests have also shown that using a single set of cloud thresholds for all conditions can result in the over analysis of cloud in regions of high humidity. High humidity causes increased concentrations of aerosols and haze, resulting in a preferential increase in atmospheric scattering at visible wavelengths relative to the near-IR. This increased scattering results in a higher measured channel 1 albedo relative to channel 2 for cloud-free areas, which could produce a false cloud signature. To account for this, the value of upper and lower thresholds are lowered to account for lower clear scene channel ratio values. Regions of potentially high humidity are identified by testing the magnitude of the predicted clear scene brightness temperature against a threshold:

$$\bullet T_{\text{pred}} > \text{THRESH}_{\text{ratio_humid}} ,$$

where T_{pred} is the predicted clear scene brightness temperature (Section 2.2.1.1) and $\text{THRESH}_{\text{ratio_humid}}$ is the high humidity threshold. In regions where this test evaluates as true, the Visible Brightness Ratio Test is defined as:

$$\bullet \text{THRESH}_{\text{ratio_lo_wet}} < A_2/A_1 < \text{THRESH}_{\text{ratio_up_wet}} ,$$

where $\text{THRESH}_{\text{ratio_lo_wet}}$ is the lower limit ratio threshold value, and $\text{THRESH}_{\text{ratio_up_wet}}$ is the upper limit ratio threshold value. In regions where the humidity test evaluates as false, the Visible Brightness Ratio Test uses a different set of thresholds:

$$\bullet \text{THRESH}_{\text{ratio_lo_dry}} < A_2/A_1 < \text{THRESH}_{\text{ratio_up_dry}} ,$$

where $\text{THRESH}_{\text{ratio_lo_dry}}$ is the lower limit ratio threshold value, and $\text{THRESH}_{\text{ratio_up_dry}}$ is the upper limit ratio threshold value.

3.2.2.5 Visible Brightness Test

The Visible Brightness Test is a single-channel threshold test that is used to discriminate relatively high cloud albedo from a predicted low-albedo background value. The predicted background albedo is derived from the clear scene Visible Background Count supporting database described in Section 2.2.2. Albedo values obtained for each grid box are first normalized to account for differences in satellite measured albedo that may occur at a given point on the Earth purely as a result of daily changes in satellite-solar viewing geometry caused by normal precession of the satellite orbit. The bi-directional reflectance characteristics of different terrestrial surfaces have been measured and can be removed through application of an Anisotropic Reflectance Factor (ARF) to account for preferential reflection off the background surface. The correction is applied as follows:

$$A'_i = A_i / \text{ARF}(\psi, \theta, \phi, M) , \quad (14)$$

where A'_i is the normalized albedo for channel i , and ψ , θ , and ϕ are the satellite zenith, solar zenith, and sun-satellite azimuth angles defined in Section 2.2.6, and M is the

Geographic Type of the Earth's surface. Evaluation of the ARF function is performed through look up tables published by Taylor and Stowe (1984).

When making a final cloud decision, results of the visible brightness test are only used in the absence of sun glint, desert, or snow/ice background conditions, all of which can produce false cloud signatures in the visible and near-IR data. Sun glint is identified by the Sun Glint Test (Section 3.1.1). Desert background surfaces are identified using the Geographic Location database (Section 2.2.5) and the Desert Background Test (Section 3.1.2). Snow/ice background surfaces are identified using the Snow and Ice Location database (Section 2.2.4) and the spectral Snow/Ice Cover Background Test (Section 3.1.3). The sensor data are normalized using Eq. 14 to remove the effects of anisotropic reflection and then compared to the corresponding VBC database value for land surfaces or to a fixed empirically defined limit for water backgrounds. If the satellite-measured albedo exceeds the expected clear-scene background value by an amount greater than an empirically defined threshold then the pixel is classified as cloudy. Separate thresholds are used for land and water backgrounds. To minimize the background surface signal in the satellite data, sensor channel selection is also a function of background surface type. Over land, channel 1 sensor albedo data are used, while over water channel 2 data are used. The Visible Brightness Test is defined as:

$$\begin{aligned} \text{If land} & \quad \bullet A_1 * \sec(\theta) - VBC > THRESH_{\text{land}}, & \quad (\text{Over Land}) \\ \text{else if water} & \quad \bullet A_2 * \sec(\theta) > THRESH_{\text{water}}, & \quad (\text{Over Water}) \end{aligned}$$

where θ is the solar zenith angle, A_1 and A_2 are the ARF corrected channel 1 and 2 albedo respectively, VBC is the corresponding surface albedo value from the Visible Background Count database (Section 2.2.2) and $THRESH_{\text{land}}$ and $THRESH_{\text{water}}$ are the visible brightness cloud detection thresholds over land and water background surface types, respectively.

3.2.3 Night Condition Cloud Tests

There are two cloud tests used during night conditions that can only be executed in the absence of solar illumination. Solar zenith angle information is used to define when night conditions exist. The AVHRR Cloud Analysis Algorithm defines night as conditions when the solar zenith angle is greater than 90° .

3.2.3.1 Fog, Low Stratus Test

The Fog, Low Stratus Test exploits the fact that at night measured channel 3 radiance is composed solely of an emitted component and that cloud emissivity at $3.7 \mu\text{m}$ is generally lower than the $10.7 \mu\text{m}$ LWIR emissivity for water droplet clouds. A cloud decision is made by comparing the T_4 value to the T_3 value. If the T_3 is lower than T_4 by an amount greater than a cloud detection threshold then the pixel is classified as cloudy. A separate cloud detection threshold is maintained for areas identified as desert background by the Geographic Type database (Section 2.2.5). The Fog, Low Stratus Test is defined as:

$$\bullet T_4 - T_3 > THRESH_{\text{fls}},$$

where $THRESH_{\text{fls}}$ is a background surface-dependent cloud detection threshold.

3.4 CLOUD DETERMINATION

Final classification of pixels as either cloud-filled or cloud-free is performed by evaluating the results of the individual AVHRR cloud tests following the procedure illustrated in Fig. 9. For nighttime situations, defined by solar zenith angles greater than 90°, the process is straightforward. If any nighttime test detects cloud, the pixel is classified as cloud-filled. These nighttime tests are the Fog, Low Stratus Test (Section 3.2.1.1) and Nighttime Thin Cirrus Cloud Test (Section 3.2.1.2).

During daytime conditions, the process for evaluating the individual cloud test results is more complex. Several of the cloud tests rely on reflected solar radiation and thus can be confused by highly reflective terrestrial backgrounds such as sun glint, snow/ice cover, and desert. These problematic backgrounds may degrade the accuracy of the AVHRR Cloud Analysis Algorithm if they are mistakenly identified as cloud. To avoid this problem individual test results that classify pixels as cloud-filled are not used to generate the final cloud product when it is likely they are erroneous due to problematic surface backgrounds. The background filter tests described in Section 3.1, and the Geographic Type database described in Section 2.2.5, are employed to filter these backgrounds from the final cloud results. Filtering is achieved by negating the results for a particular cloud test if an appropriate background filter flag has also been triggered for the pixel being evaluated. Table 8 provides a look-up matrix of the filters employed by each of the nine individual AVHRR cloud detection tests.

Table 8. Background Surface Filters for Cloud Tests

	Spectral Test Filters			Supporting Data Filters		
	Sun Glint	Snow/Ice	Desert	Snow/Ice AFGWC	Geographic Desert	Geographic Coast
Low Cloud and Fog		✓				
Precipitating Cloud		✓				
Daytime Thin Cirrus Cloud		✓				
Visible Brightness Ratio	✓	✓	✓	✓	✓	✓
Visible Brightness	✓	✓	✓	✓	✓	
Cold Cloud		✓ ¹				
Cirrus Cloud		✓ ¹				
Fog, Low Stratus						
Nighttime Thin Cirrus Cloud						

¹ During daytime conditions

In addition to background surface filters, several tests use cloud detection thresholds designed to be more restrictive over problematic background surfaces. All restrictions that must be considered when analyzing results of the individual cloud tests to produce a final cloud classification are summarized below. Included in these descriptions are solar zenith angle requirements and conditions under which the results of the individual cloud tests are not used (filtered) due to problematic background surfaces.

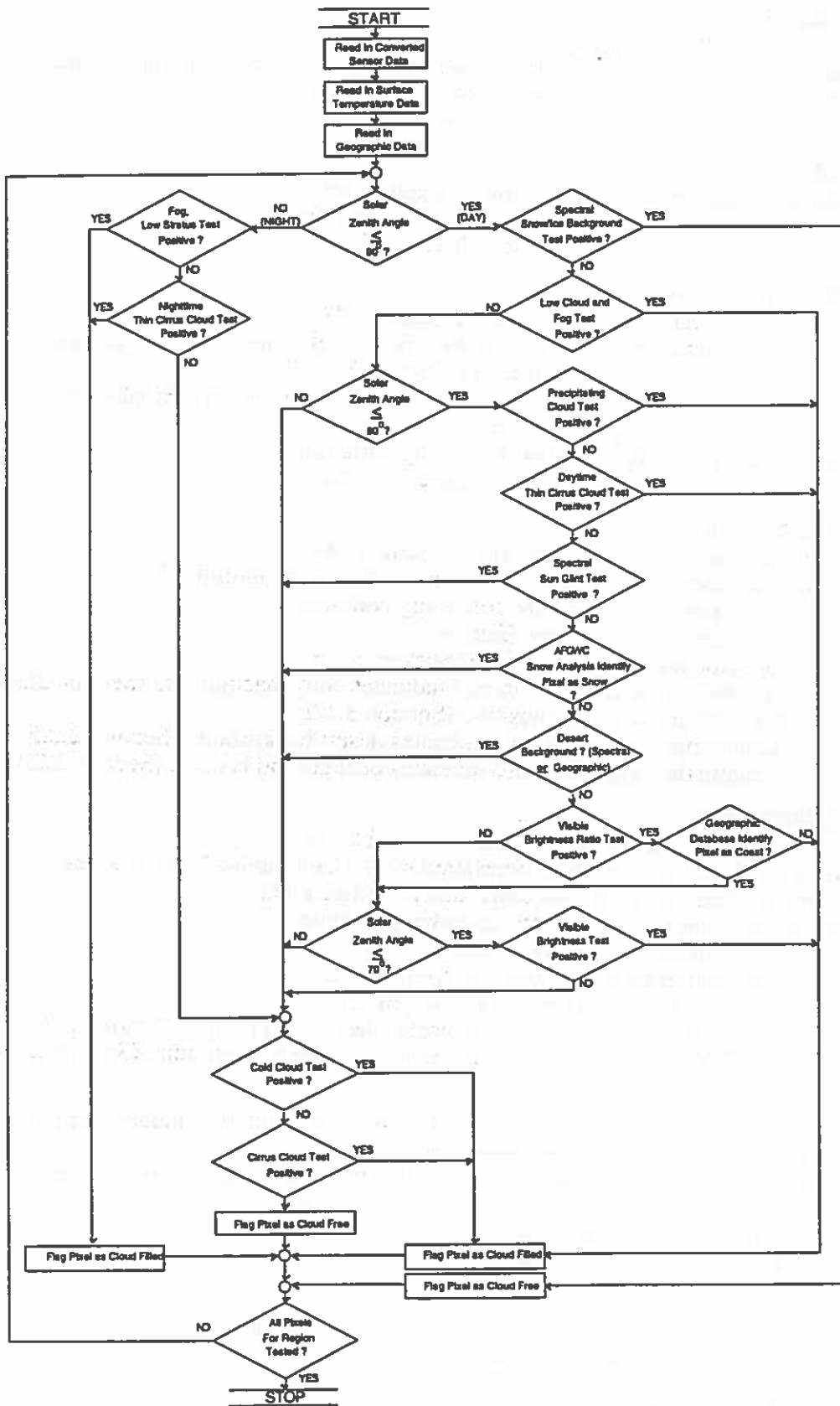


Figure 9. AVHRR Cloud Classification Procedure

Low Cloud and Fog Test

- Solar zenith angle must be less than or equal to 90°
- Desert and sun glint backgrounds require stricter cloud detection thresholds
- Test result is not used under the following conditions:
 - Spectral snow test is positive (Section 3.1.3)

Precipitating Cloud Test

- Solar zenith angle must be less than or equal to 80°
- Test result is not used under the following conditions:
 - Spectral snow test is positive (Section 3.1.3)

Daytime Thin Cirrus Cloud Test

- Solar zenith angle must be less than or equal to 80°
- Snow backgrounds identified by the AFGWC snow analysis model (Section 2.2.4) require an additional criterion (Section 3.2.1.2)
- Water and land backgrounds, identified by the Geographic Type database (Section 2.2.5), use separate cloud detection thresholds
- Test result is not used under the following conditions:
 - Spectral snow test is positive (Section 3.1.3)

Visible Brightness Ratio Test

- Solar zenith angle must be less than or equal to 80°
- Cloud detection threshold maintained as a function of humidity
- Test result is not used under the following conditions:
 - Sun glint test is positive (Section 3.1.1)
 - Spectral snow test is positive (Section 3.1.3)
 - AFGWC snow analysis model indicates snow background (Section 2.2.4)
 - Spectral desert test is positive (Section 3.1.2)
 - Geographic Type database indicates desert background (Section 2.2.5)
 - Geographic Type database indicates background is coast (Section 2.2.5)

Visible Brightness Test

- Solar zenith angle must be less than or equal to 70°
- Water and land backgrounds, identified by the Geographic Type database (Section 2.2.5), use separate cloud detection thresholds
- Test result is not used under the following conditions:
 - Sun glint test is positive (Section 3.1.1)
 - Spectral desert test is positive (Section 3.1.2)
 - Spectral snow test is positive (Section 3.1.3)
 - Geographic Type database indicates desert background (Section 2.2.5)
 - AFGWC snow analysis model indicates snow background (Section 2.2.4)

Cold Cloud Test

- Snow backgrounds identified by the AFGWC snow analysis model (Section 2.2.4) require a separate cloud detection threshold
- Water, land, coast, and desert, identified by the Geographic Type database (Section 2.2.5), use separate cloud detection thresholds
- Test result is not used under the following conditions:
 - Spectral snow test is positive (Section 3.1.3)

Cirrus Cloud Test

- Snow backgrounds identified by the AFGWC snow analysis model (Section 2.2.4) require an additional criterion (Section 3.2.1.2)
- Test result is not used under the following conditions:
 - Spectral snow test is positive (Section 3.1.3)

Fog, Low Stratus Test

- Solar zenith angle must be greater than 90°
- Desert backgrounds identified by Geographic Type database (Section 2.2.5) require a stricter cloud detection threshold

Nighttime Thin Cirrus Cloud Test

- Solar zenith angle must be greater than 90°
- Cloud detection test channel combination selected as a function of humidity

3.5 CONFIDENCE FLAG DETERMINATION

Along with the cloud analysis, the AVHRR algorithm produces information on the expected accuracy of the analysis for each pixel. Since the accuracy information is designed to provide a relative indication of how much confidence an end user can place in the analysis, the output product is termed a confidence flag. Three levels of confidence are provided: LOW, MIDDLE, and HIGH. The level of confidence assigned to each pixel is based on the strength of the cloud signature relative to the cloud threshold level for each cloud test. Signature strength is defined in terms of quanta where a quanta is based on the magnitude of the cloud threshold associated with each test. A numeric value representing the relative strength of the cloud signature in quanta is calculated for each test based on the magnitude of the cloud signal. Table 9 provides the quanta assignments for each cloud test. The convention used is that positive numbers indicate cloud-filled pixels and negative numbers indicate cloud-free pixels.

Quanta size is uniquely defined as fixed values for each cloud test with the exception of the Cirrus Cloud Test. The quanta size for the Cirrus Cloud Test is maintained as a function of the cirrus cloud threshold interpolated from Table A-2a in Appendix A. The quanta size (n) for the Cirrus Cloud Test is calculated as:

$$n = \text{Cirrus Cloud Threshold} / 2 .$$

Thus, if the Cirrus Cloud Test threshold is determined to be 7.00 then the quanta size would be set to 3.5.

For example, if the Cold Cloud Test (Section 3.2.1.1) measured a $T_{\text{pred}} - T_4$ difference of 16.3 K and the cloud threshold ($\text{THRESH}_{\text{cold}}$) were 10, then the strength of cloud signature for that test would be 2 quanta:

$$\text{diff} = 6.3 \quad (\text{i.e., } 16.3 - 10)$$

and
$$5.0 < \text{diff} \leq 10.0$$

therefore from Table 9: Quanta Magnitude = 2 and Confidence Level = MIDDLE

The procedure to assign a confidence flag to each pixel is to use the confidence level associated with the test that exhibits the strongest spectral signature. When performing the confidence flag determination, cloudy signatures always take precedence over clear signatures. Thus, if one cloud test detected cloud with a quanta magnitude of 1 and another test produced a cloud-free classification with a quanta magnitude of -2, the positive cloud result would take precedence and the pixel would be classified as cloud-filled with LOW confidence.

Table 9. AVHRR Quanta Value Classification Assignments

Cloud Test Name	Quanta Size	Spectral Signature Departure From Threshold (diff)	Quanta Magnitude	Confidence Level
Cold Cloud	5.0	$0 < \text{diff} \leq 5.0$	+1	LOW/Cloud-Filled
		$5.0 < \text{diff} \leq 10.0$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 10.0$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -5.0$	-1	LOW/Cloud-Free
		$-5.0 > \text{diff} \geq -10.0$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -10.0$	< -2	HIGH/Cloud-Free
Cirrus Cloud <i>n</i> = Cirrus Cloud Threshold / 2	<i>n</i>	$0 < \text{diff} \leq n$	+1	LOW/Cloud-Filled
		$n < \text{diff} \leq 2n$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 2n$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -n$	-1	LOW/Cloud-Free
		$-n > \text{diff} \geq -2n$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -2n$	< -2	HIGH/Cloud-Free
Low Cloud and Fog	6.0	$0 < \text{diff} \leq 6.0$	+1	LOW/Cloud-Filled
		$6.0 < \text{diff} \leq 12.0$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 12.0$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -6.0$	-1	LOW/Cloud-Free
		$-6.0 > \text{diff} \geq -12.0$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -12.0$	< -2	HIGH/Cloud-Free
Precipitating Cloud	0.05	$0 < \text{diff} \leq 0.05$	+1	LOW/Cloud-Filled
		$0.05 < \text{diff} \leq 0.10$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 0.10$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -0.05$	-1	LOW/Cloud-Free
		$-0.05 > \text{diff} \geq -0.10$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -0.10$	< -2	HIGH/Cloud-Free
Daytime Thin Cirrus	0.05	$0 < \text{diff} \leq 0.05$	+1	LOW/Cloud-Filled
		$0.05 < \text{diff} \leq 0.10$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 0.10$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -0.05$	-1	LOW/Cloud-Free
		$-0.05 > \text{diff} \geq -0.10$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -0.10$	< -2	HIGH/Cloud-Free
Visible Brightness Ratio	0.075	$0 < \text{diff} \leq 0.075$	+1	LOW/Cloud-Filled
		$0.075 < \text{diff} \leq 0.15$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 0.15$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -0.075$	-1	LOW/Cloud-Free
		$-0.075 > \text{diff} \geq -0.15$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -0.15$	< -2	HIGH/Cloud-Free
Visible Brightness	0.03	$0 < \text{diff} \leq 0.03$	+1	LOW/Cloud-Filled
		$0.03 < \text{diff} \leq 0.06$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 0.06$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -0.03$	-1	LOW/Cloud-Free
		$-0.03 > \text{diff} \geq -0.06$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -0.06$	< -2	HIGH/Cloud-Free
Fog, Low Stratus	0.75	$0 < \text{diff} \leq 0.75$	+1	LOW/Cloud-Filled
		$0.75 < \text{diff} \leq 1.50$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 1.50$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -0.75$	-1	LOW/Cloud-Free
		$-0.75 > \text{diff} \geq -1.50$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -1.50$	< -2	HIGH/Cloud-Free
Nighttime Thin Cirrus	1.0	$0 < \text{diff} \leq 1.0$	+1	LOW/Cloud-Filled
		$1.0 < \text{diff} \leq 2.0$	+2	MIDDLE/Cloud-Filled
		$\text{diff} > 2.0$	>+2	HIGH/Cloud-Filled
		$0 \geq \text{diff} \geq -1.0$	-1	LOW/Cloud-Free
		$-1.0 > \text{diff} \geq -2.0$	-2	MIDDLE/Cloud-Free
		$\text{diff} < -2.0$	< -2	HIGH/Cloud-Free

3.6 OUTPUT PRODUCT

The output product of the AVHRR Cloud Analysis Algorithm is an 8-bit quantity termed the Mask and Confidence Flag (MCF). The MCF is a bit-mapped quantity that stores cloud/no cloud information plus flags for specific types of cloud, missing data, and the confidence flag. Bit assignments for the AVHRR MCF cloud analysis algorithm output are provided in Table 10. One MCF is produced for each pixel in the input AVHRR image. Algorithm results are stored in an MCF file for subsequent use by the SERCAA Cloud Layering and Type algorithm and the Cloud Analysis Integration algorithm (see Fig. 1). MCF bit assignments are made as follows:

Table 10. AVHRR Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud
3	Precipitating Cloud	ON = Precipitating Cloud
4	Partial Cloud	Not Used By AVHRR Algorithm
5	Data Dropout	ON = Missing or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low; 2 = Middle; 3 = High
7	Flag	

Cloud Mask - Bit 0

The cloud mask bit is set to ON, indicating a cloud-filled pixel, if the pixel is determined to be cloud-filled by the final cloud classification (see Section 3.4).

Low Cloud - Bit 1

The low cloud bit is set if any of the following cloud tests are passed, as determined by the final cloud classification (see Section 3.4):

- Low Cloud and Fog Test (Section 3.2.2.1) ,
- Visible Brightness Ratio Test (Section 3.2.2.4) ,
- Visible Brightness Test (Section 3.2.2.5) ,
- Fog, Low Stratus Test (Section 3.2.3.1) .

