NASA Reference Publication 1376
Volume III

Clouds and the Earth's Radiant Energy System (CERES) Algorithm Theoretical Basis Document

Volume III—Cloud Analyses and Determination of Improved Top of Atmosphere Fluxes (Subsystem 4)

CERES Science Team
Langley Research Center • Hampton, Virginia

National Aeronautics and Space Administration
Langley Research Center • Hampton, Virginia 23681-0001

December 1995

Available electronically at the following URL address: http://techreports.larc.nasa.gov/ltrs/ltrs.html

Printed copies available from the following:

NASA Center for AeroSpace Information
800 Elkridge Landing Road
Linthicum Heights, MD 21090-2934
(301) 621-0390

National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield, VA 22161-2171
(703) 487-4650

Contents

Preface

The Release-1 CERES Algorithm Theoretical Basis Document (ATBD) is a compilation of the techniques and processes that constitute the prototype data analysis scheme for the Clouds and the Earth's Radiant Energy System (CERES), a key component of NASA's Mission to Planet Earth. The scientific bases for this project and the methodologies used in the data analysis system are also explained in the ATBD. The CERES ATBD comprises 11 subsystems of various sizes and complexities. The ATBD for each subsystem has been reviewed by three or four independently selected university, NASA, and NOAA scientists. In addition to the written reviews, each subsystem ATBD was reviewed during oral presentations given to a six-member scientific peer review panel at Goddard Space Flight Center during May 1994. Both sets of reviews, oral and written, determined that the CERES ATBD was sufficiently mature for use in providing archived Earth Observing System (EOS) data products. The CERES Science Team completed revisions of the ATBD to satisfy all reviewer comments. Because the Release-1 CERES ATBD will serve as the reference for all of the initial CERES data analysis algorithms and product generation, it is published here as a NASA Reference Publication.

Due to its extreme length, this NASA Reference Publication comprises four volumes that divide the CERES ATBD at natural break points between particular subsystems. These four volumes are

I: Overviews
- CERES Algorithm Overview
- Subsystem 0. CERES Data Processing System Objectives and Architecture

II: Geolocation, Calibration, and ERBE-Like Analyses
- Subsystem 1.0. Instrument Geolocate and Calibrate Earth Radiances
- Subsystem 2.0. ERBE-Like Inversion to Instantaneous TOA and Surface Fluxes
- Subsystem 3.0. ERBE-Like Averaging to Monthly TOA

III: Cloud Analyses and Determination of Improved Top of Atmosphere Fluxes
- Subsystem 4.0. Overview of Cloud Retrieval and Radiative Flux Inversion
- Subsystem 4.1. Imager Clear-Sky Determination and Cloud Detection
- Subsystem 4.2. Imager Cloud Height Determination
- Subsystem 4.3. Cloud Optical Property Retrieval
- Subsystem 4.4. Convolution of Imager Cloud Properties With CERES Footprint Point Spread Function
- Subsystem 4.5. CERES Inversion to Instantaneous TOA Fluxes
- Subsystem 4.6. Empirical Estimates of Shortwave and Longwave Surface Radiation Budget Involving CERES Measurements

IV: Determination of Surface and Atmosphere Fluxes and Temporally and Spatially Averaged Products
- Subsystem 5.0. Compute Surface and Atmospheric Fluxes
- Subsystem 6.0. Grid Single Satellite Fluxes and Clouds and Compute Spatial Averages
- Subsystem 7.0. Time Interpolation and Synoptic Flux Computation for Single and Multiple Satellites
- Subsystem 8.0. Monthly Regional, Zonal, and Global Radiation Fluxes and Cloud Properties
- Subsystem 9.0. Grid TOA and Surface Fluxes for Instantaneous Surface Product
- Subsystem 10.0. Monthly Regional TOA and Surface Radiation Budget
- Subsystem 11.0. Update Clear Reflectance, Temperature History (CHR)
- Subsystem 12.0. Regrid Humidity and Temperature Fields

The CERES Science Team serves as the editor for the entire document. A complete list of Science Team members is given below. Different groups of individuals prepared the various subsections that constitute the CERES ATBD. Thus, references to a particular subsection of the ATBD should specify

the subsection number, authors, and page numbers. Questions regarding the content of a given subsection should be directed to the appropriate first or second author. No attempt was made to make the overall document stylistically consistent.

The CERES Science Team is an international group led by 2 principal investigators and 19 coinvestigators. The team members and their institutions are listed below.

CERES Science Team

Bruce A. Wielicki, Interdisciplinary Principal Investigator
Bruce R. Barkstrom, Instrument Principal Investigator

Atmospheric Sciences Division
NASA Langley Research Center
Hampton, Virginia 23681-0001

Coinvestigators

Bryan A. Baum
Atmospheric Sciences Division
NASA Langley Research Center
Hampton, Virginia 23681-0001

Maurice Blackmon
Climate Research Division
NOAA Research Laboratory
Boulder, Colorado 80303

Robert D. Cess
Institute for Terrestrial & Planetary Atmospheres
Marine Sciences Research Center
State University of New York
Stony Brook, New York 11794-5000

Thomas P. Charlock
Atmospheric Sciences Division
NASA Langley Research Division
Hampton, Virginia 23681-0001

James A. Coakley
Oregon State University
Department of Atmospheric Sciences
Corvallis, Oregon 97331-2209

Dominique A. Crommelynck
Institute Royal Meteorologique
B-1180 Bruxelles
Belgium

Richard N. Green
Atmospheric Sciences Division
NASA Langley Research Center
Hampton, Virginia 23681-0001

Robert Kandel
Laboratoire de Meteorologie Dynamique
Ecole Polytechnique
91128 Palaiseau
France

Michael D. King
Goddard Space Flight Center
Greenbelt, Maryland 20771

Robert B. Lee III
Atmospheric Sciences Division
NASA Langley Research Center
Hampton, Virginia 23681-0001

A. James Miller
NOAA/NWS
5200 Auth Road
Camp Springs, Maryland 20233

Patrick Minnis
Atmospheric Sciences Division
NASA Langley Research Center
Hampton, Virginia 23681-0001

Veerabhadran Ramanathan
Scripps Institution of Oceanography
University of California-San Diego
La Jolla, California 92093-0239