Thin Cirrus Cloud - Bit 2

The thin cirrus cloud bit is set only if any of the following cloud tests are passed, as determined by the final cloud classification (see Section 3.4):

- Daytime Thin Cirrus Cloud Test (Section 3.2.2.3) ,
- Nighttime Thin Cirrus Cloud Test (Section 3.2.3.2) ,

and none of the following cloud tests are passed for the same pixel:

- Low Cloud and Fog Test (Section 3.2.2.1) ,
- Precipitating Cloud Test (Section 3.2.2.2) ,
- Visible Brightness Ratio Test (Section 3.2.2.4) ,
- Visible Brightness Test (Section 3.2.2.5) ,
- Fog, Low Stratus Test (Section 3.2.3.1) .

Precipitating Cloud - Bit 3

The precipitating cloud bit is set if any of the following cloud tests are passed, as determined by the final cloud classification (see Section 3.4):

- Precipitating Cloud Test (Section 3.2.2.2) .

Partial Cloud - Bit 4

The partial cloud bit is not used by the AVHRR Cloud Analysis Algorithm.

Data Dropout - Bit 5

The data dropout bit is set if the data for the pixel is either missing or unreliable.

Confidence Flag - Bits 6 & 7

The confidence flag bits are set to indicate LOW (1), MIDDLE (2), or HIGH (3) confidence as detailed in Section 3.5.

4. DMSP CLOUD ANALYSIS ALGORITHM DESCRIPTION

The SERCAA cloud analysis algorithm for analysis of Defense Meteorological Satellite Program (DMSP) polar orbiting satellite data consists of a single channel thermal infrared algorithm and a combined visible/infrared bispectral algorithm. These algorithms follow the approach of Gustafson and d'Entremont (1992) developed for the TACNEPH program. Both are statistical, dual threshold type algorithms designed to classify pixels as either cloud-filled, cloud-free and partially cloud-filled. Selection of whether the single channel IR or the two channel visible/IR algorithm is used is dependent on the amount of solar illumination present within the scene. A solar zenith angle threshold of 75° ($THRESH_{DMSP_solzen}$) is currently used to make this determination. Visible data collected when the solar zenith angle is greater than this threshold are not used by the algorithm due to uncertainties in the data introduced by gain control adjustments performed on-board the satellite during scan. Gain adjustments change the relationship between visible count and the apparent brightness of the surface and thus make quantitative analysis of the data problematic. The affect of gain adjustments on visible count at solar zenith angles less than 75° are relatively small compared to changes between cloud and terrestrial background and, as such, are ignored by the algorithm. However, as the scan approaches the terminator the effect of the gain adjustment becomes large. Under daytime conditions, when both visible and infrared data are available, a combination of the single channel IR algorithm and the two channel bispectral algorithm is used. During nighttime conditions the single channel functional flow diagram of the DMSP Operational Linescan System (OLS) cloud analysis approach.

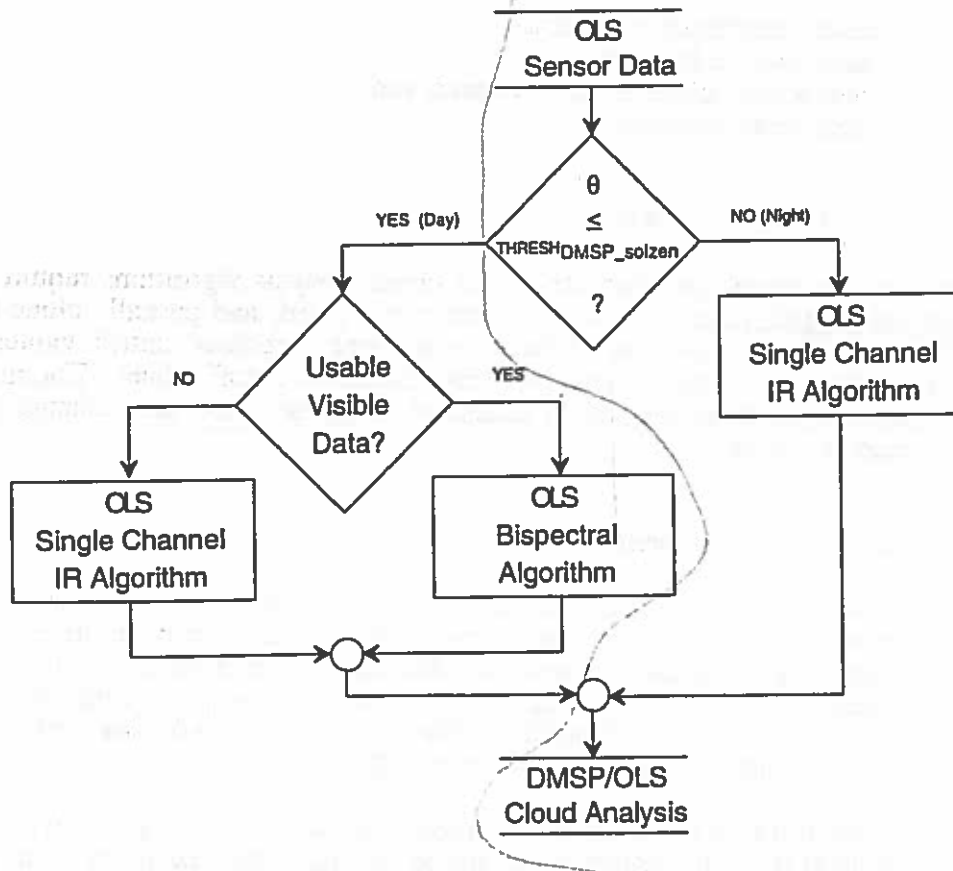


Figure 10. DMSP/OLS Cloud Analysis Algorithm Approach

Both the single channel and bispectral algorithms are designed to perform a cloud classification on each pixel within an analysis scene. An analysis scene is defined as some portion of a DMSP orbit. The size of the analysis scene is implementation dependent and will be affected by factors such as data ingest schedule, computer memory and processing resources, and time available to perform the analysis. During SERCAA, the maximum analysis scene size was a quarter orbit of 2.7 km OLS data. Processing of pixels within the analysis scene is performed by first dividing the scene into analysis boxes whose size is also implementation dependent. The SERCAA analysis box size was set at 16 x 16 pixels, however as explained in Section 2.2.1.1, this is considered the minimum size required to select a reliable reference pixel in the calculation of the predicted clear scene brightness temperature and can be increased if necessary to improve processing performance. Execution of the algorithm is performed by analyzing data on a per analysis box basis.

The threshold approach utilized by the OLS was selected because it allows for multiple uncertainties in the sensor measurements, including sensor calibration, clear scene characteristics, and atmospheric transmission, to be accounted for with a single value. An empirically derived dynamic correction factor is used to account for all sources of error collectively without the need to understand and quantify the individual contributions (see Section 2.2.1.1). The magnitude of the cloud thresholds is then dictated by the remaining uncertainty in the corrected temperatures. The performance of the algorithm is directly impacted by the ability to accurately characterize cloud-free backgrounds. This is achieved through the identification or characterization of the following:

- land/water/desert boundaries,
- snow and ice location,
- clear scene brightness temperature, and
- clear scene reflectance.

4.1 THRESHOLD CALCULATION

Both single channel and bispectral OLS cloud analysis algorithms require dual infrared threshold cutoff values to classify clear, cloud-filled, and partially cloud-filled pixels within an analysis box. In addition to infrared threshold cutoff values, the bispectral algorithm also requires a pair of visible threshold cutoff values. The method employed to select these threshold values is dependent on the sensor data channel being analyzed, infrared or visible.

4.1.1 Infrared Channel Thresholds

Infrared thresholds are based on an estimate of the clear scene brightness temperature derived from a dynamic correction to surface skin temperature estimates obtained from the AFGWC surface temperature database. Once a predicted clear scene brightness temperature is established, threshold values are computed from statistical estimates of the expected natural variability of the data. The procedure to predict clear scene brightness temperatures is detailed in Section 2.2.1.1.

Clear scene infrared statistics information is used both to establish whether a given reference pixel is cloud-contaminated and to calculate the magnitude of the clear and cloud threshold values. Thresholds are used to account for the uncertainty in the

predicted clear scene brightness temperature calculation. The cloud threshold is defined by the equation:

$$\Delta T_{\text{cld}} = |\Delta T_{\text{max}} - \Delta T_{\text{min}}| * \alpha_{\text{cld}}^{1.5}, \quad (15)$$

and the clear threshold is defined by:

$$\Delta T_{\text{clr}} = |\Delta T_{\text{max}} - \Delta T_{\text{min}}| * \alpha_{\text{clr}}^{.75}, \quad (16)$$

where ΔT_{max} and ΔT_{min} are computed from the IR-Skin Temperature Statistics internal database as described in Section 2.2.1.1 and used to represent the natural variability of the difference between the satellite observed and predicted temperature for the cloud-free background. The values of α_{cld} and α_{clr} are provided in Table A-3 in Appendix A.

Once the thresholds are calculated then cloud and clear cutoff values used in the analysis algorithms (see following sections) are calculated by subtracting the threshold values from the predicted clear scene brightness temperature calculated from Eq. 11 or 12. Thus, the cloud cutoff is defined as:

$$T_{\text{cld}} = T_{\text{pred}} - \Delta T_{\text{cld}}, \quad (17)$$

and the clear cutoff value as:

$$T_{\text{clr}} = T_{\text{pred}} - \Delta T_{\text{clr}}. \quad (18)$$

Calculation and application of the cutoff values is illustrated graphically in Fig. 11.

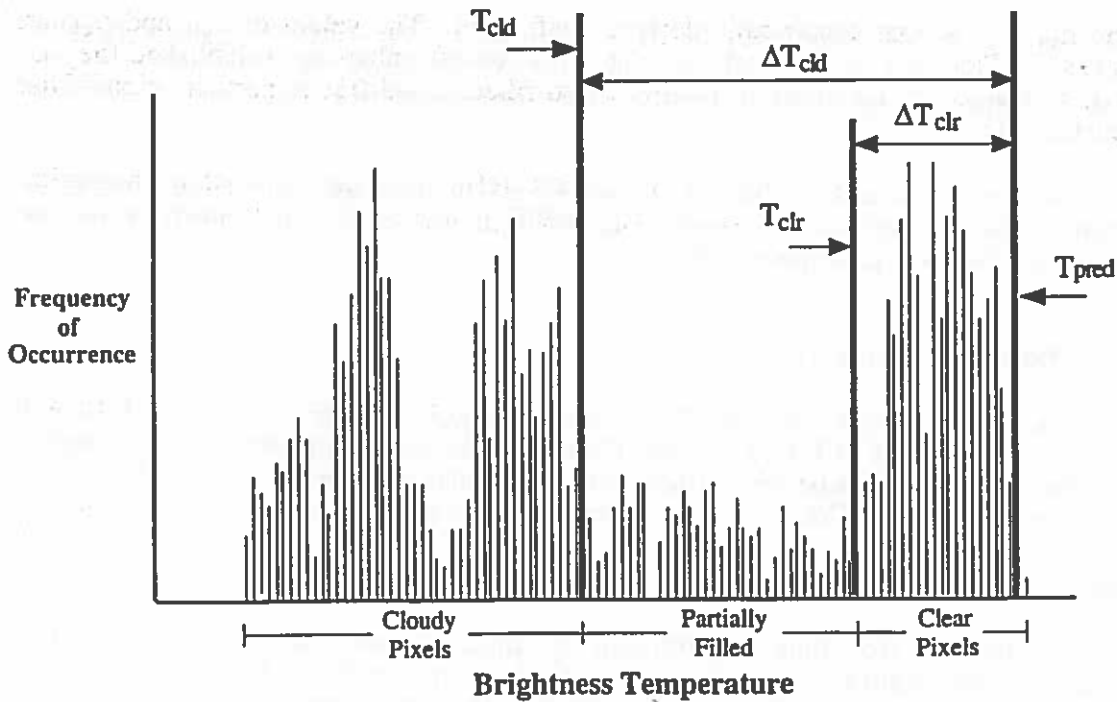


Figure 11. IR Single Channel Test Dual Threshold Classification Approach

4.1.2 Visible Channel Thresholds

The procedure for establishing cutoff values for the visible channel follow an approach similar to that for the infrared channel. Fundamental differences are that cloud and clear cutoff values are calculated on a pixel-by-pixel basis rather than over an analysis box, and the technique used to establish cutoff values varies depending on the geographic type of the scene. Over land areas, where surface reflectance is expected to change with time and geographic region, cutoff values are set with the use of the dynamically maintained Visible Background Count internal database described in Section 2.2.2. Data from the Visible Background database are used in way analogous to the way clear scene brightness temperature data are used in the infrared technique. However, since they are generated directly from the satellite observed radiances no correction factors need to be applied.

If the pixel being analyzed is located over land, then the method for determining the cloud cutoff threshold value (R_{cld}) is defined by the equation:

$$R_{cld} = R_{sfc} * \rho_{cld} , \quad (19)$$

where R_{sfc} is the brightness count from the Visible Background database that corresponds to the satellite pixel and ρ_{cld} is an empirically derived coefficient used to account for uncertainty in the background database. Note that the uncertainty coefficient is multiplied by the background value rather than added as in the IR cutoff calculation (Eq. 17). This is to account for increasing uncertainty as the value (i.e., brightness) of the background increases. Similarly, the method for determining the clear cutoff value (R_{clr}) is defined by the equation:

$$R_{clr} = R_{sfc} * \rho_{clr} , \quad (20)$$

where ρ_{cld} is a second empirically derived coefficient. The values of ρ_{clr} and ρ_{cld} are provided in Table A-3 in Appendix A. Once the cutoff values are established they are used in the bispectral algorithm to classify cloud-filled, cloud-free or partially cloud-filled pixels (Section 4.3).

Over water, where variations in surface reflectance are considered negligible compared to land, visible cutoff values (R_{cld} and R_{clr}) are fixed. These cutoff values are provided in Table A-3 in Appendix A.

4.2 SINGLE CHANNEL TEST

The DMSP single channel cloud analysis algorithm utilizes a dual threshold approach as illustrated in Fig. 11. Separate cutoff thresholds are defined to segregate pixels classified as partially cloud-filled from those that are completely cloud-filled or completely cloud-free. Cloud analysis accuracy is dependent on the accurate prediction of the clear-scene brightness temperature used to define the clear and cloudy cutoff values as described in Section 2.2.1.1.

A functional flow diagram outlining the single channel algorithm is provided in Fig. 12. This algorithm consists of two tests. First, a test is performed to determine if the brightness temperature of the IR channel is less than the cloud cutoff:

$$\bullet T_{IR} < T_{cld} ,$$

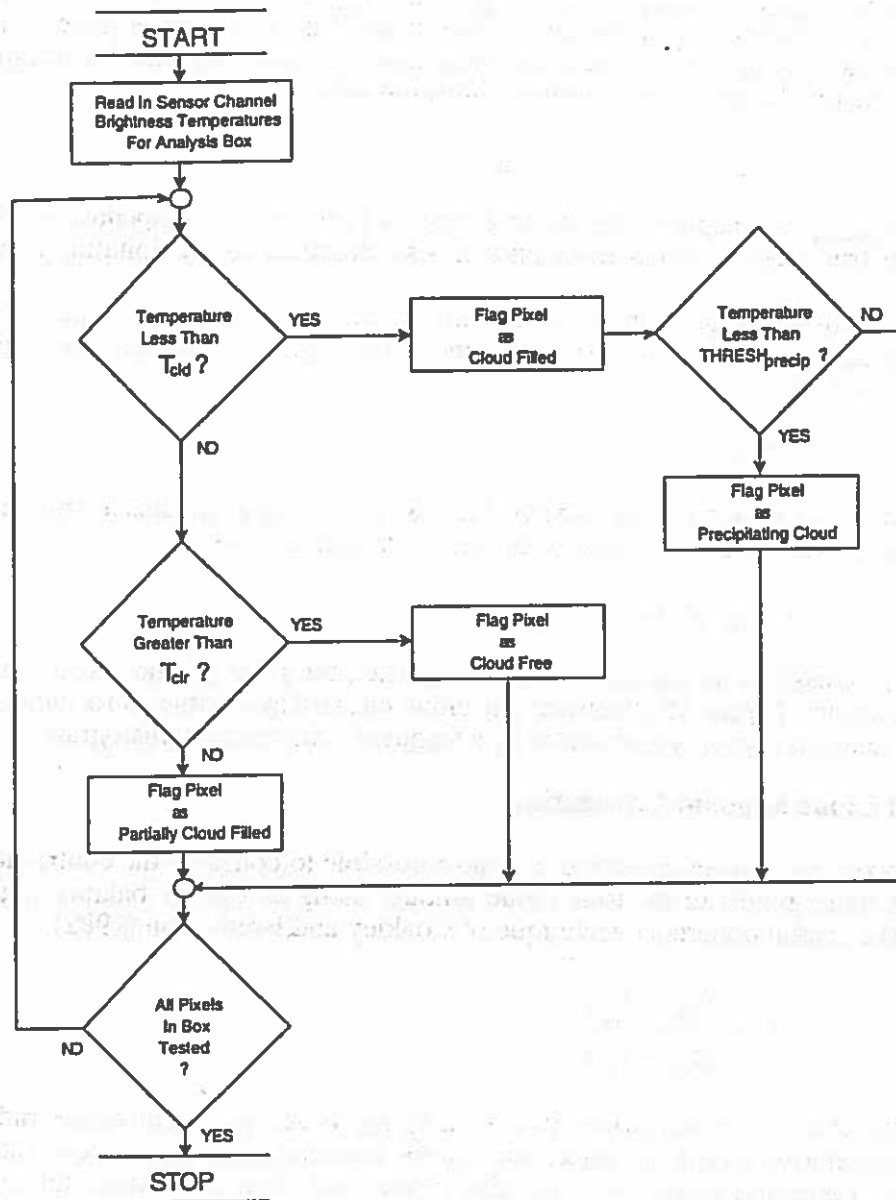


Figure 12. Single Channel Algorithm Functional Flow Diagram

where T_{IR} is the OLS infrared brightness temperature and T_{cld} is the cloud cutoff value defined by Eq. 17. Over snow or ice covered backgrounds, the measured brightness temperatures can vary significantly from the predicted clear-scene brightness temperature derived from the IR-Skin Temperature Statistics data. This is due, at least in part, to the rapid changes that can occur to the radiative characteristics of the Earth surface when new snow falls and as it melts and re-freezes. To account for the additional uncertainty in the predicted clear-scene brightness temperature used to derive T_{cld} , the magnitude of ΔT_{cld} and ΔT_{clr} used in Eqs. 17 and 18 is increased by 10% for backgrounds classified as snow or ice in the Snow and Ice Location supporting database.

If the OLS brightness temperature is less than T_{cld} then the pixel is classified as cloud-filled and a second test is performed to determine if the cloud-filled pixel is also a precipitating cloud. Precipitating clouds are identified by testing whether the brightness temperature of the pixel is less than a defined threshold value:

$$\bullet T_{IR} < THRESH_{precip} ,$$

where $THRESH_{precip}$ is a separate threshold define in Table A-3 in Appendix A. If this test evaluates as true then the cloud-filled pixel is also classified as precipitating cloud.

If the brightness temperature of the IR channel is not less than the cloud threshold (T_{cld}) then a second test is performed to determine if the brightness temperature is greater than the clear threshold:

$$\bullet T_{IR} > T_{clr} ,$$

where T_{clr} is the cloud-free cutoff defined by Eq. 18. If this test evaluates as true then the pixel is classified as cloud-free. If both of the tests evaluate as false:

$$\bullet T_{cld} \leq T_{IR} \leq T_{clr}$$

then the pixel is classified as partially cloud-filled (i.e., the FOV of the sensor contains both cloud and clear). Figure 12 illustrates the cloud classification criteria for pixels from a hypothetical analysis region accumulated in a frequency distribution histogram.

4.2.1 Partial Cloud Amount Calculation

Once pixels have been classified, it is also possible to compute the contribution of partially cloud-filled pixels to the total cloud amount using an energy balance approach adapted from the spatial coherence technique of Coakley and Bretherton (1982):

$$A_c = \frac{(I_{IR} - I_{clr})}{(I_{cld} - I_{clr})} \quad (21)$$

where A_c is the effective cloud cover ($0 \leq A_c \leq 1$), I_{IR} is the measured scene radiance, I_{cld} is the representative cloud radiance, and I_{clr} the representative clear scene radiance. I_{clr} and I_{cld} are computed from the respective cloudy and clear brightness temperature cutoff thresholds. Currently the SERCAA algorithms only use the clear, partially cloudy, and cloudy classification results from the Single Channel Test, information on partial cloud amount calculations are provided here as a possible future enhancement.

4.3 BISPECTRAL TEST

The OLS bispectral algorithm, developed for use during daytime conditions, is similar to the single channel algorithm but is applied in two spectral dimensions. Data from both visible and infrared sensor channels are analyzed simultaneously using two pairs of cutoff values, one pair for each channel. It should be noted that accurate specification of the infrared threshold is not as critical in the bispectral test as in the one channel algorithm since low (warm) liquid water clouds reflect well and will generally be detected from the visible data when not over highly reflective backgrounds.

Figure 13 provides an illustration of how the two dimensional visible-infrared space is divided into nine classification regions by the cutoff values. In this figure T_{cld} and T_{clr} represent the infrared brightness temperature cloud and clear cutoff values defined by Eqs. 17 and 18 respectively, while R_{cld} and R_{clr} represent the visible count cloud and clear cutoffs defined by Eqs. 19 and 20 respectively. Infrared temperatures that are less than the infrared cloud threshold value:

- $T_{IR} < T_{cld}$

are unambiguously classified as cloud-filled over backgrounds that are free of snow and ice. Data that are both warm in the infrared and dark in the visible channel (R_{vis}):

- $T_{IR} > T_{cld}$
- and
- $R_{vis} < R_{cld}$
- and
- $T_{IR} > T_{clr}$ or $R_{vis} < R_{clr}$

are unambiguously classified as clear. Warm bright regions:

- $T_{IR} > T_{cld}$
- and
- $R_{vis} > R_{cld}$

require an a priori clear scene classification to remove the ambiguity caused by the similarity in radiative signatures of backgrounds such as snow fields, deserts and low cloud. Data that fall between all four threshold values:

- $T_{cld} < T_{IR} < T_{clr}$
- and
- $R_{cld} > R_{vis} > R_{clr}$

are classified as partially cloud-filled. These rules are used to define the classification bins labeled in Fig. 13.

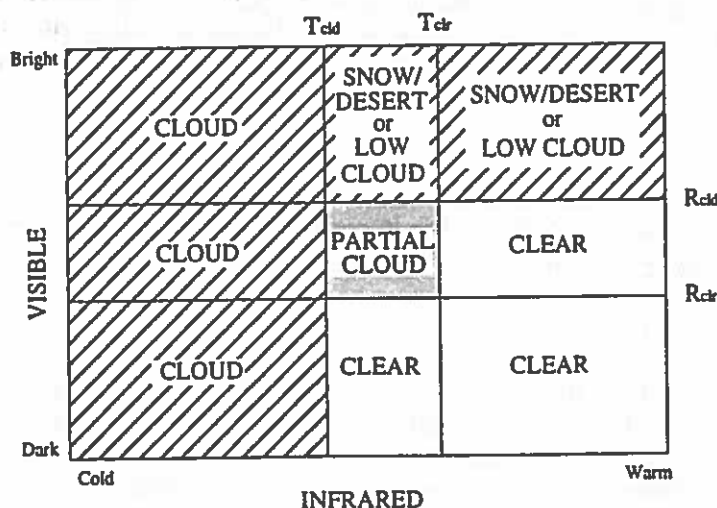


Figure 13. Bispectral Classification Approach

A functional flow diagram of the bispectral algorithm is provided in Fig. 14. The algorithm consists of several tests. Table A-3 of Appendix A lists the value of all thresholds used in the Bispectral algorithm tests.

Since visible data are used in addition to IR data in the Bispectral algorithm, a sun glint test is required over water background surfaces. Background surface type information is provided by the Geographic Type supporting database described in Section 2.2.5. Potential sun glint areas are defined by the following background surface and solar/satellite geometry tests:

- Background surface type must be water ,
- and* • $|\psi - \theta| < \text{THRESH}_{\text{zenith}}$,
- and* • $\text{THRESH}_{\text{loazimuth}} < \phi < \text{THRESH}_{\text{upazimuth}}$,

where $\text{THRESH}_{\text{upazimuth}}$, $\text{THRESH}_{\text{loazimuth}}$ and $\text{THRESH}_{\text{zenith}}$ are empirically derived threshold values that define the geographic extent over which sun glint may be expected to occur (see Section 2.2.6 for angle definitions). Note the OLS sun glint test differs from the AVHRR test described in Section 3.1.1 in that only sun-satellite geometry relationships are used, no additional spectral tests are available. If any pixel within an analysis box is determined to be located within the potential sun glint area then further processing of visible data for all pixels within that box is terminated and the infrared data are processed using the Single Channel algorithm described in Section 4.2. Bispectral processing then continues with the next analysis box.

Similarly, visible data are not processed over reflective backgrounds of desert or snow (as defined by the Geographic Type and Snow and Ice Location supporting databases respectively). If the test area is located over a snow/ice field or desert region then the single channel infrared test alone is used to classify the pixel. Thus over snow fields, desert regions, or water backgrounds that support sun glint, the OLS algorithm is dependent on the infrared signature alone to detect low cloud.

If usable visible data remain following the background dependent tests described above, then the first spectral test performed by the bispectral algorithm is to determine whether the visible count from the satellite data is greater than the cloud threshold:

$$\bullet R_{\text{vis}} > R_{\text{cld}} .$$

If the test evaluates as true then the area in question is classified as cloud-filled. Next, the single channel infrared test is performed to determine if the brightness temperature of the infrared channel is lower than the cloud threshold:

$$\bullet T_{\text{IR}} < T_{\text{cld}} .$$

If the test evaluates as true then the pixel is classified as cloud-filled. Otherwise, the algorithm data flow continues to the final test to discriminate clear and partially cloud-filled pixels. If the visible channel count is less than the clear threshold or the brightness temperature of the infrared channel is greater than the clear threshold:

$$\bullet R_{\text{vis}} < R_{\text{clr}} ,$$

or

$$\bullet T_{\text{IR}} > T_{\text{clr}} .$$

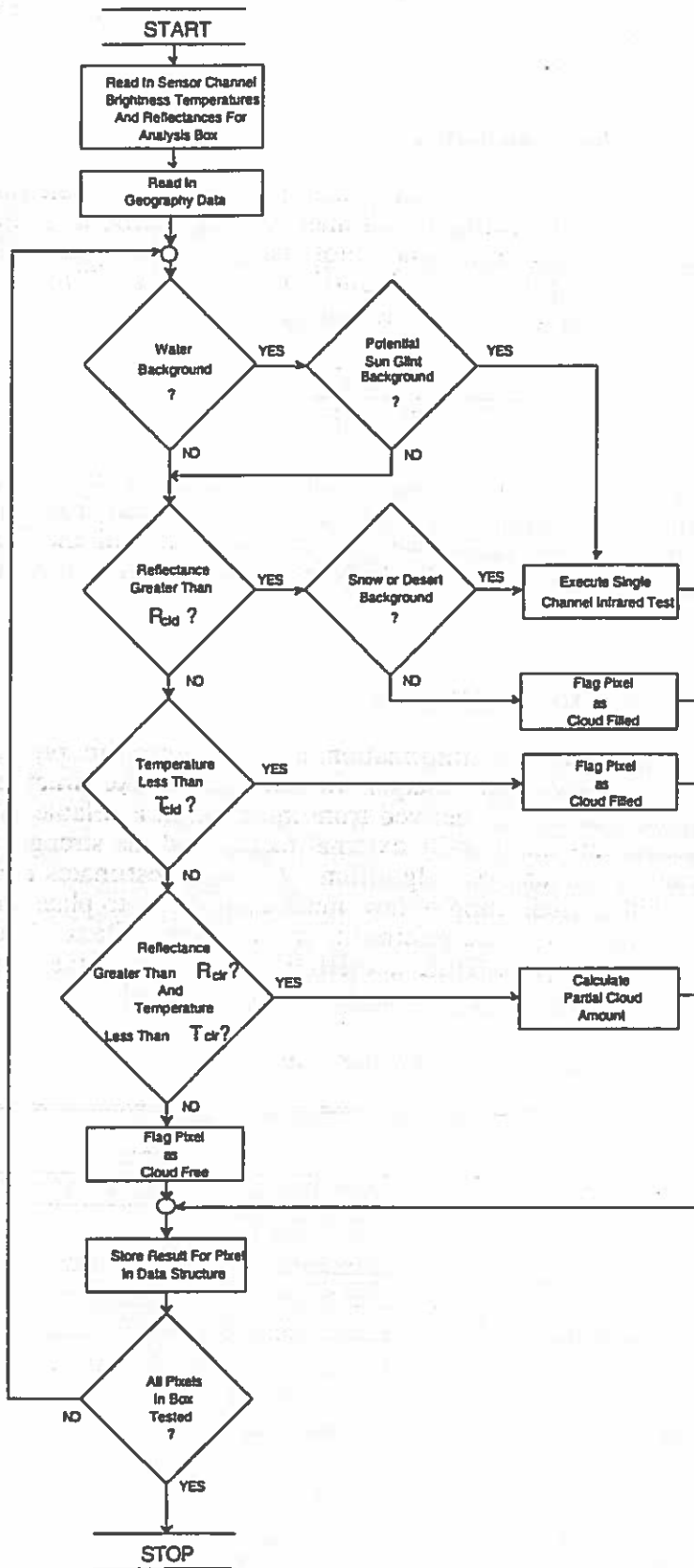


Figure 14. Bispectral Algorithm Functional Flow Diagram

If the test evaluates as true then the pixel is classified as cloud-free, otherwise, it is classified partially cloud-filled.

4.3.1 Partial Cloud Amount Calculation

Partial cloud contribution to cloud fraction can also be calculated for the Bispectral algorithm. The cloud fraction is assumed to be proportional to the distance a data point lies between the clear and cloud cutoff values in the space defined by the intersection of the four cutoff levels identified as Partial Cloud in Fig. 13. Mathematically, effective cloud cover A_c is defined as:

$$A_c = \frac{1}{2} \left(\frac{I_{IR} - I_{clr}}{I_{cld} - I_{clr}} + \frac{R_{vis} - R_{clr}}{R_{cld} - R_{clr}} \right), \quad (22)$$

where R_{vis} and I_{IR} are the measured reflectance and calculated radiance, respectively, of the partially cloud-filled data point. I_{clr} and I_{cld} are calculated from the brightness temperature thresholds, T_{clr} and T_{cld} respectively. However, as with the Single Channel algorithm this information is not used by the SERCAA algorithms and is provided here as a potential enhancement.

4.4 CONFIDENCE FLAG DETERMINATION

In addition to analyzed cloud information, the OLS algorithm provides information on the expected accuracy of the analysis for each pixel. Accuracy estimates are based on pixel attributes that can be derived from information available to the analysis algorithm such as constraints imposed by external factors and the strength of the cloud signature as measured by the analysis algorithm. Accuracy estimates are intended to provide the end user with an indication of how much confidence to place in the analysis for any given pixel and, as such, are referred to as confidence flags. Three levels of confidence are defined: LOW, MIDDLE, and HIGH. The twelve pixel attributes listed in Table 11 are used to establish the OLS analysis confidence level.

Table 11. Confidence Flag Criteria

Attribute	Source	Numeric Value
Snow/Ice Covered Background	AFGWC Snow Analysis Model	-2
Sun Glint Contamination	Geometry Tests	-1
Coast Background	Geographic Type Database	-1
Desert Background	Geographic Type Database	-1
Default Temperature Correction Used	Clear Scene Brightness Temperature	-1
Water Background	Geographic Type Database	+1
Land Background	Geographic Type Database	+1
Visible and IR Channels Available	Sensor Data	+1
Cloud and Within 15° K of T_{cld}	Cloud Algorithm	+1
Cloud and Within 10° K of T_{cld}	Cloud Algorithm	+1
Cloud and Within 5° K of T_{cld}	Cloud Algorithm	+1
Cloud and Within 3° K of T_{cld}	Cloud Algorithm	+1

A numeric value is assigned to each identifiable pixel attribute that affects confidence in the analysis (see Table 11). Note that pixels located over problematic background surface conditions (i.e., snow/ice, sun glint, coast, desert) are assumed to be more difficult to analyze and, as such, are assigned a negative value. Similarly, pixels located in an analysis box that require a default temperature correction in the calculation of the predicted clear scene brightness temperature (Eq. 11) are considered to be more suspect than those that did not use a default correction and also carry a negative value.

The numeric value is positive for attributes felt to improve the analysis accuracy. This includes cases when the cloud analysis is performed over a straight land or water background rather than one of the problematic surfaces listed above and when both visible and IR channels are available to the algorithm for analysis. The strength of the cloud signature, measured as the departure of the IR brightness temperature or visible count from the respective cloud cutoff value, is also used as a measure of confidence in the analysis.

Confidence flag values for each pixel are established by initially assigning a numeric value associated with middle level confidence and then adjusting up or down based on the attributes that apply. A final confidence value is calculated by summing the numeric value associated with all applicable attributes. For example, if the algorithm established that a given pixel had the following attributes:

Initial confidence level of MIDDLE	7
Snow/Ice covered background	-2
Land background	+1
Visible and IR channels available	+1
Within 15° of clear or cloud threshold	+1
Within 10° of clear or cloud threshold	+1
Within 5° of clear or cloud threshold	+1
Total Value	<u>10</u>

then the final numeric value assigned to that pixel would be 10. Conversion to a confidence flag value of LOW, MIDDLE, or HIGH is performed by subjecting the numeric value to the thresholds defined in Table 12. Thus for the above example, the confidence flag assigned to the pixel has a value of HIGH since the total of 10 is greater than or equal to the HIGH confidence threshold of 9.

Table 12. DMSP Confidence Flag Assignment

Confidence Level Value	Confidence Flag
$0 \leq \text{Value} \leq 5$	LOW
$6 \leq \text{Value} \leq 8$	MIDDLE
$9 \leq \text{Value}$	HIGH

4.5 OUTPUT PRODUCT

The output product of the DMSP Cloud Analysis Algorithm is a bit-mapped 8-bit value, termed the Mask and Confidence Flag (MCF), that contains cloud information and associated confidence flag information. Table 13 provides definitions of the MCF bit assignments for the DMSP Cloud Analysis Algorithm output. Information provided by

the MCF includes: cloud/no-cloud, precipitating cloud, missing data, and confidence level information. As illustrated in Fig. 1, DMSP/OLS Cloud Analysis files are created for each OLS input scene processed through the OLS cloud analysis algorithm. These files contain one MCF value for each pixel in the input image. They are subsequently accessed as required input to the Cloud Typing and Layering and the Analysis Integration Algorithms.

Table 13. DMSP Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled ; OFF = Cloud-Free
1	Low Cloud	Not Used By DMSP Algorithm
2	Thin Cirrus Cloud	Not Used By DMSP Algorithm
3	Precipitating Cloud	ON = Precipitating Cloud
4	Partial Cloud	If ON Then Bit 0 = OFF
5	Data Dropout	ON = Missing Or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

MCF bits are set as follows:

Cloud Mask - Bit 0

The cloud mask bit is set to ON, indicating a cloud-filled pixel, if the pixel is determined to be completely cloud-filled.

Low Cloud - Bit 1

The low cloud bit is not used by the DMSP Cloud Analysis Algorithm.

Thin Cirrus Cloud - Bit 2

The thin cirrus cloud bit is not used by the DMSP Cloud Analysis Algorithm.

Precipitating Cloud - Bit 3

The precipitating cloud bit is set if the single channel test (Section 4.1) detects precipitating cloud.

Partial Cloud - Bit 4

The partial cloud bit is set when partial cloud is detected by either the single channel test (Section 4.1) or the bispectral test (Section 4.2). If Bit 4 is set then Bit 0 is clear.

Data Dropout - Bit 5

The data dropout bit is set if the data for the pixel is either missing or unreliable.

Confidence Flag - Bits 6 & 7

The confidence flag bits are set to indicate LOW (1), MIDDLE (2), or HIGH (3) confidence as detailed in Section 4.4.

5. GEOSTATIONARY CLOUD ANALYSIS ALGORITHM DESCRIPTION

The SERCAA cloud analysis algorithm for geostationary satellite platforms employs a hybrid approach to detect cloud cover. Separate temporal differencing, dynamic thresholding, and spectral discriminant tests are utilized in making a determination of whether pixels within an analysis scene are cloud-filled or cloud-free. Figure 15 provides a high level data flow diagram of the geostationary cloud algorithm illustrating that each of the three tests in the hybrid algorithm are implemented as a separate processing level. As the algorithm moves down through the three processing levels the cloud analysis becomes more complete. This implies that no one processing level alone is expected to identify all clouds within the analysis scene. Rather each level is designed to build on the results from the previous level by exploiting a different cloud signature. Thus the final cloud analysis is obtained by combining the results from all the individual tests contained in the three processing levels. Operationally, the algorithm is applicable to thermal infrared sensor data alone or in combination with visible data and other infrared channels when available. The algorithm is applicable to the following satellite systems:

- GOES (Geostationary Operational Environmental Satellite - USA)
- GMS (Geostationary Meteorological Satellite - Japan)
- METEOSAT (Meteorological Satellite - Europe)

The following sections provide detailed descriptions of the algorithm modules associated with each of the three types of tests.

5.1 TEMPORAL DIFFERENCE TEST

The first level of processing utilizes a temporal differencing technique to identify new cloud development and existing cloud features that have moved over either previously clear background or lower, warmer cloud. Processing is performed on a pixel-by-pixel basis for the entire analysis scene.

This technique is applicable to the visible or infrared channel individually or may be applied simultaneously to both, in a bispectral approach. Depending on the channel chosen, the test exploits the change in infrared brightness temperature and/or visible count caused by both moving and developing cloud features in collocated pixels taken from a pair of sequential satellite images. Cloud detection is performed by identifying pixels for which the satellite-observed brightness temperature decreases and/or the visible count increases by amounts greater than expected for clear-scene conditions over the time interval between the two images. Figure 16 illustrates this concept in both spectral dimensions.

It is of primary importance that the two sequential images be co-registered as accurately as possible before the temporal differencing algorithm is applied. Any registration errors that exist between the two images can result in anomalous cloud signatures. For example, along coastlines a water pixel in one image may be misregistered to a land pixel in the next. Since it is likely that both the visible counts and IR brightness temperatures measured from the two different backgrounds will vary significantly, the temporal difference algorithm is also likely to misclassify at least one as cloud. For the SERCAA program, co-registration of sequential geostationary images was

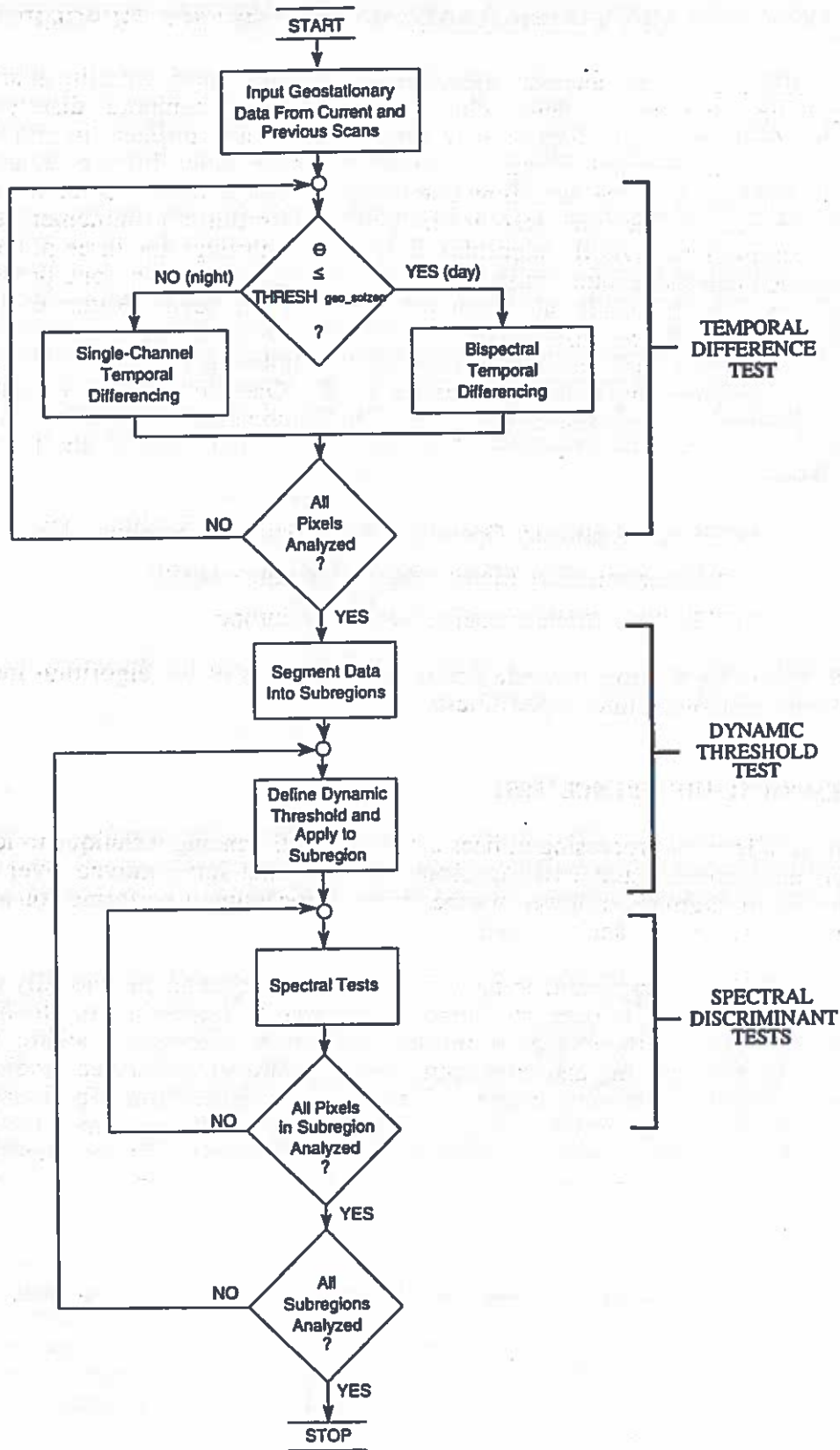


Figure 15. Geostationary Cloud Analysis Algorithm Functional Flow Diagram

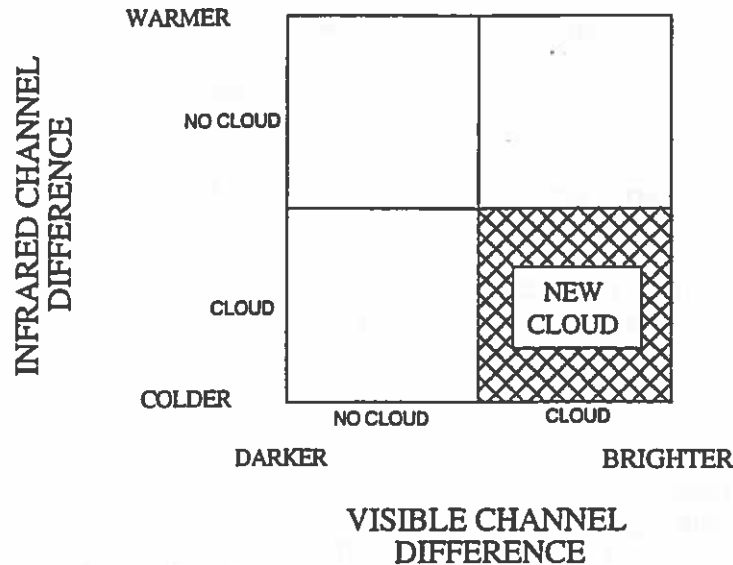


Figure 16. Visible and Infrared Temporal Difference New Cloud Algorithm Conceptual Approach

performed in an automated fashion using a two-step procedure. The first step identifies a single control latitude-longitude point within the two images, typically the satellite subpoint. Next, the satellite scan projection column and row numbers that correspond to the control point are computed for both images. If the two images are precisely collocated, these column and row numbers will match precisely from one time to the next. More typically, this is not the case. In this situation the difference between the respective control point column and row numbers serves as the offset by which one image is translated so that it lines up geographically with the other. This is performed individually for each pair of satellite images used as input to the temporal differencing technique.