David R. Randall
Colorado State University
Department of Atmospheric Science
Foothills Campus, Laporte Avenue
Fort Collins, Colorado 80523

G. Louis Smith
Atmospheric Sciences Division
NASA Langley Research Center
Hampton, Virginia 23681-0001

Larry L. Stowe
NOAA/NWS
5200 Auth Road
Camp Springs, Maryland 20233

Ronald M. Welch
South Dakota School of Mines and Technology
Institute of Atmospheric Sciences
Rapid City, South Dakota 57701-3995

Nomenclature

Acronyms

ADEOS	Advanced Earth Observing System
ADM	Angular Distribution Model
AIRS	Atmospheric Infrared Sounder (EOS-AM)
AMSU	Advanced Microwave Sounding Unit (EOS-PM)
APD	Aerosol Profile Data
APID	Application Identifier
ARESE	ARM Enhanced Shortwave Experiment
ARM	Atmospheric Radiation Measurement
ASOS	Automated Surface Observing Sites
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ASTEX	Atlantic Stratocumulus Transition Experiment
ASTR	Atmospheric Structures
ATBD	Algorithm Theoretical Basis Document
AVG	Monthly Regional, Average Radiative Fluxes and Clouds (CERES Archival Data Product)
AVHRR	Advanced Very High Resolution Radiometer
BDS	Bidirectional Scan (CERES Archival Data Product)
BRIE	Best Regional Integral Estimate
BSRN	Baseline Surface Radiation Network
BTD	Brightness Temperature Difference(s)
CCD	Charge Coupled Device
CCSDS	Consultative Committee for Space Data Systems
CEPEX	Central Equatorial Pacific Experiment
CERES	Clouds and the Earth’s Radiant Energy System
CID	Cloud Imager Data
CLAVR	Clouds from AVHRR
CLS	Constrained Least Squares
COPRS	Cloud Optical Property Retrieval System
CPR	Cloud Profiling Radar
CRH	Clear Reflectance, Temperature History (CERES Archival Data Product)
CRS	Single Satellite CERES Footprint, Radiative Fluxes and Clouds (CERES Archival Data Product)
DAAC	Distributed Active Archive Center
DAC	Digital-Analog Converter
DB	Database
DFD	Data Flow Diagram
DLF	Downward Longwave Flux

DMSP	Defense Meteorological Satellite Program
EADM	ERBE-Like Albedo Directional Model (CERES Input Data Product)
ECA	Earth Central Angle
ECLIPS	Experimental Cloud Lidar Pilot Study
ECMWF	European Centre for Medium-Range Weather Forecasts
EDDB	ERBE-Like Daily Data Base (CERES Archival Data Product)
EID9	ERBE-Like Internal Data Product 9 (CERES Internal Data Product)
EOS	Earth Observing System
EOSDIS	Earth Observing System Data Information System
EOS-AM	EOS Morning Crossing Mission
EOS-PM	EOS Afternoon Crossing Mission
ENSO	El Niño/Southern Oscillation
ENVISAT	Environmental Satellite
EPHANC	Ephemeris and Ancillary (CERES Input Data Product)
ERB	Earth Radiation Budget
ERBE	Earth Radiation Budget Experiment
ERBS	Earth Radiation Budget Satellite
ESA	European Space Agency
ES4	ERBE-Like S4 Data Product (CERES Archival Data Product)
ES4G	ERBE-Like S4G Data Product (CERES Archival Data Product)
ES8	ERBE-Like S8 Data Product (CERES Archival Data Product)
ES9	ERBE-Like S9 Data Product (CERES Archival Data Product)
FLOP	Floating Point Operation
FIRE	First ISCCP Regional Experiment
FIRE II IFO	First ISCCP Regional Experiment II Intensive Field Observations
FOV	Field of View
FSW	Hourly Gridded Single Satellite Fluxes and Clouds (CERES Archival Data Product)
FTM	Functional Test Model
GAC	Global Area Coverage (AVHRR data mode)
GAP	Gridded Atmospheric Product (CERES Input Data Product)
GCIP	GEWEX Continental-Phase International Project
GCM	General Circulation Model
GEBA	Global Energy Balance Archive
GEO	ISSCP Radiances (CERES Input Data Product)
GEWEX	Global Energy and Water Cycle Experiment
GLAS	Geoscience Laser Altimetry System
GMS	Geostationary Meteorological Satellite
GOES	Geostationary Operational Environmental Satellite
HBTM	Hybrid Bispectral Threshold Method

HIRS	High-Resolution Infrared Radiation Sounder
HIS	High-Resolution Interferometer Sounder
ICM	Internal Calibration Module
ICRCCM	Intercomparison of Radiation Codes in Climate Models
ID	Identification
IEEE	Institute of Electrical and Electronics Engineers
IES	Instrument Earth Scans (CERES Internal Data Product)
IFO	Intensive Field Observation
INSAT	Indian Satellite
IOP	Intensive Observing Period
IR	Infrared
IRIS	Infrared Interferometer Spectrometer
ISCCP	International Satellite Cloud Climatology Project
ISS	Integrated Sounding System
IWP	Ice Water Path
LAC	Local Area Coverage (AVHRR data mode)
LaRC	Langley Research Center
LBC	Laser Beam Ceilometer
LBTM	Layer Bispectral Threshold Method
Lidar	Light Detection and Ranging
LITE	Lidar In-Space Technology Experiment
Lowtran 7	Low-Resolution Transmittance (Radiative Transfer Code)
LW	Longwave
LWP	Liquid Water Path
LWRE	Longwave Radiant Excitance
MAM	Mirror Attenuator Mosaic
MC	Mostly Cloudy
MCR	Microwave Cloud Radiometer
METEOSAT	Meteorological Operational Satellite (European)
METSAT	Meteorological Satellite
MFLOP	Million FLOP
MIMR	Multifrequency Imaging Microwave Radiometer
MISR	Multiangle Imaging Spectroradiometer
MLE	Maximum Likelihood Estimate
MOA	Meteorology Ozone and Aerosol
MODIS	Moderate-Resolution Imaging Spectroradiometer
MSMR	Multispectral, multiresolution
MTSA	Monthly Time and Space Averaging
MWH	Microwave Humidity