Knowledge of the time rate-of-change of the satellite brightness temperature and visible count for the cloud-free background is required to define the temporal differencing cloud detection thresholds. Expected surface skin temperature changes are derived from the AFGWC Surface Temperature database described in Section 2.2.1. Changes in visible count are predicted using the Visible Background Count support database. VBC data are generated for each geostationary satellite based on actual satellite observations over a two-week period as described in Section 2.2.2.

During daytime conditions, when both visible and infrared sensor data are available, a bispectral temporal differencing technique is employed. At night, a one-channel version of the algorithm is used to analyze IR data alone (see Fig. 15). Day and night are defined in terms of the scene solar zenith angle:

$$\bullet \theta \leq \text{THRESH}_{\text{geo_solzen}} \bullet$$

where $\text{THRESH}_{\text{geo_solzen}}$ is the day/night cutoff threshold. The value of $\text{THRESH}_{\text{geo_solzen}}$ is provided in Table A-4. This technique makes a determination of cloud status by simultaneously examining the satellite-observed changes in infrared

brightness temperature and visible count. Figure 17 provides a schematic illustration of the visible and infrared data flow for the bispectral temporal differencing algorithm.

The first step in both the bispectral and single channel algorithms is to co-register the image data valid at the current analysis time (t) with the imagery obtained from the same satellite during the previous scan. Thus the valid time of the previous scan is $t - \Delta t$, where Δt is the time interval between consecutive scans. During SERCAA, geostationary satellite observations were available once per hour.

Both algorithms evaluate the change in the measured infrared brightness temperature that occurs between the two image times for collocated pixels. The change in brightness temperature is defined as:

$$\Delta T_{IR} = T_{IR}(t) - T_{IR}(t - \Delta t), \quad (23)$$

where $T_{IR}(t)$ represents the brightness temperature measured at the most recent observation time and $T_{IR}(t - \Delta t)$ represents the brightness temperature at the previous image time. The measured brightness temperature difference is compared to the expected change in temperature of the terrestrial surface to determine if cloud has formed or moved into the FOV during the intervening time period. To account for changes in brightness temperature due to factors other than cloud, such as diurnal cooling and heating, a value for the expected change in background temperature during the time period between $t - \Delta t$ and t is required. Temporal changes in AFGWC Surface Temperature database are used to predict the expected background temperature change:

$$\Delta T_{bck} = T_{skin}(t) - T_{skin}(t - \Delta t), \quad (24)$$

where $T_{skin}(t)$ and $T_{skin}(t - \Delta t)$ are linearly time-interpolated skin temperatures calculated from the AFGWC Surface Temperature database entries valid at times that bracket the valid times of the satellite data. The Surface Temperature values used in the time interpolation are taken from the 1/8th mesh grid point in the database that is closest to the latitude and longitude of the satellite pixel being analyzed. Skin temperatures are typically time-interpolated between the analysis and three-hour forecast fields that bound a particular geostationary satellite image valid time; however, if delays in SFCTMP data availability extend beyond three hours from the most recent SFCTMP analysis the interpolation can also be performed between the three-hour and 4.5-hour forecasts.

The infrared temporal difference test requires that the satellite-observed brightness temperature decrease over the time period by an amount greater than an empirically defined threshold (δ_{IR}):

$$\bullet \Delta T_{bck} - \Delta T_{IR} > \delta_{IR} .$$

Note that ΔT_{IR} is negative for newly developing cloud and ΔT_{bck} can be either negative or positive depending on where in the diurnal cycle the data were measured (e.g., morning heating: $\Delta T_{bck} > 0$, midday $\Delta T_{bck} \approx 0$, or nighttime cooling: $\Delta T_{bck} < 0$). The value of the threshold, δ_{IR} , is listed in Table A-4. Results of the IR test are evaluated differently by the bispectral and single-channel algorithms. If the inequality evaluates as true in the single-channel algorithm then the pixel is classified as cloud-filled and processing moves on to the next pixel in the scene. However, as illustrated in Fig. 17 the bispectral algorithm requires of pixels that pass the infrared test also to be subjected to a visible temporal difference test before they can be classified as cloud.

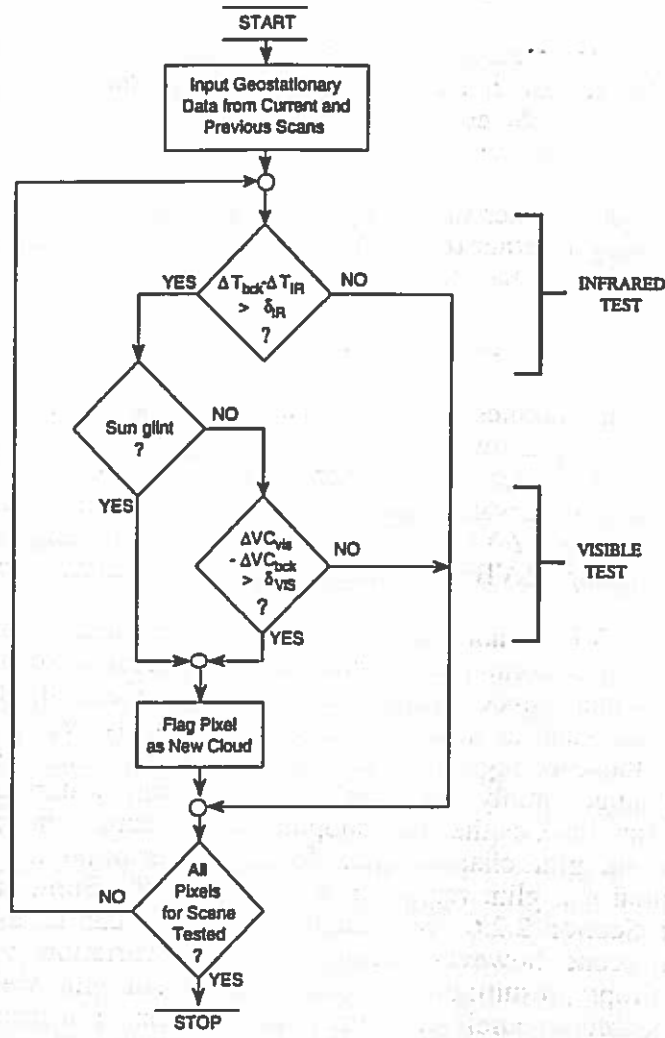


Figure 17. Bispectral Temporal Difference Functional Flow Diagram

The visible temporal difference test is similar to the IR test except that the change in satellite-measured visible counts is tested. The change in visible count for each pixel in the analysis image is defined as:

$$\Delta VC_{vis} = VIS(t) - VIS(t - \Delta t), \quad (25)$$

where $VIS(t)$ represents the visible count measured during the most recent satellite observation period and $VIS(t - \Delta t)$ represents the collocated visible count from the previous image. The measured visible count difference is compared to the expected change in brightness of the terrestrial surface to test for cloud. To account for changes in clear-scene visible count due to factors other than cloud, such as changes in solar illumination that vary with time of day, the expected change in background surface brightness over time is determined from the stored data in the VBC database. Recall from Section 2.2.2 that for each satellite separate VBC database entries are generated and maintained for every observation time throughout the day for which there are usable visible data. Using these data the expected change in the clear-scene visible background (ΔVC_{bck}) can be predicted from:

$$\Delta VC_{bck} = VBC(t) - VBC(t - \Delta t), \quad (26)$$

where $VBC(t)$ is the archived clear scene Visible Background Count valid at the data collection time t and $VBC(t - \Delta t)$ is the archived clear scene data valid at the previous image time $t - \Delta t$. Thus the archived VBC data correspond in time and location to the observed sensor data values at the two observation times.

A determination of whether the change in visible count is sufficient for cloud detection is made by testing whether the satellite observed visible count increases by an amount greater than expected for the clear scene background:

$$\bullet \Delta VC_{vis} - \Delta VC_{bck} > \delta_{VIS},$$

where δ_{VIS} is an empirically defined threshold. The value of the threshold δ_{VIS} is listed in Table A-4. The underlying assumption in the visible temporal difference test is that clouds will be brighter than the terrestrial background, thus if there is new cloud in the FOV since the last satellite observation then ΔVC_{vis} will be greater than 0. The clear-scene background count change, ΔVC_{bck} , can be either positive or negative depending on the time of day (e.g., morning: $\Delta VBC > 0$, midday: $\Delta VBC \approx 0$, afternoon: $\Delta VBC < 0$).

As discussed in AVHRR and DMSP algorithm descriptions, analysis of visible data can be problematic over some backgrounds because the clear scene can produce a visible channel signature that can be misinterpreted as cloud. Generally this includes any highly reflective surface such as snow, ice and desert. However, for the temporal difference algorithm these backgrounds are not a problem since their reflectance characteristics don't change rapidly with time. Sun glint from water backgrounds does need to be accounted for since, as the solar subpoint moves across the field of view of a geostationary satellite, the glint characteristics for any water point can change quickly. Information on potential sun glint regions is provided by the Sun-Satellite Geometry database described in Section 2.2.6. Sun-satellite geometry can be used to locate the specular point for any scene, however, because of normal variations in sea state water surfaces are rarely isotropic, resulting in the occurrence of sun glint well away from the specular point. Thus, based on purely geometric considerations, it is necessary to identify a relatively large area where the potential for sun glint exists. For the geostationary algorithm, sun glint criteria were established through thresholds applied to the solar zenith, satellite zenith, and solar-satellite azimuth angles to ensure 1) the satellite sensor is looking toward the sun, and 2) that the angle of incidence (solar zenith angle) is approximately equal to the angle of reflection (satellite zenith angle). Candidate sun glint areas are defined by the following background surface and solar-satellite geometry tests:

- Background surface type must be water ,
- and* • $|\psi - \theta| < THRESH_{zenith}$,
- and* • $THRESH_{loazimuth} < \phi < THRESH_{upazimuth}$,

where $THRESH_{upazimuth}$ and $THRESH_{loazimuth}$ are empirically derived threshold values and $THRESH_{zenith}$ defines the magnitude by which the solar zenith angle (θ) must differ from the satellite zenith angle (ψ) to support sun glint. These threshold values differ from those used by the AVHRR Cloud Analysis Algorithm and are contained in Table A-4 in Appendix A. Figure 4 provides an illustration of the solar-satellite geometry definitions used by the above tests. When values for these angles fall within a specified range for a given pixel location, the bispectral temporal difference test reverts to the single channel IR algorithm as shown in Fig. 17.

Figure 17 illustrates that the bispectral algorithm classifies a pixel as cloudy if and only if the change in both infrared brightness temperature and visible count between sequential images satisfy the respective temporal difference requirements defined above. Otherwise the pixel is considered cloud-free.

5.2 DYNAMIC THRESHOLD TEST

As illustrated in Fig. 15, the second level of processing for pixels within the analysis image is a dynamic threshold test. This test uses information from the temporal differencing tests to characterize the thermal structure of new clouds to classify the remaining pixels within a surrounding analysis area. Processing is performed by first dividing the analysis scene into subregions. During SERCAA optimal results were obtained using a subregion with a member size of 128 x 128 pixels, however, the size may be adjusted to meet specific implementation and processing requirements. A dynamic cloud threshold is established for each subregion based on the spectral characteristics of all member pixels previously classified by the temporal difference test as cloudy. The minimum and maximum brightness temperatures of the cloudy pixels are identified and then used to define an infrared brightness temperature cloud threshold (T_{cloud}). The threshold T_{cloud} is defined as:

$$T_{cloud} = T_{max} - \gamma(T_{max} - T_{min}), \quad (27)$$

where γ is a tunable factor used to eliminate anomalously warm pixels from the threshold calculation, and T_{max} and T_{min} are the maximum and minimum brightness temperatures, respectively, of the pixels classified as cloud-filled by the temporal differencing test within the image subregion. Thus, the infrared brightness temperature cloud threshold (T_{cloud}) is defined by the maximum temperature of the cloudy pixels classified by the temporal differencing test less an offset. The value of γ is listed in Table A-4.

The dynamic brightness temperature threshold is then applied to all pixels within the subregion as illustrated in Fig. 18. Thus for all pixels $i = 1, 2, 3, \dots, N$ where N is the

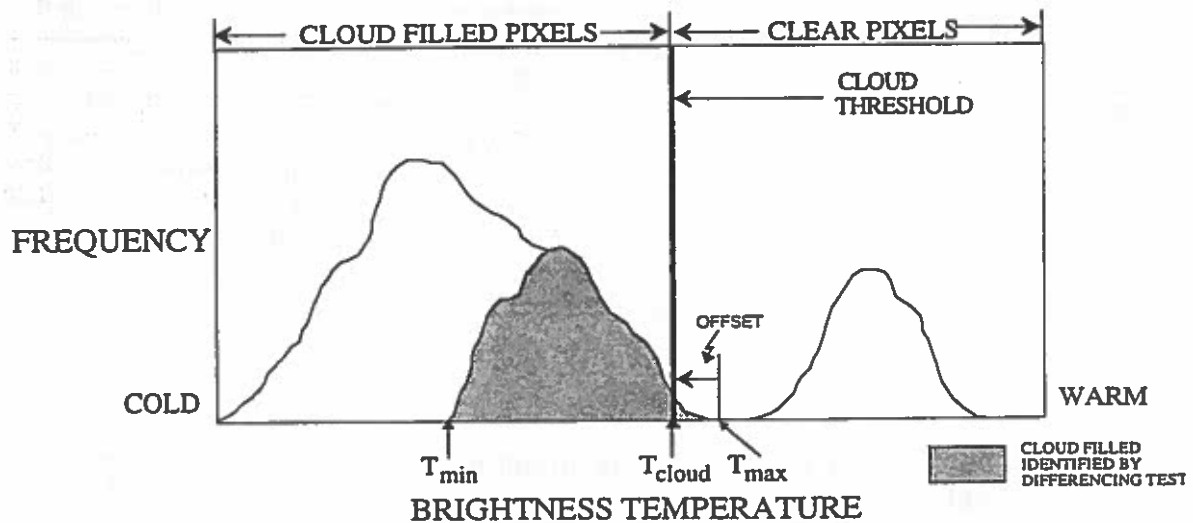


Figure 18. Dynamic Threshold Technique

number of pixels in the subregion (128^2 for SERCAA), the dynamic threshold test detects cloud in pixel i if:

$$\bullet T_i < T_{\text{cloud}},$$

where T_i is the infrared brightness temperature of the pixel i being tested.

The bispectral visible dynamic threshold technique is analogous to that for the IR. This test uses information from the temporal differencing tests to characterize the brightness attributes of new clouds to classify the remaining pixels within a surrounding analysis area. A visible count dynamic cloud threshold is established for each subregion based on the spectral characteristics of all member pixels previously classified by the bispectral temporal difference test as cloudy. The minimum and maximum visible count values of the cloudy pixels are identified and then used to define a visible count cloud threshold (VC_{cloud}). The threshold VC_{cloud} is defined as:

$$VC_{\text{cloud}} = VC_{\text{min}} + \gamma(VC_{\text{max}} - VC_{\text{min}}), \quad (28)$$

where γ is a tunable factor used to eliminate anomalously bright pixels from the threshold calculation, and VC_{max} and VC_{min} are the maximum and minimum brightness, respectively, of the pixels classified as cloud-filled by the temporal differencing test within the image subregion. Thus, the visible count cloud threshold (VC_{cloud}) is defined by the minimum brightness of the cloudy pixels classified by the temporal differencing test plus an offset. The value of γ is listed in Table A-4.

The visible dynamic brightness threshold is then applied to all pixels within the subregion as illustrated in Fig. 18. Thus for all pixels $i = 1, 2, 3, \dots, N$ where N is the number of pixels in the subregion (128^2 for SERCAA), the visible dynamic threshold test detects cloud in pixel i if:

$$\bullet VC_i > VC_{\text{cloud}},$$

where VC_i is the visible count value of the pixel i being tested.

Dynamic threshold tests are only performed when the total number of pixels in the local analysis subregion classified as cloud by the temporal differencing test exceeds a threshold percent $\text{THRESH}_{\text{td_pct}}$ of the total number of pixels in that region. This test is performed to minimize any noisy or misregistered satellite data from adversely affecting the dynamic threshold cloud detection process. For example, if the temporal difference test classifies 304 pixels in a 128×128 analysis region and the minimum required threshold is 2 percent, no dynamic thresholding will be performed on these pixels since 304 is less than $.02(128^2) = 328$. (This does not mean, however, that these pixels remain unclassified; subsequent spectral tests as described in the upcoming Section 5.3 have yet to be applied.) Values of $\text{THRESH}_{\text{td_pct}}$ are listed in Table A-4.

5.3 SPECTRAL DISCRIMINANT TESTS

The final level of geostationary algorithm processing exploits static (i.e., non-temporal) cloud spectral signatures similar to those used for the AVHRR and DMSP algorithms. For the purpose of detecting cloud, results of the spectral tests are only evaluated for pixels that have not been classified as cloud-filled by either the temporal differencing or dynamic threshold tests, although the spectral tests are applied to the entire scene on a pixel-by-pixel basis.

Multiple spectral tests are available; however, the set of tests applied to any image pixel is dependent on 1) the sensor channels available from the particular geostationary satellite that collected the image (Table 1), and 2) the scene-solar illumination. A positive result from any spectral discriminant test is sufficient to classify the pixel as cloud. Table 14 provides a listing of the spectral discriminants and the conditions under which each test is applied. Figure 19 provides a schematic illustration of the data flow for the spectral tests.

The spectral tests are segregated into night and day applications. The solar zenith angle associated with each pixel is tested against a day-night threshold to determine which set of tests can be applied. Pixels with solar zenith angles less than a threshold $THRESH_{spectral_solzen}$ are subjected to daytime tests, otherwise nighttime spectral tests are applied. The value of the spectral solar zenith angle threshold is listed in Table A-4. Note that this threshold is separate from the day/night threshold used by the temporal difference test. Sensor data are also checked at each stage of the spectral discriminant algorithm to determine if the required sensor channels are available.

5.3.1 Solar-Independent Spectral Tests

If 11 μm thermal infrared data are available (these are likely to be available for all geostationary platforms at all times) then the first geostationary spectral test, called the Cold Cloud Test, identifies clouds with an infrared brightness temperature lower than a

Table 14. Geostationary Spectral Discriminants

Test	Day Application	Night Application	Cloud Type
$VIS - VBC > THRESH_{land}$ (over land)	✓ ¹		Obvious (bright) Cloud
$VIS - VBC > THRESH_{water}$ (over water)	✓ ¹		
$\theta \leq THRESH_{geo_pcp_solzen}$ and $T_{skin} - T_{11} > THRESH_{cold}$ and $T_{3.9} - T_{11} > THRESH_{pcp_goes}$ and $VIS * \sec(\theta) \geq THRESH_{pcp_vis}$	✓		Precipitating Cloud
$T_{3.9} - T_{11} > THRESH_{LCd}$ $T_{11} - T_{3.9} > THRESH_{LCn}$	✓ ¹	✓	Low Cloud and Fog
$T_{skin} - T_{11} > THRESH_{cold}$	✓	✓	Obvious (cold) Cloud
$T_{3.9} - T_{11} > THRESH_{TCi}$		✓	Thin Cirrus Cloud
$T_{11} - T_{12} > THRESH_{(Q,\psi)}$ and $VIS * \sec(\theta) > THRESH_{DCi(2)}$ and if snow then $T_{skin} - T_{11} > THRESH_{DCi(1)}$	✓		Daytime Cirrus Cloud ²

VIS = Visible Channel Count

VBC = Visible Background Count

¹ Requires sun glint and/or snow background surface filters.

² This is a recommended daytime thin cirrus detection approach that has not yet been tested.

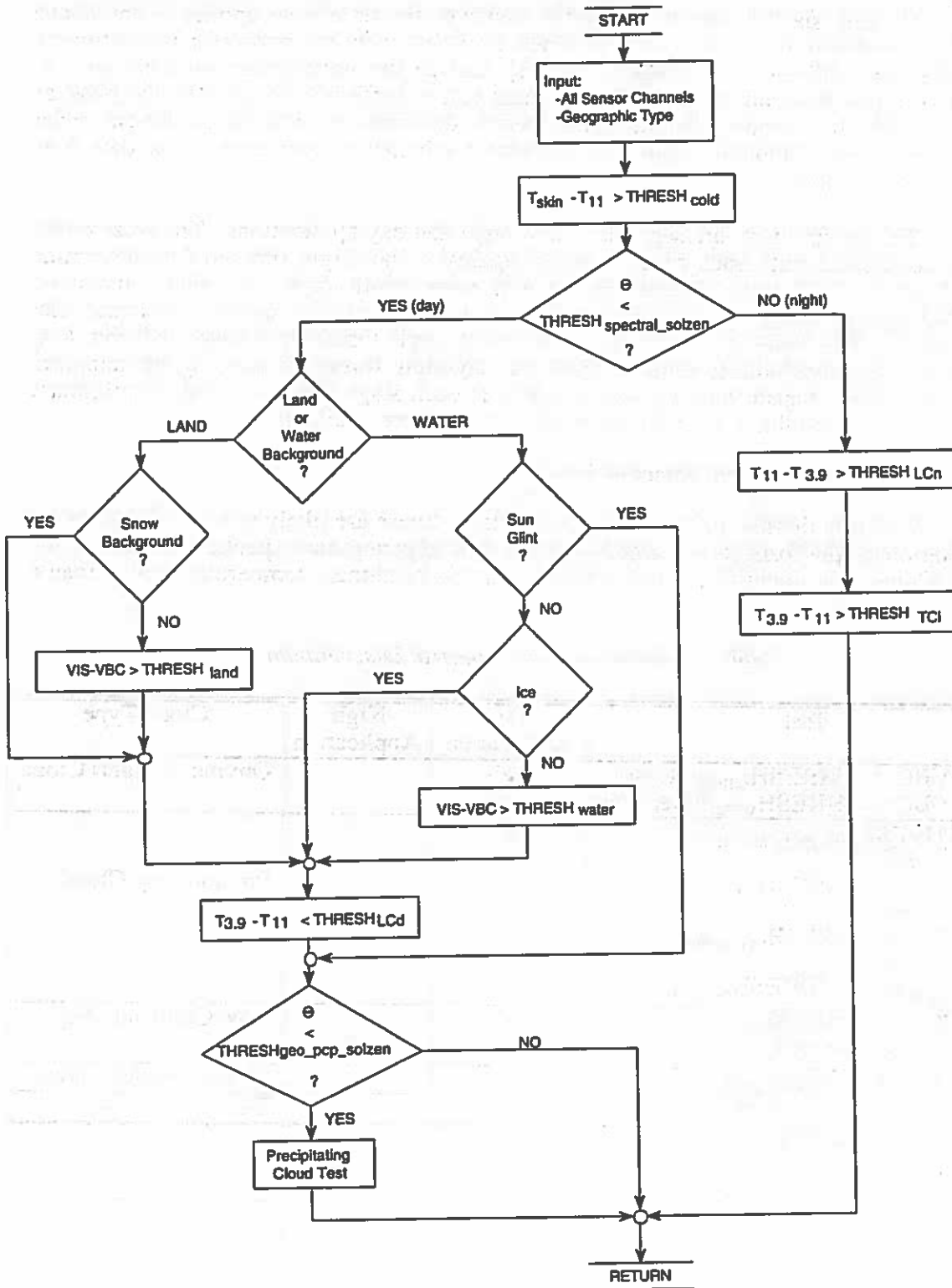


Figure 19. Spectral Test Functional Flow Diagram

predicted clear-scene brightness temperature by a pre-defined threshold. This test is executed regardless of the time of day. The Cold Cloud Test is defined as:

$$\bullet T_{\text{skin}} - T_{11} > \text{THRESH}_{\text{cold}},$$

where T_{skin} is the predicted clear scene brightness temperature, T_{11} is the 11 μm channel brightness temperature and $\text{THRESH}_{\text{cold}}$ is the cloud detection threshold. A time-interpolated skin temperature derived from the AFGWC Surface Temperature database is used as the predicted clear-scene brightness temperature, T_{skin} . No corrections are required to account for differences between the satellite observed and modeled skin temperatures (e.g., atmospheric attenuation, calibration error, etc.) because the test is only required to detect cloud with a strong thermal signature and, as such, the threshold $\text{THRESH}_{\text{cold}}$ is large.

5.3.2 Day Condition Spectral Tests

Daytime tests require 1) solar zenith angles less than the spectral solar zenith angle threshold, $\text{THRESH}_{\text{spectral_solzen}}$ and 2) data from a visible channel and/or middle- and thermal-infrared channels. All geostationary satellites have a visible channel. However, at present only GOES-7 has the additional middle (3.9 μm) and thermal-infrared (11 and 12 μm) channel combinations (refer to Table 1). Should such channel combinations be added to other platforms in the future, these tests will be equally applicable to those data sets. Three daytime tests are used: the first identifies obvious bright cloud; the second, low cloud and fog; and the third, precipitating clouds.

The first daytime cloud test requires visible channel data only and tests for obvious bright cloud. This test is dependent on background surface type and snow or ice cover determined from the Geographic Type and Snow and Ice Location support databases. If the background type is land, a test is performed to determine if the visible channel count is higher than the stored VBC by an amount greater than a pre-defined threshold. Snow covered locations are eliminated. The test is defined as:

$$\bullet \text{VIS} - \text{VBC} > \text{THRESH}_{\text{land}},$$

and

$$\bullet \text{not snow covered},$$

where VIS is the observed visible channel count value, VBC is the count value stored in the clear-scene Visible Background Count database (refer to Section 2.2.2), and $\text{THRESH}_{\text{land}}$ is the difference threshold value. If the VIS - VBC difference exceeds the cloud threshold value and the region is not identified as a snow, then the pixel is classified as cloud.

If the background type is water, a test is performed to determine if the visible count is higher than a static threshold count value for water. Ice locations and sun glint are not processed. The test is defined as:

$$\bullet \text{VIS} - \text{VBC} > \text{THRESH}_{\text{water}},$$

and

$$\bullet \text{not ice covered},$$

and

$$\bullet \text{not sun glint},$$

where VIS is the observed visible channel count value, VBC is the count value stored in the clear-scene Visible Background Count database (refer to Section 2.2.2), and $\text{THRESH}_{\text{water}}$ is the cloud detection threshold count value. If the observed visible channel count value is greater than the cloud/no-cloud threshold value and the region is

identified as being uncontaminated by potential sun glint and/or ice backgrounds then the pixel is classified as cloud-filled. Sun glint regions are determined geometrically as described in Section 5.1.

The second daytime test, which identifies low cloud and fog, is used when middle- and thermal-IR data are simultaneously available. This test checks whether the 3.9 μm channel brightness temperature is higher than the corresponding 11 μm temperature. The Low Cloud and Fog test is defined as:

- and
- $T_{3.9} - T_{11} > \text{THRESH}_{\text{LCd}}$,
 - not sun glint,

where $T_{3.9}$ is the 3.9 μm channel brightness temperature, T_{11} is the 11 μm channel brightness temperature, and $\text{THRESH}_{\text{LCd}}$ is the cloud detection threshold whose value is listed in Table A-4. If the brightness temperature difference exceeds the cloud threshold value and the pixel location is not in a sun glint area, then the pixel is classified as cloud-filled.

The third daytime test is the precipitating cloud test that predominantly identifies cumulonimbus clouds. The geostationary precipitating cloud test is identical in theoretical basis and algorithm flow to the AVHRR precipitating cloud test described in Section 3.2.2.2, however there are two implementation details that are different. The first is that the geostationary precipitating cloud test is executed only when the local solar zenith angle is less than a preset threshold:

- $\theta < \text{THRESH}_{\text{geo_pcp_solzen}}$,

where $\text{THRESH}_{\text{geo_pcp_solzen}}$ is the solar zenith angle threshold defined in Appendix A, Table A-4. The second difference is that the individual spectral signature thresholds are different than those for the AVHRR algorithm:

- and
- $T_{\text{skin}} - T_{11} > \text{THRESH}_{\text{cold}}$,
 - $T_{3.9} - T_{11} > \text{THRESH}_{\text{pcp_goes}}$,
 - $\text{VIS} * \sec(\theta) \geq \text{THRESH}_{\text{pcp_vis}}$,
- and

where $\text{THRESH}_{\text{cold}}$, $\text{THRESH}_{\text{pcp_goes}}$, and $\text{THRESH}_{\text{pcp_vis}}$ are empirically defined cloud thresholds defined in Table A-4.

Finally, an untested but recommended approach for detecting thin cirrus in the daytime when simultaneous visible, 11 μm , and 12 μm data are available (currently only with GOES) is presented here. This test is outlined in Table 14 and is directly analogous to the AVHRR daytime thin cirrus test described in Section 3.2.2.3. First, the $T_{11} - T_{12}$ brightness temperature difference is compared to a cirrus threshold:

- $T_{11} - T_{12} > \text{THRESH}(Q, \psi)$,

where $\text{THRESH}(Q, \psi)$ is a table of cloud detection thresholds which are functions of total atmospheric precipitable water, Q , and path length, characterized by ψ . Theoretically derived threshold values are contained in Table A-2b in Appendix A.

If the $T_{11} - T_{12}$ test evaluates as true, then a second test is performed to eliminate clouds with a high visible brightness from being classified as thin cirrus. This is analogous to the visible and near-IR channel brightness checks performed by the

AVHRR test (Section 3.2.2.3), but with the exception that geostationary satellites have only a single visible channel. Thus the visible count is normalized to a sun overhead condition and then compared to a visible count threshold:

$$\bullet \text{VIS} * \sec(\theta) < \text{THRESH}_{\text{Dci}(2)}$$

where VIS is the satellite-derived visible count, θ is the satellite zenith angle, and $\text{THRESH}_{\text{Dci}(2)}$ is the cloud threshold defined in Table A-4.

A final test is required to check for snow-covered backgrounds is identical to the corresponding AVHRR test. If the Snow and Ice Location database (Section 2.2.4) identifies the surface background as being snow or ice covered then the pixel is subjected to an additional test to ensure that the signature detected by the $T_{11} - T_{12}$ difference test was not the underlying snow or ice background rather than cirrus cloud:

$$\bullet T_{\text{skin}} - T_{11} > \text{THRESH}_{\text{Dci}(1)},$$

where T_{skin} is the time-interpolated skin temperature from AFGWC Surface Temperature database (Section 2.2.1) and $\text{THRESH}_{\text{Dci}(1)}$ is the cirrus cloud detection threshold defined in Table A-4.

5.3.3 Night Condition Spectral Tests

Nighttime tests use data from channels in the 3.7 - 3.9 μm middle infrared and 10 - 12.5 μm long wave thermal infrared window regions. There are two nighttime tests: the first identifies low cloud and fog and the second identifies thin cirrus.

The nighttime Low Cloud and Fog Test determines whether the 11 μm brightness temperature is higher than the 3.9 μm channel by an amount greater than a pre-defined threshold. If the brightness temperature difference exceeds the threshold value then the pixel is classified as cloud-filled. The nighttime Low Cloud and Fog Test is defined as:

$$\bullet T_{11} - T_{3.9} > \text{THRESH}_{\text{LCn}},$$

where T_{11} is the 11 μm channel brightness temperature, $T_{3.9}$ is the 3.9 μm channel brightness temperature, and $\text{THRESH}_{\text{LCn}}$ is the cloud detection threshold whose value is listed in Table A-4.

The Thin Cirrus Cloud Test uses the reverse signature and checks whether the 3.9 μm channel brightness temperature is higher than that at 11 μm . If the brightness temperature difference exceeds a pre-defined threshold then the pixel is classified as cloud-filled. The Thin Cirrus Cloud Test is defined as:

$$\bullet T_{3.9} - T_{11} > \text{THRESH}_{\text{TCi}},$$

where $T_{3.9}$ is the 3.9 μm channel brightness temperature, T_{11} is the 11 μm channel brightness temperature, and $\text{THRESH}_{\text{TCi}}$ is the cloud detection threshold whose value is listed in Table A-4.

5.4 CONFIDENCE FLAG DETERMINATION

In addition to analyzed cloud information, the geostationary algorithm provides information on the expected accuracy of the analysis for each pixel. Accuracy estimates are intended to provide the end user with an indication of how much confidence to place in the analysis for any given pixel and, as such, are referred to as confidence flags. Three levels of confidence are defined: LOW, MIDDLE, and HIGH. However, in the geostationary algorithm, only two of these flags are used: MIDDLE and HIGH. The level of confidence assigned to each pixel is based on the results of the temporal differencing and spectral signature tests. Cloud-filled pixels that were identified only by a spectral test (Section 5.3), are assigned a MIDDLE confidence. Cloud-filled pixels that were detected by either the temporal differencing or dynamic threshold tests (Sections 5.1 and 5.2) are assigned a HIGH confidence.

Confidence flag values reflect the overall characteristic of the geostationary algorithm to not over analyze cloud. The temporal differencing algorithm will only detect newly developed or moving clouds, however since it is only minimally dependent on knowledge of the absolute clear-scene background characteristics testing has shown it to be extremely reliable. Similarly, the dynamic threshold is based on actual satellite observations of the clouds themselves, arguably the most accurate information available on the actual cloud characteristics since they are resolved by the satellite. Because of this no assumptions are required to correct for atmospheric attenuation, calibration errors, or radiative characteristics of the cloud. Thus the only major limitation on these two algorithms is that they will only detect new or moving clouds, plus surrounding clouds with the same thermal and reflective characteristics. So, while they may not detect all cloud in a scene, there is high confidence that only clouds are detected.

5.5 OUTPUT PRODUCT

The output product of the Geostationary Cloud Analysis Algorithm is a bit-mapped 8-bit MCF identical to those produced by the AVHRR and DMSP algorithms. One MCF is produced for each pixel in the geostationary imagery and is stored in an MCF file for later use by the Layer and Type and Cloud Analysis Integration Algorithms (see Fig. 1).

Information stored in the MCF includes: a cloud/no-cloud flag, flags for low cloud, thin cirrus cloud, and precipitating cloud, a missing or bad data flag, and the confidence level. These geostationary MCF file bit assignments shown in Table 15 and are discussed below.

Table 15. Geostationary Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled ; OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud
3	Precipitating Cloud	ON = Precipitating Cloud
4	Partial Cloud	Not Used By Geostationary Algorithm
5	Data Dropout	ON = Missing Or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

Cloud Mask - Bit 0

The cloud mask bit is set to ON, indicating a cloud-filled pixel, if the pixel is determined to be cloud-filled by at least one of the geostationary temporal differencing, dynamic threshold, or spectral cloud tests described in Sections 5.1, 5.2, and 5.3.

Low Cloud - Bit 1

The low cloud bit is set if the following cloud test is passed:

- Low Cloud and Fog Test (Section 5.3.2) .

Thin Cirrus Cloud - Bit 2

The thin cirrus cloud bit is set if *only* the following cloud test is passed:

- Nighttime Thin Cirrus Cloud Test (Section 5.3.3) .

Precipitating Cloud - Bit 3

The precipitating cloud bit is set if the following cloud test is passed:

- Precipitating Cloud Test (Section 5.3.2) .

Partial Cloud - Bit 4

The partial cloud bit is not used by the Geostationary Cloud Analysis Algorithm.

Data Dropout - Bit 5

The data dropout bit is set to ON if the data for the pixel are missing or unreliable.

Confidence Flag - Bits 6 & 7

The confidence flag bits are set to indicate LOW (1), MIDDLE (2), or HIGH (3) confidence as detailed in Section 5.4.

6. CLOUD TYPING AND LAYERING ALGORITHM DESCRIPTION

The Cloud Typing and Layering Algorithm is a two-stage process that operates on output from the cloud analysis algorithms and sensor data from all polar and geostationary satellites (Fig. 1). Pixels classified as cloudy in the native satellite scan projection are analyzed to retrieve layer information and then remapped to a polar stereographic projection. Statistics describing total cloud and cloud layer distributions are accumulated over $1/16^{\text{th}}$ mesh grid cells and written to a final output file.

The first processing stage involves segregation of pixels identified as cloudy by each of the cloud analysis algorithms into cloud layers, using long wave IR data. Layers are then classified into cumuliform or stratiform cloud types. Processing is performed in the raw satellite projection over large regions of data to minimize the occurrence of artificial cloud boundaries that may appear as artifacts of the analysis algorithm. A single set of algorithms is applied to data from all satellite platforms, the only differences being the magnitude of thresholds used for distinguishing cloud type. Cloud typing is based on scale length and thresholds are chosen so that the transition from cumuliform to stratiform occurs at roughly the same physical size for each satellite.

The second stage of processing determines the number of cloud layers and associated attributes in an individual $1/16^{\text{th}}$ mesh grid cell. In order to move from a satellite projection (on which the cloud typing is based) to a polar stereographic projection (on which the $1/16^{\text{th}}$ mesh grids are based), intermediate files of $1/16^{\text{th}}$ mesh grid coordinates (i, j) are created for each input image to facilitate the remapping process. Up to four floating cloud layers are identified for each $1/16^{\text{th}}$ mesh grid cell and a fractional cloud amount, type, and cloud top temperature are calculated for each. A schematic illustration of the cloud typing and layering algorithm is provided in Fig. 20.

A key goal in the design of these algorithms has been to minimize discontinuities in cloud layers between adjacent grid cells, while still allowing sufficient flexibility for layers to "float", or change mean heights, over longer distances. An additional constraint is imposed by the requirement that there be no more than four layers identified in each $1/16^{\text{th}}$ mesh grid cell.

The concern in regard to layer discontinuity at the $1/16^{\text{th}}$ mesh grid cell level is addressed in both the typing and layering algorithms, but in a slightly different manner for each. The typing procedure is applied to relatively large regions of imagery in the original satellite projection. Image sizes are selected based on computational expediency. Consequently, there are no $1/16^{\text{th}}$ mesh grid cell boundaries imposed on the results. The layering algorithm, on the other hand, must operate within grid cell boundaries to allow enforcement of the four layer requirement. However, the area of analysis is extended to adjacent $1/16^{\text{th}}$ mesh grid cells (over a 3×3 grid cell region) although results are applied only to the center grid cell. This allows the layering results in a particular grid cell to be influenced by adjacent data to minimize discontinuities between grid cells. Since the pixels contributing to the layering results are from a 3×3 grid cell area, layers are allowed to float as the moving "layering" window is applied to each grid cell. The 3×3 $1/16^{\text{th}}$ mesh grid cell analysis region is simply a starting point which can be enlarged for greater continuity if needed.

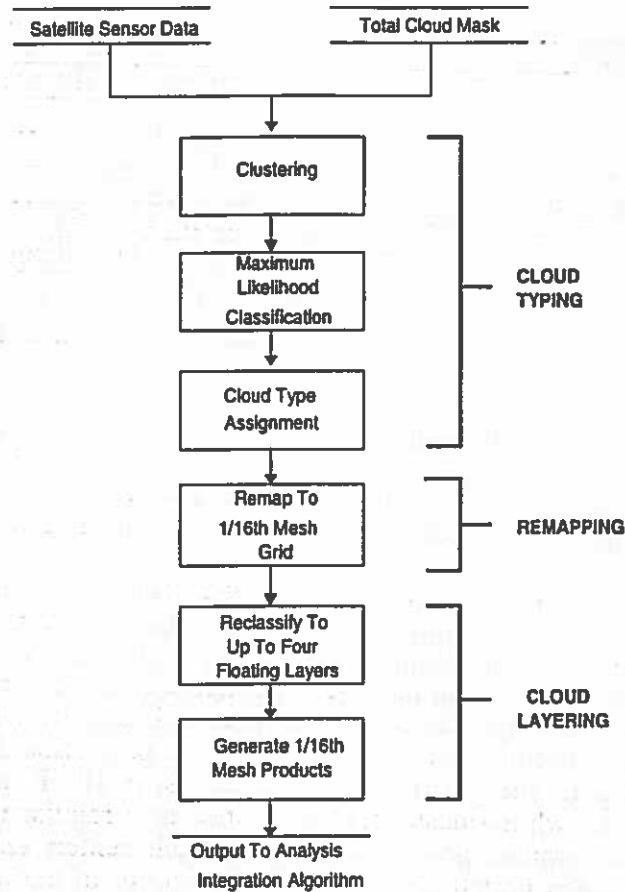


Figure 20. SERCAA Cloud Typing and Layering Algorithm

6.1 DATA REQUIREMENTS AND INPUTS

The specific data inputs used in the typing and layering algorithms include:

- Calibrated and byte-scaled LWIR sensor data in the native satellite projection (AVHRR channel 4; GMS / GOES / DMSP / METEOSAT 10-12 μm channels).
- An MCF cloud mask for each satellite IFOV obtained as output from the respective cloud analysis algorithms (refer to Table 16).
- A location map that provides 1/16th mesh i, j grid coordinates for each satellite IFOV (see below).

Cloud mask data are obtained from the respective cloud analysis algorithm output files as bit-mapped 8-bit MCF values that contain cloud information for each IFOV in the original satellite image. The MCF bit-mapping key is defined in Table 16. Cloud information includes: a cloud/no cloud flag, flags indicating specific cloud types identified by the AVHRR and GOES algorithms (i.e., low, cirrus and precipitating cloud), a missing data bit and a confidence flag in the range of 1-3 indicating low, middle, or high confidence in the cloud analysis, respectively.

Table 16. Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled; OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud
3	Precipitating Cloud	ON = Precipitating Cloud
4	Partial Cloud (From DMSP)	If ON Then Bit 0 = OFF
5	Data Dropout	ON = Missing or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

The cloud layering and remapping steps require *i* and *j* coordinate maps to identify which 1/16th mesh grid cell each satellite pixel belongs. The values of *i*, *j* are 1/16th mesh grid coordinates for each pixel. These data are used for both layer determination and processing of the cloud product information within grid cells.

The AFGWC standard secant polar stereographic projection is used to map all SERCAA analysis products to a common database. This projection is generated geometrically by positioning a secant plane normal to the Earth's axis at a "standard" or "true" latitude of 60°. Lines of constant latitude are concentric circles around the center grid point and lines of constant longitude are straight lines radiating from the pole. In the AFGWC polar stereographic northern (southern) hemispheric projection, the center of the grid is the North (South) pole, the positive *i*-axis (columns) is 10° E, and the positive *j*-axis (rows) is 100° E. A "whole-mesh grid" is defined as a regular rectangular grid overlaid on the polar stereographic projection with grid cell centers exactly 381.0 km apart at the true latitude. Other nested grids are simply fractions of the whole-mesh grid size: 1/2 mesh (190.5 km), 1/4 mesh (95.3 km), 1/8th mesh (47.6 km), 1/16th mesh (23.8 km), etc. The SERCAA integrated analysis output grid resolution is 1/16th mesh. A more complete description of the AFGWC grids is provided by Hoke et al. (1981). All conversions between polar stereographic (*i*, *j*) and Earth (lat, lon) are made in adherence to these conventions.