MWP	Microwave Water Path
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NESDIS	National Environmental Satellite, Data, and Information Service
NIR	Near Infrared
NMC	National Meteorological Center
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
OLR	Outgoing Longwave Radiation
OPD	Ozone Profile Data (CERES Input Data Product)
OV	Overcast
PC	Partly Cloudy
POLDER	Polarization of Directionality of Earth's Reflectances
PRT	Platinum Resistance Thermometer
PSF	Point Spread Function
PW	Precipitable Water
RAPS	Rotating Azimuth Plane Scan
RPM	Radiance Pairs Method
RTM	Radiometer Test Model
SAB	Sorting by Angular Bins
SAGE	Stratospheric Aerosol and Gas Experiment
SARB	Surface and Atmospheric Radiation Budget Working Group
SDCD	Solar Distance Correction and Declination
SFC	Hourly Gridded Single Satellite TOA and Surface Fluxes (CERES Archival Data Product)
SHEBA	Surface Heat Budget in the Arctic
SPECTRE	Spectral Radiance Experiment
SRB	Surface Radiation Budget
SRBAVG	Surface Radiation Budget Average (CERES Archival Data Product)
SSF	Single Satellite CERES Footprint TOA and Surface Fluxes, Clouds
SSMI	Special Sensor Microwave Imager
SST	Sea Surface Temperature
SURFMAP	Surface Properties and Maps (CERES Input Product)
SW	Shortwave
SWICS	Shortwave Internal Calibration Source
SWRE	Shortwave Radiant Excitance
SYN	Synoptic Radiative Fluxes and Clouds (CERES Archival Data Product)
SZA	Solar Zenith Angle
THIR	Temperature/Humidity Infrared Radiometer (Nimbus)

TIROS	Television Infrared Observation Satellite
TISA	Time Interpolation and Spatial Averaging Working Group
TMI	TRMM Microwave Imager
TOA	Top of the Atmosphere
TOGA	Tropical Ocean Global Atmosphere
TOMS	Total Ozone Mapping Spectrometer
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
TSA	Time-Space Averaging
UAV	Unmanned Aerospace Vehicle
UT	Universal Time
UTC	Universal Time Code
VAS	VISSR Atmospheric Sounder (GOES)
VIRS	Visible Infrared Scanner
VISSR	Visible and Infrared Spin Scan Radiometer
WCRP	World Climate Research Program
WG	Working Group
Win	Window
WN	Window
WMO	World Meteorological Organization
ZAVG	Monthly Zonal and Global Average Radiative Fluxes and Clouds (CERES Archival Data Product)

Symbols

A	atmospheric absorptance
$B_\lambda(T)$	Planck function
C	cloud fractional area coverage
CF_2Cl_2	dichlorofluorocarbon
$CFCl_3$	trichlorofluorocarbon
CH_4	methane
CO_2	carbon dioxide
D	total number of days in the month
D_e	cloud particle equivalent diameter (for ice clouds)
E_o	solar constant or solar irradiance
F	flux
f	fraction
G_a	atmospheric greenhouse effect
g	cloud asymmetry parameter
H_2O	water vapor

I	radiance
i	scene type
m_i	imaginary refractive index
$\hat{N}$	angular momentum vector
N_2O	nitrous oxide
O_3	ozone
P	point spread function
p	pressure
Q_a	absorption efficiency
Q_e	extinction efficiency
Q_s	scattering efficiency
R	anisotropic reflectance factor
r_E	radius of the Earth
r_e	effective cloud droplet radius (for water clouds)
r_h	column-averaged relative humidity
S_o	summed solar incident SW flux
S_o'	integrated solar incident SW flux
T	temperature
T_B	blackbody temperature
t	time or transmittance
W_{liq}	liquid water path
w	precipitable water
$\hat{x}_o$	satellite position at t_o
x, y, z	satellite position vector components
$\dot{x}, \dot{y}, \dot{z}$	satellite velocity vector components
z	altitude
z_{top}	altitude at top of atmosphere
α	albedo or cone angle
β	cross-scan angle
γ	Earth central angle
γ_{at}	along-track angle
γ_{ct}	cross-track angle
δ	along-scan angle
ε	emittance
Θ	colatitude of satellite
θ	viewing zenith angle
θ_o	solar zenith angle
λ	wavelength
μ	viewing zenith angle cosine

μ_o	solar zenith angle cosine
ν	wave number
ρ	bidirectional reflectance
τ	optical depth
$\tau_{aer}(p)$	spectral optical depth profiles of aerosols
$\tau_{H_2O\lambda}(p)$	spectral optical depth profiles of water vapor
$\tau_{O_3}(p)$	spectral optical depth profiles of ozone
Φ	longitude of satellite
ϕ	azimuth angle
$\tilde{\omega}_o$	single-scattering albedo

Subscripts:

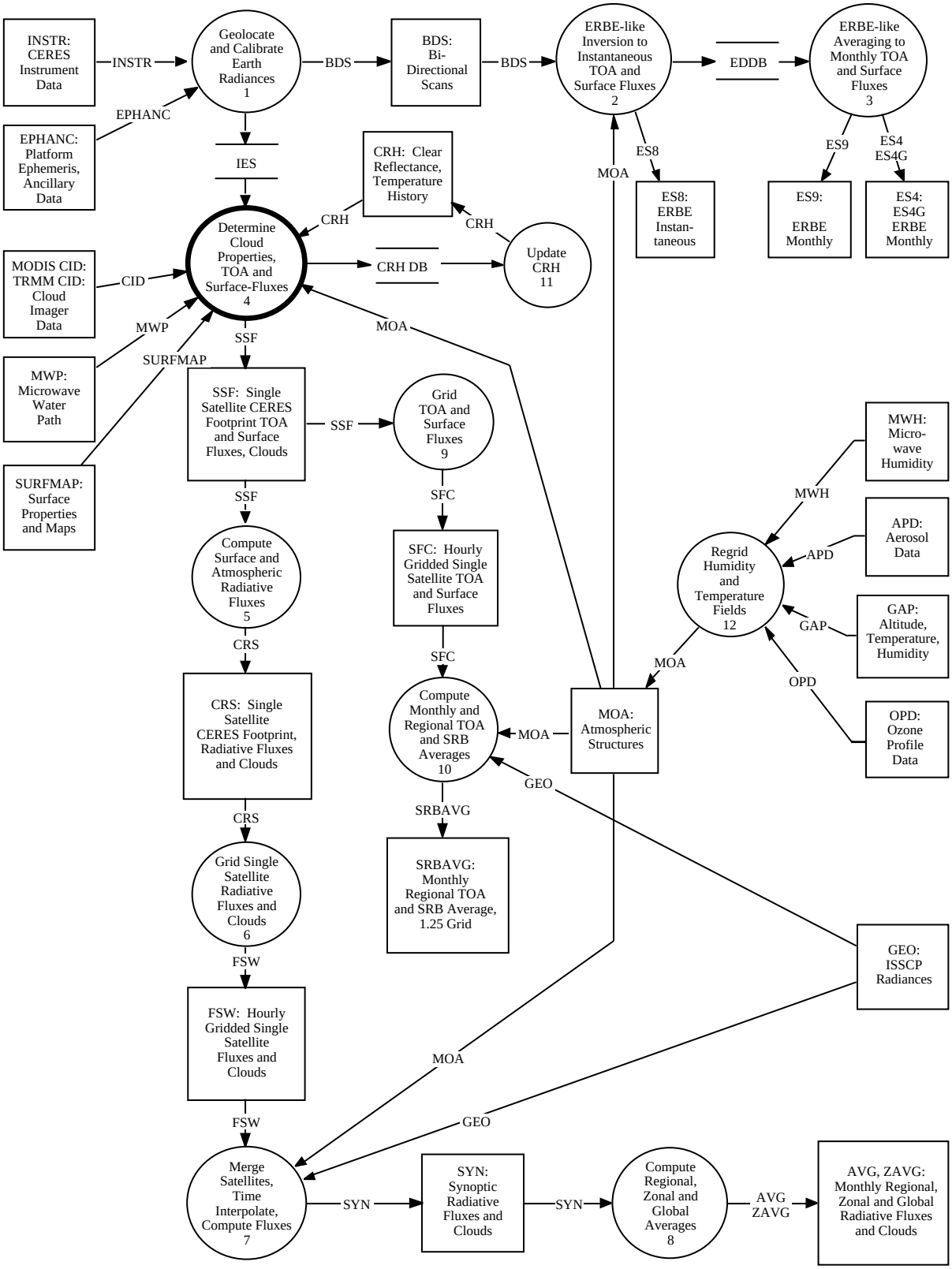
c	cloud
cb	cloud base
ce	cloud effective
cld	cloud
cs	clear sky
ct	cloud top
ice	ice water
lc	lower cloud
liq	liquid water
s	surface
uc	upper cloud
λ	spectral wavelength

Units

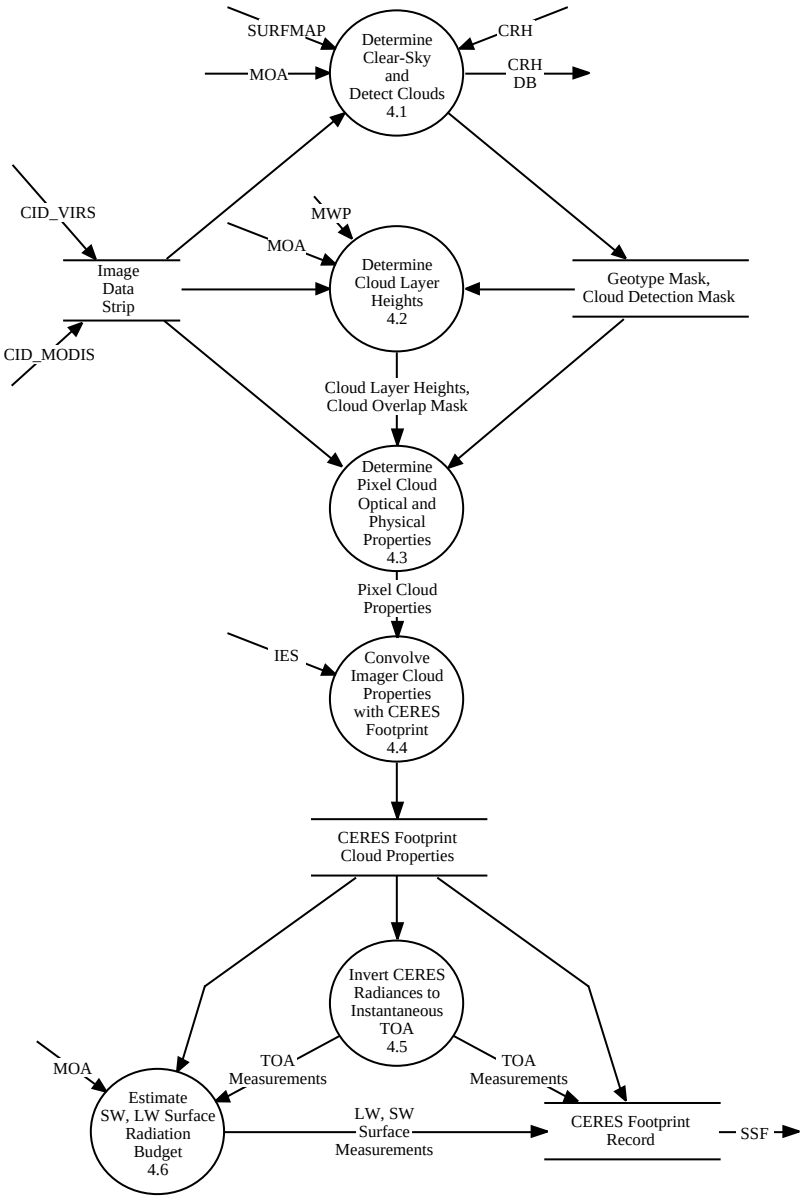
AU	astronomical unit
cm	centimeter
cm-sec^{-1}	centimeter per second
count	count
day	day, Julian date
deg	degree
deg-sec^{-1}	degree per second
DU	Dobson unit
erg-sec^{-1}	erg per second
fraction	fraction (range of 0–1)
g	gram
g-cm^{-2}	gram per square centimeter
g-g^{-1}	gram per gram
g-m^{-2}	gram per square meter

h	hour
hPa	hectopascal
K	Kelvin
kg	kilogram
kg-m^{-2}	kilogram per square meter
km	kilometer
km-sec^{-1}	kilometer per second
m	meter
mm	millimeter
µm	micrometer, micron
N/A	not applicable, none, unitless, dimensionless
ohm-cm^{-1}	ohm per centimeter
percent	percent (range of 0–100)
rad	radian
rad-sec^{-1}	radian per second
sec	second
sr^{-1}	per steradian
W	watt
W-m^{-2}	watt per square meter
W-$m^{-2}sr^{-1}$	watt per square meter per steradian
W-$m^{-2}sr^{-1}µm^{-1}$	watt per square meter per steradian per micrometer