6.2 CLOUD TYPING

The Cloud Typing Algorithm operates on large sections of imagery to provide continuity over large cloud formations. While the image size selected is resource driven, reasonable results have been obtained for image sizes that range from several hundred to several thousand kilometers across a scene. Thus, the image size selected is not conditional to the operation of the algorithm and may be adjusted to meet implementation requirements.

The cloud typing procedure is a two-step process in which the cloud-filled pixels are first stratified by the LWIR brightness temperature into layers. The size of connected pixels in each layer is then used to make a cumuliform or stratiform cloud type determination. The process of coalescing cloud type classes, from one layer to the next, is illustrated in Fig. 21. The processing steps are described in detail in the sections that follow.

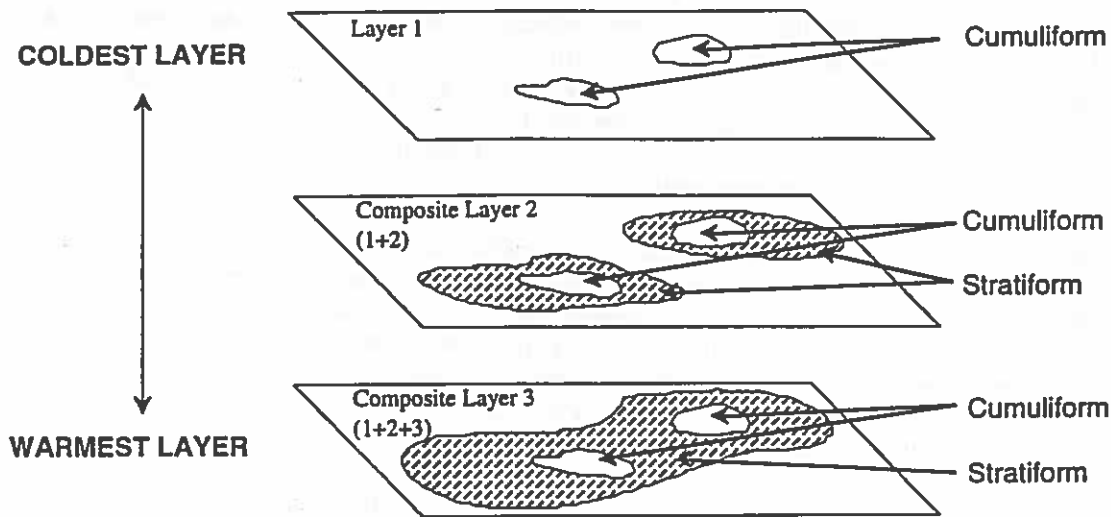


Figure 21. Cloud Type Determination

6.2.1 Height Stratification

Height stratification of sensor data for the purpose of determining cloud type is achieved by applying an unsupervised clustering routine and a Bayesian maximum likelihood classifier to the sensor LWIR channel. Implicit in this approach is the assumption that LWIR clusters are stratified by height. A myriad of generic clustering routines exist from which one could achieve the same or similar results as the algorithm described below. In general, these routines vary in their approach to pixel selection and merging criteria. The criteria for our selection of the algorithms described below were based on processing efficiency and speed. The particular routines selected for testing of the algorithm described here were obtained from a publicly-distributed image processing software package named the Image Processing Workbench (IPW) (Frew, 1990). The IPW programs `ustats` and `bayes` served as our unsupervised clustering routine and maximum likelihood classifier, respectively.

In general, unsupervised clustering involves the definition, identification, and mapping of spectral values into homogeneous clusters. The SERCAA clustering algorithm is generic and designed to define and identify natural groupings within the spectral domain of pixels in feature space. Unsupervised classification determines the number of clusters present in an image and provides descriptive statistics such as cluster means, variances and inter-cluster covariance. These are used as input to a maximum likelihood classifier that uses the data to assign each image pixel to the cluster of which it has the highest probability of being a member.

Height stratification begins by removing all clear pixels from the cluster processing. This is accomplished by applying the cloud and missing data masks generated by the cloud analysis algorithms (i.e., MCF Bits 0 and 5 from Table 16) to the raw IR image. Cloud-free or missing LWIR pixel values are replaced with zero so that only cloudy pixels are clustered and subsequently classified.

The IPW program `ustats` is used to perform the unsupervised clustering that generates the cluster statistics used by the IPW maximum likelihood classifier, `bayes`. The program `ustats` requires as inputs the number of output clusters desired (denoted by

the argument **-c**) and the cluster threshold radius (**-r**). The cluster threshold radius is defined as the standard Euclidean distance between a pixel and a cluster centroid in spectral space. Pixels whose distance from a cluster centroid exceeds this threshold are not merged with the cluster. An optional **ustats** argument (**-x**), utilized in the SERCAA Cloud Typing and Layering Algorithm, is the exclusion of the digital number zero from processing. As zeros represent non-cloud pixels and are not desired in the statistical analysis, their exclusion reduces processing time.

After empirical testing, a cluster threshold radius of 15 digital numbers (DNs) has been selected. Given this threshold, **ustats** invokes an iterative procedure to determine the number of clusters in a scene. So as not to constrain the number of clusters **ustats** might find, the input number of clusters (**-c**) is set to an arbitrarily large value, 100. This is to permit the algorithm to determine the number of clusters naturally found in the scene based solely on their spectral characteristics. For a full scene (regardless of sensor), the typical number of clusters found is between 7 and 14.

The clustering routine progresses in two phases. The first phase involves selection of pixels to form intermediate clusters. The approach is to successively sample the image in increasing resolution to obtain an accurate sampling of the total population of pixels. The image is divided in quadtree fashion, with each quad successively divided into smaller quads until single-pixel quads exist. After each division, the pixel from the upper-left-hand corner of each previously unsampled quad is selected to either form the kernel of a new cluster or to merge with an existing cluster. This selection method ensures that there will be no resampling of the same pixels.

In the second phase, the selected pixels are processed by the clustering routine in the following manner (see also Fig. 22):

- 1) Determine if the pixel is clear or cloudy (MCF Bit-0 value of 0 or 1, respectively). If clear, the pixel is excluded from further analysis. If cloudy, continue to next step.
- 2) Locate the nearest intermediate cluster that is within the cluster threshold radius of the current pixel. These clusters are referred to as 'intermediate' because they will continue to grow and shift as more pixels are added in this phase. The maximum number of intermediate clusters allowed is 10 times the maximum number of output clusters desired (**-c**). (This number may be modified for implementation needs.)
- 3) If such a cluster exists, then add the current pixel to it, adjusting the cluster centroid.
- 4) Else:
 - a) If the current number of intermediate clusters is less than the total permitted create a new intermediate cluster containing only the current pixel.
 - b) Else, ignore this pixel and continue with the next pixel.
- 5) Continue until all pixels have been evaluated.
- 6) Select the 100 most populous clusters (as specified by the input parameter **-c**; in practice this number ranges between 7 and 14) and generate the mean and variance for the resulting clusters. Write the information to an ASCII file.

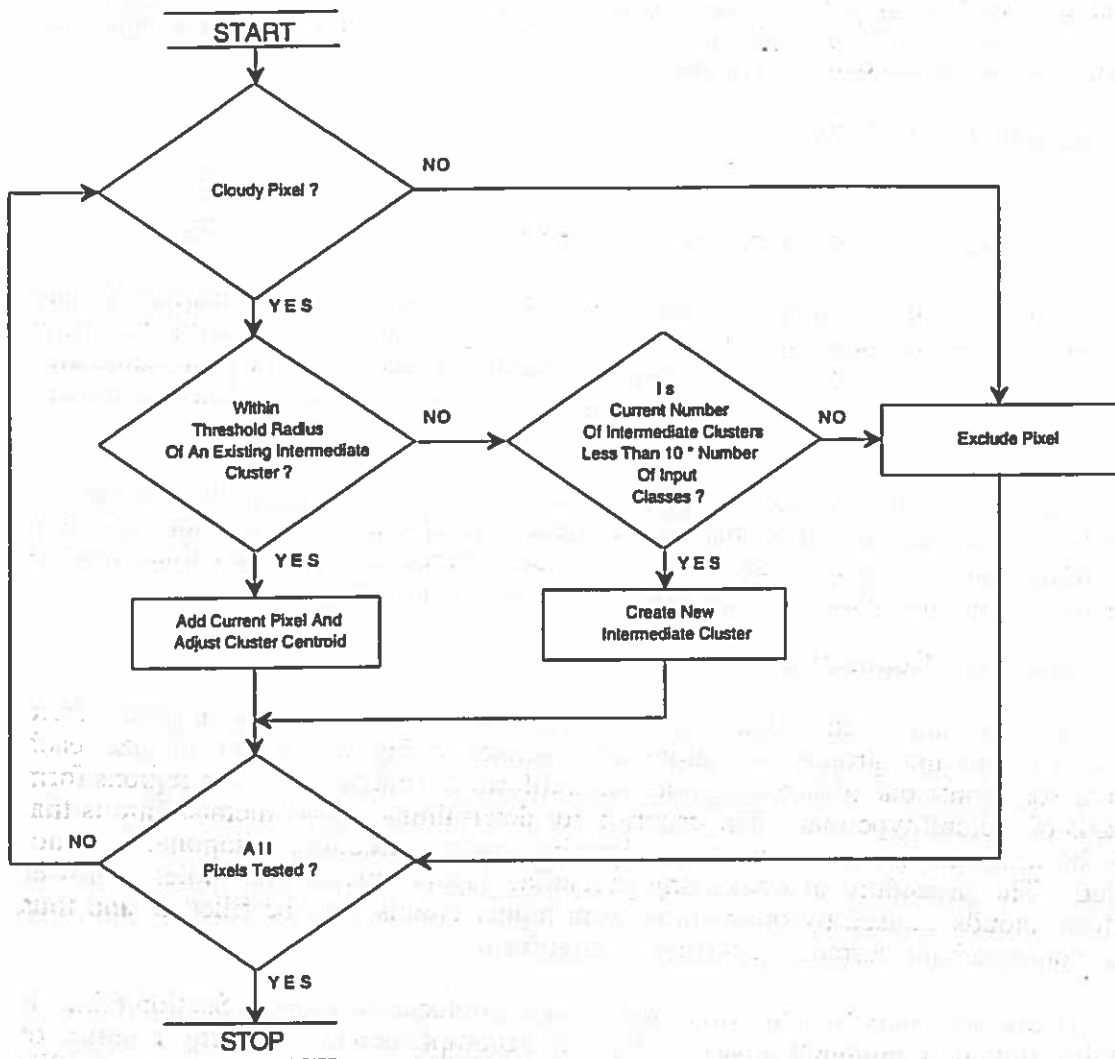


Figure 22. Unsupervised Clustering Pixel Selection

The ASCII file produced in Step 6 above provides the basis by which the Bayesian maximum likelihood classifier (*bayes*) divides the image into layers. Inputs to *bayes* include the statistics file generated by *ustats* (-s), the probability threshold (-t), the DN to be excluded from the analysis (-e), the option to use raw data (-r) (as opposed to floating point data) and selection of a DN to which unclassified pixels are mapped. To prevent the occurrence of unclassified pixels, we use a probability threshold (-t) of 0. This threshold is the chi-square (χ^2) value below which a desired percentage of pixels is mapped as unclassified. To reduce processing time in an operational scheme, 0% unclassified pixels is desired. As in *ustats*, non-cloud pixels (DN = 0) are excluded from analysis to reduce processing time. The resulting map contains layer membership for each cloudy pixel, as well as the mean temperature and variance for each layer.

Each image pixel is characterized in multispectral space by a measurement vector x ; where x contains the measured value from each sensor channel. Recall that the SERCAA cloud typing algorithm uses only the LWIR channel from each satellite, thus x reduces to the scalar x containing only the DN from that one channel. The class to which

a particular pixel belongs is determined using a decision rule that assigns x to a particular class, ω_i , if the probability of the pixel's occurring in that class is greater than the probability of the pixel's occurring in any other class.

According to Richards (1986):

$$x \in \omega_i \text{ if } p(\omega_i | x) > p(\omega_j | x) \text{ for all } j \neq i \quad (29)$$

and
$$p(x | \omega_i) = \sigma_i^{-1} \exp\{-1/2 (x - m_i)^2 / \sigma_i^2\} \quad (30)$$

where i is the total number of classes analyzed by the unsupervised classification, σ_i and σ_i^2 are the standard deviation and variance, respectively, of class ω_i , and m_i is the mean radiance vector for class ω_i . The maximum likelihood decision rule first calculates $p(x | \omega_i)$ for each class $i = 1, 2, \dots, M$ and then assigns x to the class that has the highest calculated probability.

Output from this processing takes the form of a classified image that defines to which class value, ω_i , each pixel has been assigned. Each class is related logically to a height stratification, or layer, in the image. They are ordered such that the coldest layer is represented by the first class (ω_1), and the warmest by the last class.

6.2.2 Top-Down Connectivity

The next step of the Cloud Typing process involves coalescing of pixels from successive layers into groups, or regions as illustrated in Fig. 21. Based on size, each region is sorted into one of two categories, cumuliform or stratiform. These regions form the basis of a cloud type map. The criterion for determining region membership is that pixels be adjacent in one of the four adjoining spatial directions; diagonals are not counted. The procedure of coalescing preceding layers ensures that holes in lower stratiform clouds, caused by obstruction from higher clouds, will be filled in and thus appear connected and correctly identified as stratiform.

As the first step, the classified layer image produced by bayes (Section 6.2.2) is separated into its component classes. This is accomplished by applying a series of 2-column ASCII look-up tables to the classified image using a sequence of IPW commands (`interp`, `mklut`, `lutx`). The `interp` command interpolates between ASCII X-Y pairs of integer breakpoints. As an example, the following look-up table:

```
0 1
3 5
5 7
```

when interpolated, yields the following:

```
0 1
1 2
2 4
3 5
4 6
5 7
```

The IPW program `mklut` creates an IPW look-up table (a single-line image) and writes it to the standard output. In the SERCAA implementation, `mklut` is used without any arguments. The `lutx` function then applies the look-up table to the classified layer image

and maps each pixel according to the class (layer) information contained therein. These three commands are strung together by pipes to form the following UNIX command:

```
interp < {look-up table} | mklut | lutx -i {classified image}
```

The result of this process is to produce a separate binary image for each class/layer where a DN of 1 represents pixels that fall within that layer and all other pixels have a DN of 0.

To assist the region-growing process and to help further define region boundaries, a variance-based texture derivative (Woodcock and Ryherd, 1989) is formed from the binary layer image using the IPW program `texture`. Inputs to the program are window size (-w), minimum window size (-m), scaling factor (-s) and selection of an adaptive window (-a). The following is a typical command line used to generate a texture image:

```
texture -a -m3 -w3 -i {input image} -s10 -f {output image}
```

After empirical trials, we have selected to use an adaptive window, 3 x 3 pixels in dimension. The minimum window size, or the minimum number of pixels for a window to be considered, is set to 3. This is important in instances when a complete window is not possible (as is the case on the edge of an image). Unlike conventional texture routines that center a window around a pixel, `texture` employs an adaptive window that analyzes all possible windows to which a pixel belongs. Texture is calculated for all windows and the 'best' value is assigned to that pixel. The 'best' value is defined as the lowest standard deviation calculated for any window to which the pixel could belong. Each pixel is analyzed in this fashion. The advantage of the adaptive window is to greatly reduce the artificial blockiness resulting from conventional texture routines. The scaling factor (-s) is set to 10 and is used to increase the dynamic range of the output texture values. The texture derivative is band-interleaved by pixel with the binary layer image (using the IPW program `mux`).

Each class/layer is segmented into regions using the IPW function `segment`, a multi-pass, region-growing segmentation algorithm. The multiple pass approach of `segment` allows for incremental growth of regions, pixel-by-pixel. During each pass, regions that are more similar than a stated threshold are allowed to merge, however each region is only allowed one merge per pass. This ensures that each merge is optimal for that region. The maximum number of pixels allowed in any given region is constrained only by the number of pixels in the image.

A simple source code modification to `segment` has been made to facilitate its use in other SERCAA scripts. The output filename extension is now limited to the base name '.rmap'. Specifically, the changes as they appear in the code are:

```
Line 259: (void) strcat(rfname, ".rmap");
```

```
Line 261: printf("%s.rmap contains the region map image for tolerance %f\n\n",
```

The `segment` function operates on the band interleaved binary layer and texture image. The algorithm and its inputs are described in detail by Woodcock and Harward (1992). It is appropriate, however, to mention briefly the arguments used in the SERCAA Cloud Typing and Layering Algorithm. As an example, the command, with arguments, might appear as:

```
segment.ext -t1 -o seg.layer.1 -m.10 -n1 -M layer.1 layer.1.tmux
```

where:

- t Tolerance that determines the minimum spectral similarity between all the neighbors of each region
- o Prefix to be attached to the output filenames
- m Merge coefficient which restricts the degree of merging on each pass ($0 < m \leq 1$)
- n The minimum number of pixels required in all final regions after the tolerance (-t) has been reached
- M Use a mask image in which any pixels that correspond to a 0 in the mask are prevented from merging with any region

A tolerance of 1 is chosen to ensure that only pixels with identical DNs in one or both bands (i.e., the binary layer image and the texture derivative image) or identical DNs in one band and a difference of no more than 1 DN in the other are merged. The merge coefficient has been set low (0.1) in the current implementation to ensure the careful growth of regions. In an operational scheme, this merge coefficient can be raised to reduce processing time. The minimum number of pixels in a region is set to 1 while there is no restriction placed on the maximum number of pixels. Finally, the binary layer image is used as a mask to prevent any pixels belonging to other layers from being included in the region growing process. Output from the segmentation routine includes a region map image in which each pixel possesses the DN of the region to which it belongs.

A histogram of the region map is generated using the IPW programs `hist` and `xyhist`. The program `hist` reads an IPW image and generates a histogram which is written to the standard output as a single-line IPW image. The program `xyhist` reads the histogram from the standard input and writes out a two-column histogram in ASCII format. The first column indicates the region ID and the second indicates the number of pixels belonging to that region. This forms what is referred to as a region table. The cumuliform/stratiform decision is made by sorting the table into two groups based on the number of pixels per region. The thresholds used to sort the regions vary depending on satellite source, but represent approximately the same physical size for each IFOV. They were established empirically through inter comparison with manual analysis. The thresholds for each satellite are: AVHRR = 500; DMSP/OLS = 1100; GMS = 320; METEOSAT = 320; GOES = 500. Smaller regions (i.e., those not exceeding the threshold) are labeled cumuliform and are remapped to a DN of 1, while larger regions are labeled stratiform and assigned a DN of 2. The labels are stored as an IPW look-up table and are applied to the region map by again applying `interp`, `mklut`, and `lutx` as described above. This procedure produces a cumuliform/stratiform mask image.

Each height layer is processed in order starting with the highest layer (ω_1) and working down. Processing of successive layers, however, requires two additional steps. Initially, the binary mask for a lower layer is added (logical OR) with the masks for all preceding (higher) layers (e.g., when processing the layer associated with ω_3 , the masks for layers ω_1 and ω_2 , are first added the mask for ω_3). This forms a composite binary mask (or composite layer) from which a texture derivative is calculated and on which `segment` is run. Combining layers in this way ensures that there are contiguous areas in the binary images that would otherwise be missing due to obstruction from a higher layer. Otherwise the `segment` algorithm would interpret these "holes" as region boundaries and produce an artificial ringing-effect in the cumuliform/stratiform mask.

The second added step is the extraction of pixels belonging to the current layer from the segmented composite image. This is accomplished by applying the original binary layer mask (logical AND) to the composite image such that only pixels belonging to the current layer are retained. The result is a cloud type image for each layer containing up to three digital numbers: 0 (clear pixels and those not contained in the layer), 1 (cumuliform) and 2 (stratiform). The resulting cloud type images for all layers are combined (logical OR) to form a composite cloud type map for all pixels in the scene.

To illustrate a case, processing the second layer in a height stratified image is explained here. From the classified (height stratified) image, a binary mask of Layer 2 (ω_2) is formed by applying a look-up table to the classified image. In this example, the look-up table would appear as follows:

0	0
1	0
2	1
3	0

The binary mask previously generated for Layer 1 is combined with the Layer 2 mask (using a logical OR statement) to form a composite binary image. With the union of the two layers taken, a texture derivative is calculated. The texture image and the composite binary image are band-interleaved by pixel and the segmentation routine is run. A histogram is generated of the resulting region map and sorted into cloud type according to region size. A region table is created indicating which region belongs to which cloud type (1 = cumuliform; 2 = stratiform) and is applied to the region map. The result is a cloud type image for Layers 1 and 2 combined. To isolate the cloud type information for Layer 2 (because the binary image on which these procedures are run is a composite of both Layers 1 and 2), the original binary mask for Layer 2 is applied (logical AND) to the cloud type image just formed. The result is a cloud type image in which pixels with DN's greater than 0 belong to Layer 2. This process is repeated for all layers generated by the classification routine. The cloud type images for all layers in the height stratified image are combined (logical OR) to produce the final cloud type map for the image.

6.3 CONVERSION FROM SENSOR TO POLAR STEREOGRAPHIC PROJECTION

All processing to this point has been on the original satellite IFOVs in scan projection at the sensor resolution. The final step for determination of layer parameters and location requires mapping all cloud parameters to a polar stereographic $1/16^{\text{th}}$ mesh grid. To move from the satellite projection to a $1/16^{\text{th}}$ mesh grid, the i, j coordinate maps discussed in Section 6.1 are employed. These maps serve as a template to identify $1/16^{\text{th}}$ mesh grid cell membership for each image pixel. Cloud layer determination is performed by analyzing LWIR sensor data along with individual cloud analysis results and cloud type information, all mapped to the $1/16^{\text{th}}$ mesh projection. Also, a number of cloud layer parameters, including cloud fraction and cloud test confidence measures, are calculated by averaging or summing over all pixels contained within a $1/16^{\text{th}}$ mesh grid cell.

At this stage of the SERCAA processing, the available information associated with each pixel in an original image (independent of satellite) is summarized in Table 17. This information is obtained from the i, j coordinate map (Section 6.1), the LWIR image (Section 6.1), the cloud type image (Section 6.2.2), and the MCF file (Table 16). Remapping is accomplished by accumulating this pixel information into a Grid Cell

Table (GCT). The GCT is organized such that there is one GCT entry for every 1/16th mesh grid cell. Each entry is populated with the Table 17 attributes for all pixels whose i, j coordinates match those of the grid cell. Note that since the 1/16th mesh grid spacing (\cong 24 km) is generally greater than the spatial resolution of the satellite data, a single table entry will usually contain information describing more than one pixel.

Table 17. Pixel Attributes

Array Element	Description
i, j	1/16 th mesh grid coordinates
LWIR Value	digitized brightness temperature from LWIR sensor channel
MCF Value	Mask and Confidence flag (Table 16)
Cloud Type	Stratiform or Cumuliform

6.4 CLOUD LAYERING

SERCAA requirements specify a maximum of four floating layers for each 1/16th mesh grid cell. In Section 6.2.1, cloud layers were identified over large regions of a satellite image using clustering and classification routines to ultimately produce cloud type maps. Recall that the number of resulting layers found for each region typically varies between 7 and 14, independent of the satellite sensor. The probability that a single grid cell within one of these regions would contain four or more layers is not negligible. Also, layers defined over large regions reflect conditions at the scale of the scene and are not necessarily representative of the local area defined by a particular i, j grid cell.

To address these issues, a separate cloud layering procedure was developed to operate on the smaller, 1/16th mesh scale, while still minimizing artificial discontinuities between layers at grid cell boundaries. This is achieved by performing a second, local scale clustering and cloud layer classification process that operates on the cloudy pixels within a floating 3 x 3 window of 1/16th mesh grid cells centered on the grid cell currently being analyzed. The process is conceptually similar to the large-scale layering procedure described in Section 6.2.1; unsupervised clustering and maximum likelihood classification are performed using the LWIR pixels that fall within the floating window. However, rather than the IPW functions *ustats* (unsupervised clustering) and *bayes* (maximum likelihood classifications), a simplified unsupervised clustering routine with a migrating means algorithm is used.

The layer analysis for each 16th mesh grid cell requires the Table 17 information in the contained in the GCT entries for all grid cells in the 3 x 3 floating window centered on that cell. The LWIR values for all cloudy pixels contained within the window region are retrieved and sorted in increasing order. For example, a hypothetical data set of 29 pixels located within a given window produces the following sorted distribution of LWIR DN values:

1 1 2 2 2 4 5 7 7 8 9 10 10 10 11 17 17 17 18 18 18 18 19 19 20 21 21 24 25

The initial step in the clustering process is to provide a first guess at the location of cluster centers in the sorted DN data. The first cluster center is taken as the pixel with a DN that is just less than a user-specified radius away from the minimum DN in the sorted pixel array. In the example, if the cluster radius (*r*) were, say, 3 then the first cluster

center would be chosen at DN=4, since the next DN (5) is more than 3 units from the minimum DN (1):

1 1 2 2 2 4 5 7 7 8 9 10 10 10 11 17 17 17 18 18 18 18 19 19 20 21 21 24 25
 ^#1

From this point the clustering algorithm lays out the remaining first-guess cluster centers such that their minimum separation distance is at least $2r$ units. Thus, in the example they would be placed at the following locations in the DN array:

1 1 2 2 2 4 5 7 7 8 9 10 10 10 11 17 17 17 18 18 18 18 19 19 20 21 21 24 25
 ^#1 ^#2 ^#3 ^#4

Selection of the last cluster center requires special processing. If the last element in the sorted data distribution is more than r units from the previously chosen cluster, it is used as the last cluster center. In the example above, if 21 were the last element in the array, it would qualify as a new cluster center even though it is not $2r$ units from the last cluster center, 17.

Note that the cluster radius only defines a minimum allowable separation between initial cluster centers. If the range of DN values is large relative to the minimum radius then the clusters will not be spread evenly over the data range. Consequently, for a given set of data to be clustered the algorithm uses either the specified cluster radius (currently defined as 8 for OLS and 4 for all other sensors) or one-eighth of the data range, whichever is larger. One-eighth of the data range is used because it is the magnitude of the cluster radius that would just fit four (the maximum allowable number of layers) non-overlapping classes in the data set. Even in these situations, however, it is likely that less than four cluster centers will be identified if the data range is small or there are few pixels within the window.

After the initial cluster centers have been chosen, pixel values are assigned to clusters based on the center DN value to which they are closest. From this first guess, cluster definitions are refined based on a two-step iterative migrating means process. The first step is to redefine cluster centers as the mean value of the member pixel DNs. The second step is to reassign pixels to clusters based on distance from the newly defined cluster centers. After each iteration the sum of the square of the deviations of pixel values from their cluster centers is calculated. The magnitude of this statistic decreases with each iteration as clusters become more internally similar and more externally dissimilar. The process continues until no further change is found in the deviation statistic. Once final clusters have been established, the mean and variance of each are calculated. At this point, cluster attributes are assumed to be representative of individual cloud layers.

The moving 3×3 window of $1/16^{\text{th}}$ mesh grid cells is used to establish local cloud layer attributes for the center grid cell in order to minimize artificial layer discontinuities across grid cell boundaries. Cloudy pixels in the center grid cell only are now assigned to a unique cloud layer using a maximum likelihood classifier (MLC). The MLC relies on the mean and variance (calculated as described above from the final clusters established over the broader 3×3 window) to define the layer attributes. The classifier used in this application has the same governing equation used in the bayes program described in Section 6.2.1 (Eq. 30). Note that when only one cloud-filled pixel is located in a grid cell the computed variance is 0.0; in this case a default value of 0.1 is substituted, otherwise calculation of the discriminant in Eq. 30 would have 0.0 as a

denominator. Final cloud layer statistics are then computed using only the pixels contained in the center grid cell; they include the mean, variance, and number of layers.

6.5 OUTPUT PRODUCT

The cloud parameters produced for each 1/16th mesh grid cell are listed in Table 18. The Grid Cell Table described in Section 6.3 is used to determine which satellite pixels belong to a given grid cell. Output cloud parameters are calculated from the information maintained for each member pixel in the grid cell table (Table 17) and the output of the layering process described above in Section 6.4. These parameters are used subsequently as input to the Analysis Integration Algorithm (see Fig. 1) in which results from multiple satellite platforms obtained at different times are integrated into a single consistent cloud analysis.

Layer summary statistics are calculated only for pixels classified as cloudy by the appropriate cloud analysis algorithm with the exceptions of the total number of pixels in the layer (NPL) and the total number of pixels in the grid cell (NPIX). The total number of pixels in a layer, the mean IR temperature (CTT) and variance (CTTV) are computed directly by the cloud layering algorithm (Section 6.4). Cloud type (TYP) is determined based on plurality (i.e., each layer is assigned the cloud type that most frequently occurs in the pixels belonging to that layer). Although the analysis algorithms provide a confidence flag for all pixels in a scene, only the confidence flags for pixels analyzed as cloudy are used to calculate the mean confidence flag for the layer (ICF).

In addition to layer statistics, Table 18 also contains parameters that apply globally to the entire grid cell. Both clear and cloudy pixels for all layers are used to calculate the total number of pixels (NPIX) and the total number of data dropouts (IDD) in a grid cell. Information on individual cloud types is obtained from the cloud analysis specific MCF output products (Table 16). Total number of pixels classified by the respective analysis algorithms as low cloud (LCC), thin cirrus (TCC), precipitating cloud (PCC), and partial cloud (PTC) are summed over all members of the grid cell as defined by the grid cell table.

Table 18. Cloud Typing and Layering Output

Column	Parameter	Description
1		i-Coordinate for Grid Cell
2		j-Coordinate for Grid Cell
3		Layer of Grid Cell for Which the Statistics Pertain
4	(CTT)	Cloud Top Mean IR Temperature of Pixels in Layer
5	(CTTV)	Cloud Top IR Temperature Variance of Pixels in Layer
6	(NPL)	Total Number of Pixels in Layer
7	(NPIX)	Total Number of Pixels in Grid Cell
8	(IDD)	Total Number of Data Dropouts in Grid Cell
9	(TYP)	Cloud Type of Layer
10	(ICF)	Mean Confidence Flag for Layer
11	(LCC)	Total Number of Low Cloud Pixels Detected in Cloud Analysis
12	(TCC)	Total Number of Thin Cirrus Pixels Detected in Cloud Analysis
13	(PCC)	Total Number of Precipitating Cloud Pixels Detected in Cloud Analysis
14	(PTC)	Total Number of Partial Cloud Pixels Detected in Cloud Analysis

7. ANALYSIS INTEGRATION ALGORITHM DESCRIPTION

This section describes the analysis integration portion of the SERCAA program. In this stage of the algorithm, independent cloud analyses from one or more satellite platforms are integrated to produce a single optimum analysis.

The general conceptual approach to the integration problem is one that utilizes both rule-based concepts as well as principles from statistical objective analysis. The unique nature of the satellite-derived cloud parameters and constraints on computational complexity drive the way in which the data are processed. For example, some cloud parameters such as cloud type and number of layers are discrete quantities and cannot be "averaged" in any physically meaningful way. Computational concerns also argue for applying rule-based ideas that allow the preferential selection of one satellite analysis over all others, and avoiding weighted averaging of the data as much as possible. The integration technique employed here is a blend of rules and a simplified optimum interpolation technology, described in detail by Hamill and Hoffman (1993).

Cloud parameters that are integrated during this stage are: total cloud fraction (CFT), layer cloud fraction (CF), layer cloud top IR temperature (CTT), layer cloud type (ITY), number of cloud layers (NLAY, up to 4 floating layers), and analysis confidence flag index (ICF). In addition, indices for the detection of thin cirrus cloud (ICI), precipitating cloud (ICB), and low cloud (ILO), the estimated error in total cloud fraction (ECFT), estimated error in layer cloud fraction (ECF), and local standard deviation of the analyzed cloud top IR temperature (CTTSD) are used during the integration, but are not themselves quantities to be integrated. While ICF and ICB are derived for all sensor analyses, the parameters ICI and ILO are only provided for analyses generated from NOAA/AVHRR and GOES sensor data; all confidence and cloud indices are based on cloud tests performed during the sensor-specific cloud analysis processing (refer to Sections 3 through 5). Table 19 contains a summary of all parameters processed in the Analysis Integration Algorithm; the final output parameters produced by the algorithm are: NLAY, CFT, CF, CTT, ICF, ITY, ECFT, and ECF.

With the exception of the estimated cloud fraction errors (ECF and ECFT), all integrated cloud parameters are computed within the integration module based on the statistics described in Table 18 that are output by the Cloud Typing and Layering Algorithm (refer to Section 6). CFT is calculated from the output of the layering process by summing the number of cloudy pixels in each layer (NPL) and dividing by the total number of pixels in the grid cell (NPIX). Similarly, CF is calculated by dividing the number of cloudy pixels in a layer (NPL) by NPIX. ITY is specified by combining layer height information with the cumuloform/stratiform determination (TYP) of the layering/typing step. NLAY is also determined within the Analysis Integration Algorithm from the number of cloud layer records present at each 1/16th mesh grid cell, as specified in the output of the Cloud Typing and Layering Algorithm. ILO, ICI, and ICB are indices representing the average value of the low cloud, thin cirrus, and precipitating cloud flags over all the cloudy pixels in each layer. Specifically, $ILO = LCC/NPL$, $ICI = TCC/NPL$, and $ICB = PCC/NPL$.

Estimated error statistics, ECF and ECFT, are determined from the sensor-dependent analysis error growth models discussed in Section 7.4. Based on the estimated errors of the input analyses and the type of blending used (e.g., optimum interpolation), an estimate of analysis error for the integrated analysis of total and layer cloud fraction is produced by the integration algorithm.

Table 19. Analysis Integration Processed Parameters

Parameter	Description	Dimensions	Integrated Product
NLAY	Number of Cloud Layers	NX*NY	Yes
CFT	Total Cloud Fraction	NX*NY	Yes
CF	Layer Cloud Fraction	NX*NY*NZ	Yes
CTT	Layer Cloud Top IR Temperature	NX*NY*NZ	Yes
ICF	Analysis Confidence Flag Index	NX*NY*NZ	Yes
ITY	Layer Cloud Type	NX*NY*NZ	Yes
ECFT	Estimated Error in Total Cloud Fraction	NX*NY	Yes
ECF	Estimated Error in Layer Cloud Fraction	NX*NY*NZ	Yes
CTTSD	Local Standard Deviation of Analyzed Cloud Top IR Temperature	NX*NY*NZ	No
ICB	Precipitating Cloud Detection Index	NX*NY*NZ	No
ICI	Thin Cirrus Cloud Detection Index	NX*NY*NZ	No
ILO	Low Cloud Detection Index	NX*NY*NZ	No

NX=number of columns in analysis grid
 NY=number of rows in analysis grid
 NZ=maximum number of layers (4)

During the layering and typing process described in Section 6, cloud analysis data derived from high resolution satellite imagery are analyzed and remapped to produce gridded analyses of all relevant cloud parameters on a relatively low resolution 1/16th mesh polar stereographic grid. Resolution of the satellite data varies with sensor characteristics and viewing geometry, thus the number of image pixels that fall within any particular 1/16th mesh analysis grid cell also varies. Sensor-dependent acceptance thresholds are applied to minimize under sampling, such that analysis parameters are set to missing for any grid cell with less than the minimum required pixels (clear + cloudy). Table 20 shows the minimum required number of pixels for each satellite.

Table 20. Grid Box Minimum Pixel Requirements

Satellite	Minimum Number of Pixels
NOAA/AVHRR	5
DMSP/OLS	10
GEOSTATIONARY ¹	5

¹ i.e., GOES, METEOSAT, GMS

The Analysis Integration Algorithm blends all available cloud analyses to form a single optimum analysis of the cloud parameters contained in Table 19. The integration is performed on cloud analysis products derived separately from each satellite platform. Thus, the analyzed cloud parameters are obtained using algorithms tailored to extract the maximum information contained in the sensor data of each satellite.

Figure 23 contains a flow diagram illustrating the rule-based approach and the data flow through the integration module for both total and layer cloud parameters. Note that since data availability and quality vary from one grid point to another the flow

diagram represents the data processing stream for a single 1/16th mesh analysis grid box. All analysis points are treated independently during the integration procedure.

7.1 INTEGRATION OF TOTAL CLOUD FRACTION

The integration of total cloud amount precedes integration of layer quantities since: 1) the estimates of total cloud fraction are believed to be more reliable than any individual layer fraction (due to small sample sizes and height assignment errors), and 2) the total cloud fraction for a 1/16th mesh grid point will constrain the sum of the layer cloud fractions that come out of the layer integration step (refer to Eq. 35).

A key element governing use of the satellite analyses within the integration algorithm is that of data timeliness. Timeliness is defined in terms of the actual age of the satellite analysis and a timeliness criterion, such that any analysis older than the timeliness criterion is no longer timely and, therefore, not used by the integration algorithm.

As shown in Fig. 23, the first step is to read in the previous integrated analysis (if available), along with any new satellite analyses. If it is determined that no new analyses exist, the old analysis is persisted. If new analyses are available, a check is made to determine if more than one are timely. If only one timely analysis is available, total cloud fraction and the estimated analysis error of total cloud are set to the value of this analysis.

If more than one analysis satisfies timeliness requirements, these analyses are examined to determine if all the analyses are completely cloudy. If so, total cloud fraction is set to 100 percent and the estimated analysis error is computed based on the formalism of optimum interpolation (OI) as discussed in Section 7.6. If all timely analyses are not totally cloudy, they are examined to determine if they are all completely clear. If they are clear, the total cloud fraction is set to 0 percent and the estimated analysis error is again computed using OI. In this case all other cloud layer parameters are left as missing.

If the analyses are neither all completely clear nor completely cloudy, the estimated error of each sensor analysis (refer to Section 7.4 for details) is examined to determine if the most recent analysis also has the lowest estimated error (and is therefore the "best" analysis). If this is true, then the total cloud fraction and estimated error is set to this analysis. Finally, if the best analysis is not the most recent, an OI algorithm is applied to obtain a blended estimate of total cloud fraction and analysis error (see Section 7.6).

7.2 INTEGRATION OF LAYER CLOUD PARAMETERS

Once integration of total cloud fraction is complete, integration of cloud layer parameters is performed. In cases where a single, optimal analysis can be identified from the multiple satellite analyses, the layer parameter integration follows the total cloud integration procedure described above. Thus in cases of only one timely analysis, all analyses indicating clear conditions, or when the most timely analysis is also the most accurate, then the integrated layer cloud parameters are simply the layer parameters of the selected most accurate analysis. However, in cases where multiple satellite analyses are combined to produce the integrated analysis, then the layer integration algorithm departs from the total cloud procedure.

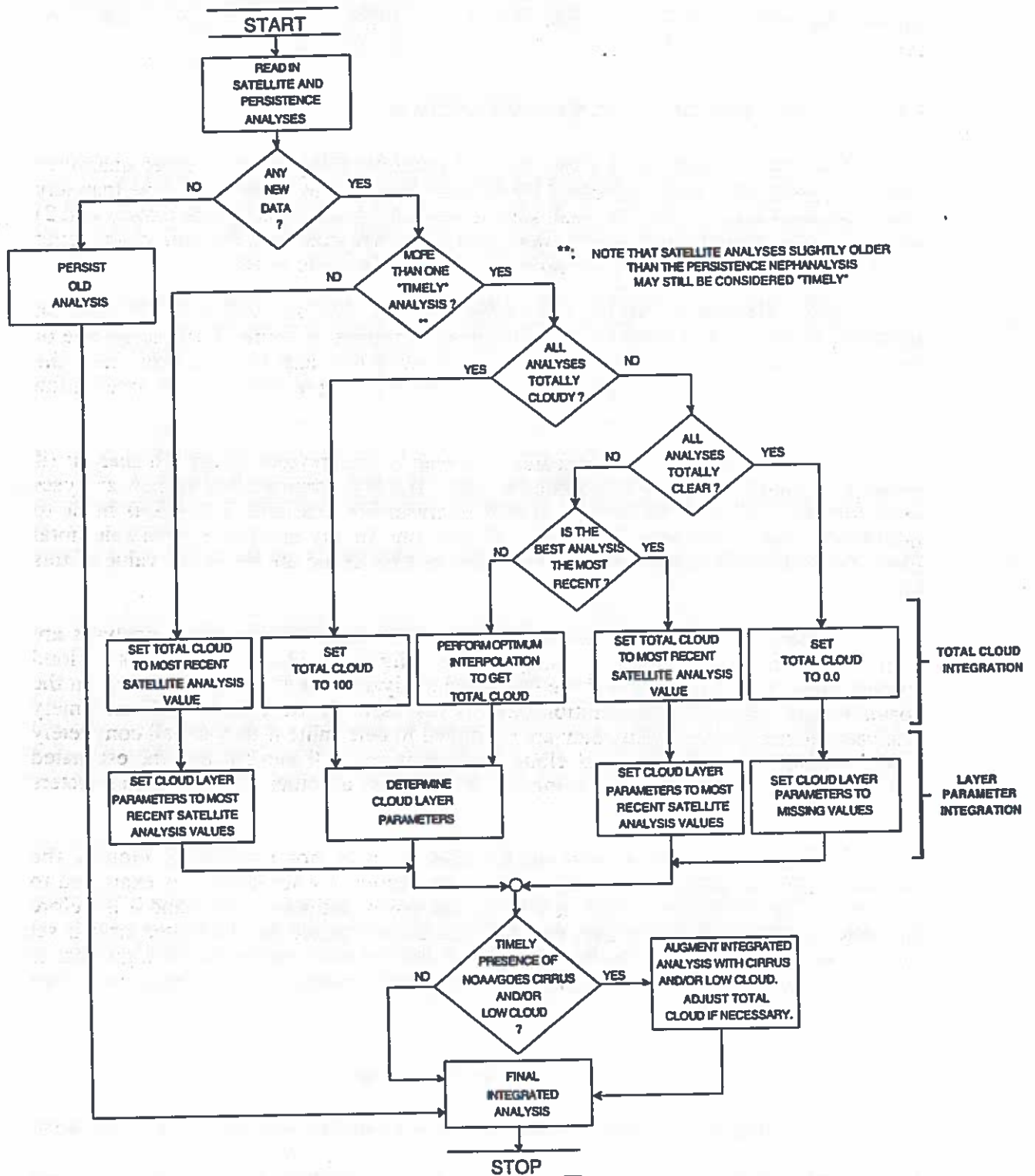


Figure 23. Cloud Analysis Integration Functional Flow

As indicated in Fig. 23, there are two cases when the layer integration algorithm differs from the total cloud integration procedure: 1) when there are multiple timely analyses that indicate 100 percent cloud cover, and 2) when the most recent analysis does not have the lowest estimated error. The reason these are special cases is that the vertical distribution of cloudiness and type is likely to vary among multiple cloudy analyses derived from different satellites and these differences need to be resolved to produce the single integrated analysis. In situations of multiple input analyses, the layer integration algorithm selects one analysis as a master profile by identifying the most recent of the timely analyses that also contains a non-zero cloud amount. For the discrete quantities of cloud type and number of layers the integrated analysis profiles take on the values of the master analysis. For the continuously varying parameters of layer cloud fraction, cloud top temperature and confidence flag index, an OI blending is performed by matching layers in the other timely analyses with the levels in the master analysis to which they are closest. Distance is calculated using the distance metric given by Eq. 31.