CERES Top Level Data Flow Diagram



Subsystem 4.0 Top Level Data Flow Diagram



Clouds and the Earth's Radiant Energy System (CERES)

Algorithm Theoretical Basis Document

Imager Clear-Sky Determination and Cloud Detection

(Subsystem 4.1)

CERES Science Team Cloud Retrieval Working Group

Bryan A. Baum[1]
Ronald M. Welch[2]
Pat Minnis[1]
Larry L. Stowe[3]
James A. Coakley, Jr.[4]

Algorithm Implementation, Data Analysis, and Data Management

James Titlow[5]
Vasanth Tovinkere[5]
Pat Heck[6]
Qing Trepte[5]
Dave Doelling[6]
Shalini Mayor[6]

Automated Classification Techniques and Algorithm Development

Todd Berendes[2]
Qingyuan Han[1]
Sundar A. Christopher[1]
Kwo-Sen Kuo[1]
Manuel Penaloza[1]
Antonette Logar[1]
Paul Davis[3]

[1]Atmospheric Sciences Division, NASA Langley Research Center, Hampton, Virginia 23681-0001
[2]South Dakota School of Mines and Technology, Rapid City, South Dakota 57701-3995
[3]Satellite Research Laboratory, National Oceanic and Atmospheric Administration, 5200 Auth Road, Camp Springs, Maryland 20746
[4]Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331-2209
[5]Science Applications International Corporation (SAIC), Hampton, Virginia 23666
[6]Analytical Services & Materials, Inc., Hampton, Virginia 23666

4.1. Imager Clear-Sky Determination and Cloud Detection

4.1.1. Overview

This document outlines the methodology for the CERES Release 1 global cloud mask. The cloud mask will be applied to the appropriate imager data stream for TRMM or EOS (VIRS or MODIS, respectively). More precisely, the goal of this effort is to determine those imager pixels that are unobstructed between the top of atmosphere (TOA) and the surface. The output from this algorithm is a pixel-level mask that includes information about which cloud masking tests were applied, whether each test did or did not indicate cloud, and a final decision of whether cloud was or was not present for each pixel.

The cloud mask is being designed for the narrowband channels on both the AVHRR and VIRS instruments. The additional capabilities afforded by the MODIS instrument will be addressed in Release 2 of this document. The members of the CERES cloud mask development team are closely coordinating their activities with the MODIS cloud mask algorithm development. Therefore, close coordination between the CERES and MODIS efforts will be maintained.

The CERES cloud masking algorithm will rely heavily upon a rich heritage of both NASA and NOAA experience with global data analysis. Initial algorithm design will incorporate the approaches used by ISCCP (International Satellite Cloud Climatology Project) (Rossow and Garder 1993), CLAVR (Clouds from AVHRR) (Stowe et al. 1991), and SERCAA (Support of Environmental Requirements for Cloud Analysis and Archive). The ISCCP algorithms are based upon two channels, one in the visible wavelength region and one in the infrared. The CLAVR approach uses all five channels of the AVHRR instrument. The CLAVR multispectral threshold approach with narrowband channel difference and ratio tests will be used, including dynamic threshold specification with clear-sky radiation statistics. The SERCAA algorithm is operational at the Phillips Laboratory, Hanscom Air Force Base, and uses all five AVHRR radiometric channels. The SERCAA is sponsored jointly by the Department of Defense, Department of Energy, and Environmental Protection Agency Strategic Environmental Research and Development Program. When appropriate, the spatial coherence method (Coakley and Bretherton 1982) will be used to improve the clear-sky spectral values. Artificial intelligence classification approaches will be applied for complex scene analysis, especially in polar, rugged terrain, and coastal regions.

The cloud mask algorithm will be tested on two months of global AVHRR GAC data. It will be modified as needed for Release 2 and then delivered for use with VIRS data on the TRMM mission.

4.1.2. Data and Assumptions

4.1.2.1. Assumptions

Anyone who has worked with data measured in the field quickly comes to realize that the real world is less than perfect. A number of assumptions may be listed that attempt to place boundaries on the cloud mask task.

1. Satellite data used as input to the cloud mask algorithm is calibrated.
2. Satellite level 1-B data, for some imaging instruments, may be striped (like the GOES scanner) or have some “smearing” at high viewing scan angles. We assume that the data contains no striping or smearing.
3. The mask will be provided for “good” data only, i.e., for those narrowband channels that have radiometric integrity. For instance, the AVHRR 3.7-μm channel is sometimes too noisy to permit accurate analysis of the radiometric data. This assumption implies that there may be holes in the mask if the data are incomplete.
4. The system level integration issues associated with implementation of this algorithm will not be raised in this subsystem document.

5. Sea surface temperature, surface snow/ice coverage, and operational National Meteorological Center gridded analysis products are assumed to be available for the operational cloud mask algorithm.
6. Smoke from forest fires, dust storms over deserts, and other surface phenomena that result in obstructing the field of view between the surface and the satellite will be considered as "cloud" if such pixels pass the threshold tests. When new tests are developed that distinguish between these phenomena, they will be incorporated into the algorithm.