Figure 24 illustrates a case of multiple timely analyses with different cloud layer properties. Assume analysis A was found to be most timely but with a higher estimated analysis error than analysis B. According to the rule-based approach described above, sensor A is chosen as the master analysis into which the sensor B analysis is blended. The number of cloud layers and associated cloud types in the integrated analysis are set to those of sensor A. The layer properties of cloud amount and height are obtained by matching levels in analysis B with those in A to which they are nearest. The distance metric is a function of both the mean and variance of the analyzed layer temperatures (Richards, 1986) and is expressed as:

$$d_{ij} = \left(\frac{1}{\text{CTTSD}_i^2} + \frac{1}{\text{CTTSD}_j^2} \right) (\text{CTT}_i - \text{CTT}_j)^2 \quad (31)$$

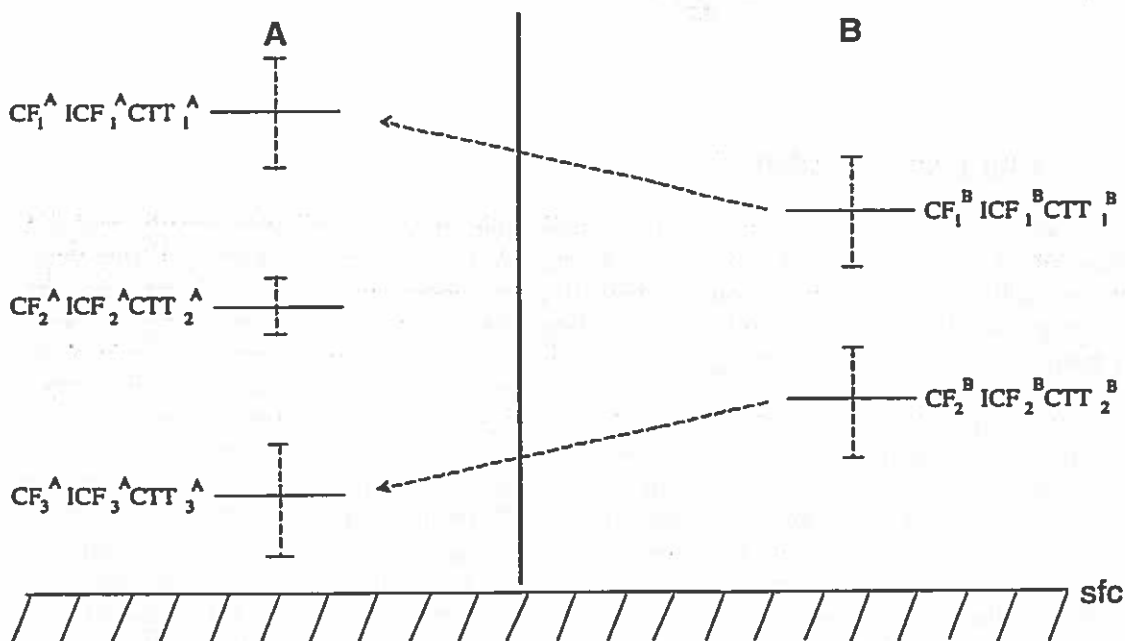


Figure 24. Cloud Analysis Integration Example

where d_{ij} is the square of the distance (normalized by variance) between the mean temperatures of layers i and j . Thus in some situations layers with a high cloud top temperature variance may be associated with layers different than would be the case if only temperature means were used. In the example cited in Fig. 24, if index i is used to refer to layers in the master analysis, A, then index j will refer to the analysis B layers.

Applying this approach to the example in Fig. 24, the two layers analyzed in B are matched to analysis A using CTT and CTTSD information from both analyses. In this case, layers 1 and 2 in B are matched to layers 1 and 3, respectively in the master analysis (sensor A). CF and CTT are then blended as follows:

$$CF_1^{INT} = W_1^A CF_1^A + W_1^B CF_1^B \quad (32)$$

$$CF_2^{INT} = CF_2^A \quad (33)$$

$$CF_3^{INT} = W_3^A CF_3^A + W_2^B CF_2^B \quad (34)$$

The analysis weights (W) are determined from the OI equations (Section 7.6).

After blending, consistency is enforced between CF and CFT by rescaling all layer cloud fractions (CF) assuming no overlap of cloud layers, i.e.

$$CF_i' = \left(\frac{CFT}{CFSUM} \right) CF_i \quad (35)$$

where CF_i' is the rescaled cloud fraction in layer i and

$$CFSUM = \sum_{i=1}^{NLAY} CF_i \quad (36)$$

7.3 CLOUD TYPE ASSIGNMENT

Table 21 contains the nine cloud types that may be defined during analysis integration. Determination of cloud type is achieved by a combination of information made available to the integration algorithm from both the cloud analysis and the layering/typing algorithms, as well as from independent cloud height information. Cloud layer height can be estimated using information on the local temperature profile and the observed cloud top temperature. Operationally this would be available from the Upper Air Database maintained at AFGWC (see Section 2.2.3). In practice a climatological profile that varies with season and latitude was used during Phase I. Essentially, cloud height information can be combined with the cumuliform/stratiform assignment produced in the Cloud Typing and Layering Algorithm (TYP from Table 18) to obtain up to six cloud types: cirrus, cirrostratus, altocumulus, altostratus, cumulus, and stratus. Additional use of the cloud type flags ICB, ICI, and ILO available from the cloud analysis algorithms allow detection of cumulonimbus, cirroform, and low cloud, respectively. Note that insufficient information is currently available to permit assignment of either nimbostratus or stratocumulus.

Table 21. Cloud Types Processed During Analysis Integration

ITY	Cloud Type	Height Range (meters)
1	Cirrus	$z \geq 6700$
2	Cirrostratus	$z \geq 6700$
3	Alto cumulus	$2000 \leq z < 6700$
4	Altostratus	$2000 \leq z < 6700$
5	Stratocumulus	$z < 2000$
6	Stratus	$z < 2000$
7	Cumulus	$z < 2000$
8	Cumulonimbus	$z \geq 6700$
9	Nimbostratus	$z < 2000$

7.4 ESTIMATED ANALYSIS ERRORS

Analysis errors for total cloud and cloud layer fraction (ECFT and ECF, respectively) are estimated assuming an initial analysis error plus some additional error growth which is a linear function of time. The analysis error is thus estimated from:

$$E_a = E_0 + \left(\frac{dE}{dt}\right)\Delta t, \quad (37)$$

where E_0 is the estimated analysis error at the initial analysis time, dE/dt is the analysis error growth rate, Δt is the time difference between the sensor analysis time and the integrated analysis time, and E_a is the estimated analysis error at the integrated analysis time.

In the case when a first guess analysis is available based on persistence of a previous integrated analysis, the initial error itself is both a function of the estimated error in the integrated analysis and the difference between the valid times of the previous analysis and the current analysis. This assumes the same error growth rate used for the individual sensor analyses (Table 22).

The variables E_0 and dE/dt are sensor-dependent and have been derived from satellite-based cloud analyses for both polar orbiting and geostationary platforms. Presently, estimated errors for total and layer cloud fraction are assumed to be equal during analysis integration. Table 22 contains current values used during analysis integration. These represent estimates based on limited data and are subject to refinement. For example, a higher order error growth function, or dependence on cloud type may be incorporated.

Table 22. Estimated Analysis Errors

Analysis Source	E_0 (percent)	dE/dt (percent/hour)
NOAA/AVHRR	10	5
DMSP/OLS	15	5
GEOSTATIONARY ¹	20	5
First Guess From Previous Analysis Integration	Function of E_a and Δt	5

¹ i.e., GOES, METEOSAT, GMS

7.5 CONFIDENCE FLAG INDEX AND CLOUD TYPE INDEX

As described in the sections pertaining to the sensor-specific cloud analysis algorithms (refer to Sections 3 through 5), when an image pixel is found to be cloud-filled, a data confidence flag is set (1-3) based on the various cloud tests conducted. This flag indicates low, moderate, or high confidence in the cloud determination. The Cloud Typing and Layering Algorithm forms an analysis confidence flag index which is the average confidence flag derived from all the cloudy pixels in each cloud layer (ICF from Table 18). This is passed directly to the Analysis Integration Algorithm. Lying in the range 1 to 3 it provides an estimate of the reliability of the analyzed layer cloud fraction(s) in an analysis grid box.

Similarly, cloud type indices are computed within the integration algorithm for each of the specific cloud type flags generated by the cloud algorithms: low, thin cirrus, and precipitating cloud. The first two are only available from NOAA/AVHRR and GOES analyses. These indices are the fraction of pixels within the layer that have the respective cloud type (e.g., ILO = LCC/NPL; refer to Table 18). Analogous to the confidence flag index, cloud type indices can be interpreted as confidence flags for the detection of specific cloud types, since they reflect the degree to which specific cloud type tests were passed within the sensor-specific cloud analysis algorithms (Sections 3 through 5). Acceptance thresholds are applied to these cloud type indices in assigning the type cumulonimbus (refer to Section 7.3 above), and for adding AVHRR or VAS-derived cirrus and/or low cloud (refer to Section 7.7 below). Table 23 contains acceptance thresholds used for cloud flag indices. Derived values must be greater than or equal to the indicated thresholds for the diagnosis of cloud and/or type to be valid.

Table 23. Acceptance Thresholds

Index	Threshold
ILO	0.3 (range: 0-1)
ICI	0.3 (range: 0-1)
ICB	0.3 (range: 0-1)

7.6 OPTIMUM INTERPOLATION

As described above, in situations when one sensor analysis cannot be unambiguously selected as the most accurate and timely, the integration procedure will blend all timely analyses using OI. OI (Gandin, 1963; Schlatter, 1975; Lorenc, 1981) is a well-established procedure used for the objective analysis of common meteorological variables such as geopotential height and wind. OI provides a formalism for synthesizing multiple observations into a single consistent, accurate analysis. Accuracy is achieved by weighting the data based upon their error characteristics such that less accurate observations are assigned less weight relative to more accurate ones, with the objective of minimizing the root mean squared error at each analysis point. In the case when all input observations (i.e. the sensor analyses) are valid at the same location and there is no assumed correlation between the analysis errors of different sensors, the OI simplifies to:

$$CF_{INT} = \sum_{i=1}^{NANL} W_i * CF_i \quad (38)$$

where NANL is the number of input analyses, CF_{INT} is the integrated analysis, W_i is the weight of the i^{th} input analysis, and CF_i is the cloud fraction observed in the i^{th} input sensor analysis.

W_i are the OI averaging weights given by:

$$W_i = \frac{1}{\sum_{j=1}^{NANL} \left(\frac{ECF_i}{ECF_j} \right)^2} \quad (39)$$

where ECF_i is the estimated analysis error of the i^{th} input sensor analysis, and ECF_j are the estimated errors for each of the available sensor analyses which are to be blended. Thus the weight assigned to an analysis is inversely proportional to the square of its estimated error.

When the combination of timeliness and accuracy warrants, the OI blending is performed for CFT, CF, and CTT, and ICF with the weights (W_i) calculated from the estimated errors for total cloud fraction (Eq. 39). The same weights are used for the blending of each of these parameters.

Likewise, the estimated error of the OI analysis can be obtained from:

$$ECF_{INT} = \sqrt{\frac{1}{A}} \quad (40)$$

where

$$A = \sum_{i=1}^{NANL} \left(\frac{1}{ECF_i^2} \right) \quad (41)$$

Thus, both ECF and ECFT in the integrated analysis are computed using Eqs. 40 and 41.

7.7 CIRRUS AND/OR LOW CLOUD FROM NOAA/AVHRR AND GOES/VAS

The integration algorithm is designed to take advantage of the additional information provided by multispectral observations from the NOAA/AVHRR and GOES/VAS. In particular, cloud algorithms developed for these sensors are better able to detect the presence of low clouds as well as thin semi-transparent cirrus clouds relative to the other sensor systems available to SERCAA. Once the integration of all cloud parameters has been performed, the analysis grid box is then checked for the possible addition of these cloud types.

The process of accounting for information on cirrus and/or low cloud is straightforward. First, since these clouds are more reliably detected from the AVHRR and VAS data than from other sensor data, timeliness constraints are not as strict for cirrus and low cloud derived from these sources than for clouds analyzed from other sensors. Aside from AVHRR or VAS-detected cirrus and low cloud, any data older than 2 hours is no longer timely. For AVHRR or VAS-detected cirrus this timeliness constraint is set to 3 hours; for AVHRR or VAS-detected low cloud a 4-hour timeliness constraint is used. These thresholds are tunable and may be varied. To determine if

Table A-1. AVHRR Cloud Test Thresholds

Threshold Name	Cloud Detection Value	Cloud Clearing Value	Description
THRESH _{ci}	5.0 K	5.0 K	Cirrus Cloud Test snow/ice filter threshold
THRESH _{ci snow}	280 K	280 K	Cirrus Cloud Test potential snow background threshold
THRESH _{cold}	9.0 K	7.0 K	Cold Cloud Test threshold over water
	10.0 K	8.0 K	Cold Cloud Test threshold over land
	20.0 K	10.0 K	Cold Cloud Test threshold over coast
	10.0 K	10.0 K	Cold Cloud Test threshold over desert
	15.0 K	10.0 K	Cold Cloud Test threshold over snow
THRESH _{dci w}	0.2	N/A	Daytime Thin Cirrus Cloud Test threshold over water
THRESH _{dci l}	0.2	N/A	Daytime Thin Cirrus Cloud Test threshold over land
THRESH _{desert}	0.2	0.2	Desert Background Test reflectance threshold
THRESH _{desert lo diff}	7.0 K	7.0 K	Desert background Test lower limit channel difference threshold
THRESH _{desert lo diff}	17.0 K	17.0 K	Desert Background Test upper limit channel difference threshold
THRESH _{desert lo ratio}	0.85	0.85	Desert Background Test lower limit ratio threshold
THRESH _{desert up ratio}	1.05	1.05	Desert Background Test upper limit ratio threshold
THRESH _{fls}	1.0 K	0.6 K	Fog, Low Stratus Test threshold
	2.0 K	0.6 K	Fog, Low Stratus Test threshold over desert
THRESH _{glint(1)}	20.0 K	20.0 K	Sun Glint Test threshold
THRESH _{glint(2)}	309 K	309 K	Sun Glint Test threshold
THRESH _{land}	0.25	0.20	Visible Brightness Test threshold over land
THRESH _{lcf}	12.0 K	8.0 K	Low Cloud, and Fog Test threshold over non-desert
	20.0 K	15.0 K	Low Cloud, and Fog Test threshold over desert
	54.0 K	8.0 K	Low Cloud, and Fog Test threshold over potential sun glint regions
THRESH _{loazimuth}	120°	N/A	Lower azimuth threshold (Sun Glint Test)
THRESH _{precip(1)}	20.0 K	N/A	Precipitating Cloud Test threshold
THRESH _{precip(2)}	30.0 K	N/A	Precipitating Cloud Test threshold
THRESH _{precip(3)}	0.45	N/A	Precipitating Cloud Test threshold
THRESH _{ratio lo dry}	0.75	0.7	Visible Brightness Ratio Test lower limit threshold
THRESH _{ratio up dry}	1.1	1.15	Visible Brightness Ratio Test upper limit threshold
THRESH _{ratio humid}	295 K	295 K	Visible Brightness Ratio Test high humidity threshold
THRESH _{ratio lo wet}	0.70	0.70	Visible Brightness Ratio Test lower limit threshold (High Humidity)
THRESH _{ratio up wet}	1.0	1.15	Visible Brightness Ratio Test upper limit threshold (High Humidity)
THRESH _{snow(1)}	278 K	278 K	Snow/Ice Cover Background Test threshold
THRESH _{snow(2)}	9.0 K	9.0 K	Snow/Ice Cover Background Test threshold
THRESH _{snow(3)}	9.0 K	9.0 K	Snow/Ice Cover Background Test threshold
THRESH _{snow land}	0.2	0.2	Snow/Ice Cover Background Test threshold over land
THRESH _{snow water}	0.1	0.1	Snow/Ice Cover Background Test threshold over water
THRESH _{temp desert(1)}	300 K	300 K	Desert Background Test temperature threshold
THRESH _{temp desert(2)}	10 K	10 K	Desert Background Test temperature threshold
THRESH _{tc1}	4.0 K	3.0 K	Nighttime Thin Cirrus Cloud Test threshold
THRESH _{tc1 humid}	290 K	290 K	Nighttime Thin Cirrus Cloud Test high humidity threshold
THRESH _{upazimuth}	240°	240°	Upper azimuth threshold (Sun Glint Test)
THRESH _{water}	0.16	0.12	Visible Brightness Test threshold over water
THRESH _{zenith}	40.0°	40.0°	Zenith angle threshold (Sun Glint Test)
THRESH _(T4,y)	Refer to Section 3.2.3		Cirrus Cloud Test threshold

Table A-2a. AVHRR T4 - T5 Threshold Table

T4	1.00	1.25	1.50	1.75	2.00	sec (ψ)
260	0.55	0.60	0.65	0.90	1.10	
270	0.58	0.63	0.81	1.03	1.13	
280	1.30	1.61	1.88	2.14	2.30	
290	3.06	3.72	3.95	4.27	4.73	
300	5.77	6.92	7.00	7.42	8.43	
310	9.41	10.74	11.03	11.60	13.39	

Table A-2b. AVHRR T4 - T5 Threshold Table

Precipitable Water (cm)	Viewing Angle (deg)							Scan Angle Zenith Angle
	0	10	20	30	40	50	55	
0.452	0.0	0.1	0.1	0.0	0.0	0.0	0.0	
2.11	1.2	1.2	1.2	1.3	1.3	1.5	1.6	
4.20	1.8	1.8	1.8	1.9	2.0	2.2	2.3	

Table A-3. DMSP Cloud Test Thresholds

Threshold Name	Value		Description
R _{cld}	Water: 37	Land: See Section 4.1.2	Visible channel cloud threshold
R _{clr}	Water: 40	Land: See Section 4.1.2	Visible channel cloud threshold
THRESH _{DMSP solzen}	75°		Day/Night solar zenith angle threshold
THRESH _{loazimuth}	120°		Lower azimuth threshold (Potential sun glint)
THRESH _{precip}	230 K		Single Channel Precipitating Cloud Test threshold
THRESH _{upazimuth}	240°		Lower azimuth threshold (Potential sun glint)
THRESH _{zenith}	40°		Zenith angle threshold (Potential sun glint)
α_{cld}	1.5		Infrared channel cloud threshold offset value
α_{clr}	0.75		Infrared channel clear threshold offset value
ρ_{cld}	0.4		Visible channel cloud threshold adjustment over land
ρ_{clr}	0.1		Visible channel clear threshold adjustment over land

Table A-4. Geostationary Cloud Test Thresholds

Threshold Name	Value	Description
THRESH _{cold}	25 K	Cold Cloud threshold
THRESH _{geo_solzen}	85°	Day/Night solar zenith angle threshold
THRESH _{geo_pcp_solzen}	65°	Precipitating Cloud Test solar zenith angle threshold
THRESH _{land}	30 (counts)	Bright cloud over land threshold
THRESH _{loazimuth} ¹	150°	Lower azimuth threshold (Potential sun glint)
THRESH _{pcp_goes}	8 K	Precipitating Cloud IR threshold
THRESH _{pcp_vis}	170 (counts)	Precipitating Cloud visible threshold
THRESH _{water}	30 (counts)	Bright cloud over water threshold
THRESH _{DCi(1)}	10 K	Daytime Cirrus Cloud threshold
THRESH _{DCi(2)}	170 (counts)	Daytime Cirrus Cloud threshold
THRESH _{LCd}	8 K	Low Cloud and Fog threshold (Day application)
THRESH _{LCn}	2 K	Low Cloud and Fog threshold (Night application)
THRESH _(O,ψ)	See Section 5.3.2	Daytime Cirrus Cloud threshold
THRESH _{spectral_solzen}	85°	Spectral discriminant tests Day/Night threshold
THRESH _{TCi}	3 K	Thin Cirrus Cloud threshold
THRESH _{td_pct}	1 (percent)	Minimum threshold count threshold
THRESH _{upazimuth} ¹	210°	Upper azimuth threshold (Potential sun glint)
THRESH _{zenith} ¹	15°	Zenith angle threshold (Potential sun glint)
δ _{IR}	6 K	Infrared Temporal Differencing threshold
δ _{VIS}	4 (counts)	Visible Temporal Differencing threshold
γ	0.3	Dynamic Threshold Test adjustment value

¹These limits have not yet been fully tested.

APPENDIX B

ACRONYMS

AFGWC	Air Force Global Weather Central
ASCII	American Standard Code for Information Interchange
AVHRR	Advanced Very High-Resolution Radiometer
DMA	Defense Mapping Agency
DMSP	Defense Meteorological Satellite Program
DN	Digital Number
EBBT	Equivalent Black Body Brightness Temperature
EUMETSAT	European Organization for the Exploration of Meteorological Satellites
FOV	Field of View
GCT	Grid Cell Table
GMS	Geostationary Meteorological Satellite, Japan
GOES	Geostationary Operational Environmental Satellite
IFOV	Instantaneous Field of View
IPW	Image Processing Workbench
IR	Infrared
LWIR	Long-Wave Infrared
MCF	Mask and Confidence Flag file
METEOSAT	Geostationary Satellite (EUMETSAT)
MLC	Maximum Likelihood Classifier
MWIR	Middle-Wave Infrared
NOAA	National Oceanic and Atmospheric Administration
OI	Optimum Interpolation
OLS	Operational Linescan System
PL	Phillips Laboratory
RTNEPH	Real-Time Nephanalysis
SERCAA	Support of Environmental Requirements for Cloud Analysis and Archive
SFCTMP	AFGWC Surface Temperature Model
SSM/I	Special Sensor Microwave Imager
TACNEPH	Tactical Nephanalysis
UTC	Universal Time Coordinated
VAS	Visible Infrared Spin Scan Radiometer and Atmospheric Sounder
VBC	Visible Background Count database
VISSR	Visible Infrared Spin Scan Radiometer