4.1.2.1.1. Input data. The primary input data sets for subsystem 4.1 are the AVHRR GAC (global area coverage) satellite data and the following ancillary data sets:

- 1-min resolution coastline map, with lakes, rivers, islands, state/country boundaries
- 10-min resolution topographical map (see section 4.1.5.1. for further information)
- 10-min resolution ecosystem map
- 18-km resolution U.S. Navy/NOAA weekly Sea Ice Product
- 150-km or better resolution weekly NOAA Snow Data Product
- NMC gridded meteorological analysis product
- NOAA gridded weekly sea surface temperature product

The spatial resolution of the AVHRR GAC data is about 4 km at nadir. The spectral data include AVHRR channels 1 (0.55–0.68 μm), 2 (0.725–1.1 μm), 3 (3.55–3.93 μm), 4 (10.5–11.5 μm), and 5 (11.5–12.5 μm), which include visible, near-infrared, and infrared window regions. The NOAA-11 central wave numbers for the AVHRR IR channels are (see Kidwell 1991)

Table 4.1-1. Central Wave Numbers as a Function of Temperature for the NOAA-11 AVHRR NIR and IR Channels

Temperature Range (K)	Ch 3 (cm^{-1})	Ch 4 (cm^{-1})	Ch 5 (cm^{-1})
180-225	2663.50	926.80	837.75
225-275	2668.15	927.34	838.08
275-320	2671.40	927.80	838.40
270-310	2670.95	927.73	838.35

The values shown in Table 4.1-1 are slightly different for other sensors in this series of instruments. The VIRS instrument has a 720-km swath width with spectral measurements at channels 1 (0.63 ±0.05 μm), 2 (1.60 ±0.03 μm), 3 (3.75 ±0.19 μm), 4 (10.80 ±0.5 μm), and 5 (12.00 ±0.5 μm).

4.1.2.1.2. Output data. The output from the cloud mask algorithm will be a pixel by pixel product; i.e., a cloud mask will be derived for each imager pixel. The mask will be derived for the highest spatial resolution data available. There will be a final decision as to whether the pixel was obstructed or not that will be based upon the various cloud mask tests applied during the course of the algorithm. The final obstruction/no obstruction decision is stored in the variable denoted by "cloud fraction" in the imager pixel output data structure (Table 4.4-4). For the VIRS instrument, the cloud fraction will be either a "0" or a "1." For validation purposes only, a separate output data structure will be implemented that stores the results from the individual tests. If there are 10 tests applied to identify cloud, there will be 10 results saved for each pixel.

4.1.3. Cloud Masking Algorithms

4.1.3.1. Overview

Clouds generally are characterized by higher albedos and lower temperatures than the underlying surface. However, there are numerous conditions when this characterization is inappropriate, most notably over snow and ice. Of the cloud types, cirrus, low stratus, and small cumulus are the most difficult to detect. Likewise, cloud edges are difficult to recognize when they do not completely fill the field of view (FOV) of the imager pixel. The cloud mask effort builds upon operational experience of several groups that will now be discussed.

The NOAA CLAVR algorithm (Phase I) uses all five channels of AVHRR to derive a global cloud mask (Stowe et al. 1991). It examines multispectral information, channel differences, and spatial differences and then employs a series of sequential decision tree tests. Cloud-free, mixed (variable cloudy), and cloudy regions are identified for 2×2 GAC pixel arrays. If all four pixels in the array fail all the cloud tests, then the array is labeled as cloud-free (0% cloudy); if all four pixels satisfy just one of the cloud tests, then the array is labeled as 100% cloudy. If one to three pixels satisfy a cloud test, then the array is labeled as mixed and assigned an arbitrary value of 50% cloudy. If all four pixels of a mixed or cloudy array satisfy a clear-restoral test (required for snow/ice, ocean specular reflection, and bright desert surfaces) then the pixel array is reclassified as "restored-clear" (0% cloudy). The set of cloud tests is subdivided into daytime ocean scenes, daytime land scenes, nighttime ocean scenes, and nighttime land scenes.

Subsequent phases of CLAVR, now under development, will use dynamic clear/cloud thresholds predicted from the angular pattern observed from the clear sky radiance statistics of the previous 9-day repeat cycle of the NOAA satellite for a mapped 1° equal area grid cell (Stowe et al. 1994). As a further modification, CLAVR will include pixel by pixel classification based upon different threshold tests to separate clear from cloud contaminated pixels, and to separate cloud contaminated pixels into partial and total (overcast) cover. Cloud contaminated pixels will be radiatively "typed" as belonging to low stratus, thin cirrus, and deep convective cloud systems. A fourth type is middle mixed which includes all other cloud types.

The International Satellite Cloud Climatology Project (ISCCP) cloud masking algorithm is described by Rossow (1989), Rossow and Gardner (1993), and Seze and Rossow (1991a, b). Only two channels are used, the narrowband VIS (0.6 μm) and the IR (11 μm). Each observed radiance value is compared against its corresponding Clear-Sky Composite value. This step uses VIS radiances, not VIS reflectances. Clouds are assumed to be detected only when they alter the radiances by more than the uncertainty in the clear values. In this way the "threshold" for cloud detection is the magnitude of the uncertainty in the clear radiance estimates. As such this algorithm is not a constant threshold method such as used in Phase I of the CLAVR algorithm.

The ISCCP algorithm is based on the premise that the observed VIS and IR radiances are caused by only two types of conditions, "cloudy" and "clear," and that the ranges of radiances and their variability that are associated with these two conditions do not overlap (Rossow and Garder 1993). As a result, the algorithm is based upon thresholds, where a pixel is classified as "cloudy" only if at least one radiance value is distinct from the inferred "clear" value by an amount larger than the uncertainty in that "clear" value. The uncertainty can be caused both by measurement errors and by natural variability. This algorithm is constructed to be "cloud-conservative," minimizing false cloud detections but missing clouds that resemble clear conditions.

The ISCCP cloud-detection algorithm consists of five steps (Rossow and Garder 1993):

1. Space contrast test on a single IR image
2. Time contrast test on three consecutive IR images at constant diurnal phase

3. Cumulation of space/time statistics for IR and VIS images
4. Construction of clear-sky composites for IR and VIS every 5 days at each diurnal phase and location
5. Radiance threshold for IR and VIS for each pixel

Parts of the ISCCP scheme will be incorporated into the CERES cloud mask. Some modifications are necessary since all the AVHRR channels will be used, not just the visible and infrared channels (AVHRR channels 1 and 4).

The Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA) algorithm from the Air Force uses multispectral AVHRR data to derive a global cloud mask. The SERCAA cloud decision tree consists of a series of cloud tests and background filter tests to identify cloudy and clear scenes using multispectral data and empirical thresholds. The algorithm is performed on a pixel-by-pixel basis. Percent albedo of channel 1 and channel 2 used in SERCAA has been changed to reflectance for CERES analysis.

The spatial coherence method (Coakley and Bretherton 1982) is especially useful in determining clear and cloudy sky radiances. It is applicable to single-layered and sometimes multilayered cloud systems that extend over moderately large regions, greater than 250 km × 250 km, and which have completely cloudy and completely clear pixels. Using the local spatial structure of the IR radiances, difficulties arise when interpreting situations involving multilayered cloud systems, subpixel-sized clouds, and clouds with variable emissivities.

4.1.3.2. ISCCP Space Contrast Test

This test, described in Rossow and Garder (1993), is similar to that of spatial coherence in that it is applied only to IR brightness temperatures. It is based upon the fact that clear pixels tend to have higher temperatures than cloudy pixels and to exhibit less spatial variability. First a small local region is defined, composed of pixels with the same ecosystem. The spatial domain is approximately 450 km × 450 km over ocean, 90 km × 90 km over ice-covered water, and 90 km × 90 km over land. The pixel in the local region with the largest IR (11-μm) brightness temperature ($T_{B\max}$) is found, consistent with the spatial coherence test. All pixels with temperatures lower than the spatial contrast threshold defined by

$$T_B < T_{B\max} - \text{Threshold}_{cs} \tag{4.1-1}$$

are labeled as cloudy; all others are labeled as “undecided.” Since cloud variability can be as small as surface variability, values of Threshold_{cs} = 3.5 K are chosen over ocean and Threshold_{cs} = 6.5 K over both ice-covered water and land. Note that not only is it important that the class of pixels be identical (land or ocean), but also that the size of the region be chosen carefully. All coastal regions and all land regions containing mixed land and water pixels are excluded from this test, since the inherent contrast between land and water surface radiances would dominate the results. For regions that are too large, there is increased likelihood of spatial variations in surface parameters. The shape of the test regions also can be important, since meridianal gradients in surface temperature generally are larger than zonal gradients. The size of the contrast threshold must be larger than the magnitude of natural variations at the surface and smaller than that caused by clouds.

4.1.3.3. ISCCP Spatial/Temporal Analysis

In decreasing order of magnitude, temporal variations of IR and VIS radiances are caused by: (1) formation/dissipation of clouds, or advection of clouds, (2) diurnal heating and cooling (IR) and changes in solar illumination (VIS), (3) variations of surface parameters at synoptic and seasonal time scales, and (4) atmospheric conditions. Investigations of temporal variability and regional variations are reported by Minnis and Harrison (1984a, b), Seze and Desbois (1987), Gutman et al. (1987) and Seze

and Rossow (1991a, b). Diurnal variations of surface temperature on land can be large enough to prevent a sensitive test. However, this can be partially overcome by modeling the full diurnal cycle based on the clear measurements in daytime (Minnis and Harrison, 1984a). The IR radiances are corrected to an approximate nadir view by using a radiative transfer model based upon zonal, monthly mean atmospheric conditions derived from TOVS data.

The following has been adapted from Rossow and Garder (1993); in this implementation, we have included tests for AVHRR channel 2 and 3 reflectances. The first test in this set examines the temperature and channels 1 through 3 reflectance values on the present and previous clear days. If

$$\left|T_B(i) - T_B(cs)\right| < \text{Threshold}_T^{\min} = 2.5 \text{ K} \tag{4.1-2}$$

$$\left|\rho_1(i) - \rho_1(cs)\right| < \text{Threshold}_1^{\min} \tag{4.1-3}$$

$$\left|\rho_2(i) - \rho_2(cs)\right| < \text{Threshold}_2^{\min} \tag{4.1-4}$$

$$\left|\rho_3(i) - \rho_3(cs)\right| < \text{Threshold}_3^{\min} \tag{4.1-5}$$

then the pixel is labeled as probably clear, where $T(i)$ and $T(cs)$ are the measured (i) and clear sky (cs) temperatures, respectively, and $\rho(i)$ and $\rho(cs)$ are measured and clear sky reflectance values for each of the three channels. Obviously, $\text{Threshold}_T^{\min}$, $\text{Threshold}_1^{\min}$, $\text{Threshold}_2^{\min}$, and $\text{Threshold}_3^{\min}$ must be larger than the natural surface variability. On the other hand, if these values are set too low, then relatively smooth low-level broken cloudiness will be classified as clear.

Likewise, if

$$\left|T_B(i) - T_B(cs)\right| > \text{Threshold}_T^{\max} = 6 \text{ K} \tag{4.1-6}$$

$$\left|\rho_1(i) - \rho_1(cs)\right| > \text{Threshold}_1^{\max} \tag{4.1-7}$$

$$\left|\rho_2(i) - \rho_2(cs)\right| > \Delta\text{Threshold}_2^{\max} \tag{4.1-8}$$

$$\left|\rho_3(i) - \rho_3(cs)\right| > \text{Threshold}_3^{\max} \tag{4.1-9}$$

then the pixel is labeled as probably cloudy. Pixels which do not satisfy either of the above two tests are labeled as undecided. The reflectance values defined above are different for different ecosystems.

Events associated with storms may cause large changes in surface temperature. Snow and precipitation and wetting the ground cause large changes in surface reflectances. Therefore, if the previous day was cloudy and all of the neighboring pixels of the same ecosystem were cloudy, then it is assumed that a major storm may have occurred and that the above tests are suspect. If some of the neighboring pixels of the same ecosystem were clear on the previous day, then no major storm event took place. In this case, the average values of the neighboring clear pixels for the previous day are used in these tests.

A final spatial/temporal test examines characteristic variations of clear conditions of the same ecosystem type over larger spatial scales and at longer periods of time. The short-term (ST) period of time is approximately 9 days; the long-term (LT) period of time is approximately 25 days. The actual ST and LT time scales vary according to ecosystem (see Table 4.1-11). The short-term period of time approximates the natural time scale for significant variability of the local surface temperature and reflectances and is the repeat cycle for the AVHRR sensor. The long-term period of time is consistent with variations of more persistent cloud cover and covers three AVHRR cycles. Statistics of the mean (M) and standard deviation (σ) are computed for both ST and LT over approximately 32×32 pixel regions of

the same ecosystem. The minimum and maximum values of ST and LT over this region also are found. If the present-day values are labeled as clear and if these values lie within one standard deviation of the ST and LT values, then the pixel is labeled as definitely clear. If the present day mean value lies between $M + \sigma$ and the maximum value or between $M - \sigma$ and the minimum value, then the pixel is labeled as probably clear. A similar test is made if the present day pixel is labeled as cloudy. Otherwise the pixel is labeled as undecided.

4.1.3.4. CLAVR Reflectance Uniformity Test

The reflectance uniformity test is applied by computing the maximum and minimum values of AVHRR channel 1 or channel 2 reflectances within a 2×2 pixel array. Pixel arrays with channel 1 reflectance differences greater than 9% over land or channel 2 reflectance differences greater than 0.3% over ocean are labeled as mixed (Davis et al. 1993). The value over ocean is low because a cloud-free ocean is almost uniformly reflective, while nonuniformity is assumed to be caused by cloudiness.

Note that this test is being refined; first, by requiring that the ecosystem be the same for the pixel array. Second, the mean and standard deviation of reflectance values for each of the 59 ecosystems (see section 4.1.4.1.) will be computed for channels 1 through 3 as a function of season. It is expected that this test can be substantially improved.

4.1.3.5. Cirrus Cloud Tests

a. SERCAA. The brightness temperature difference between channel 4 and channel 5 ($T_{B4} - T_{B5}$, or BTD^{45}) exhibits a persistent cirrus cloud signature based on the fact that cirrus cloud brightness temperatures are consistently higher at 10.7 μm than at 11.8 μm. However, in the absence of cloud, water vapor attenuation can cause a positive BTD^{45} that could be mistaken for a cloud signature. Thus, the cloud detection threshold is defined as a function of the channel 4 brightness temperature T_{B4} (as a surrogate for water vapor loading) and viewing zenith angle θ (to count for atmospheric path length). Table 4.1-2 contains the threshold values for a range of T_{B4} and θ developed by Saunders and Kriebel (1988) is used as the basis in the Cirrus Cloud Test.

Table 4.1-2. Thresholds for SERCAA Cirrus Cloud Test

	Threshold for sec(θ) of—				
T_{B4}	1.00	1.25	1.50	1.75	2.00
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

The cirrus cloud test is defined as

$$T_{B4} - T_{B5} > \text{Threshold}(T_{B4}, \theta) \tag{4.1-10}$$

It can apply to both daytime and nighttime.

When the background is classified as snow or ice covered, an additional test is required based on the assumption that channel 4 brightness temperatures measured from cirrus clouds are lower than the terrestrial background temperature. This test is defined as:

$$T_{cs} - T_{B4} > \text{Threshold}_{ci} \tag{4.1-11}$$

where T_{cs} is the clear sky brightness temperature, and Threshold_{ci} = 5.0 K, is the cirrus cloud snow/ice filter threshold.

b. CLAVR. The CLAVR brightness temperature difference between channels 4 and 5 (BTD^{45}) (Stowe et al. 1993; Inoue 1987, 1989) is particularly effective in detecting cirrus clouds. Stowe et al. suggest the following thresholds for oceans and land:

$$\text{Threshold}_{45}(\text{oceans}) = \sum_{i=0}^{5} a_i T_{B4}^{i} \tag{4.1-12}$$

$$\text{Threshold}_{45}(\text{land}) = \sum_{i=0}^{4} a_i T_{B4}^{i} \tag{4.1-13}$$

where the coefficients a_i are given in Table 4.1-3. If the value for Threshold_{45} is greater than the threshold determined from equations (4.1-12) or (4.1-13), the pixel is labelled as being cloudy. If T_{B4} is less than 260 K over a water surface, the threshold is set to zero.

Table 4.1-3. Coefficients Used to Determine Thresholds for CLAVR Cirrus Test

Coefficient	Ocean	Land
a_0	9.27066×10^{4}	-1.34436×10^{4}
a_1	-1.79203×10^{3}	194.945
a_2	13.8305	–1.05635
a_3	–0.0532679	2.53361×10^{-3}
a_4	1.02374×10^{-4}	-2.26786×10^{-6}
a_5	-7.85333×10^{-8}	0

4.1.3.6. Cold Cloud Test

The Cold Cloud Test uses a single IR channel to discriminate the thermal signature of midlevel clouds from the terrestrial background. A cloud decision is made by comparing the channel 4 brightness temperature T_{B4}, with the clear scene brightness temperature T_{cs}. When T_{B4} is lower than T_{cs} by a amount greater than a preset threshold, the pixel is classified as cloudy. The test is defined as:

$$T_{cs} - T_{B4} > \text{Threshold}_{\text{cold}} \tag{4.1-14}$$

where $\text{Threshold}_{\text{cold}}$ is the surface-dependent threshold shown in the following table:

Table 4.1-4. Thresholds for Cold Cloud Test

Surface background	Threshold (K)
Water	9.0
Land	10.0
Coast	20.0
Desert	10.0
Snow	15.0

4.1.3.7. Daytime Low Cloud and Fog Test

The Low Cloud and Fog Test is based on the different radiative characteristics of liquid water clouds at AVHRR channel 3 (3.7 μm) and channel 4 (10.8 μm) wavelengths. During daytime, the radiance from channel 3 is a combination of both emitted and reflected energy, while channel 4 is only emitted energy. The test assumes that a liquid water cloud will reflect enough solar energy at 3.7 μm to make the channel 3 brightness temperature, T_{B3}, significantly higher that T_{B4}. The test is defined as the difference between the 3.7- and 10.8-μm brightness temperatures (BTD_{34}):

$$T_{B3} - T_{B4} > \text{Threshold}_{lcf} \tag{4.1-15}$$

where Threshold_{lcf} is a surface-dependent cloud detection threshold given in Table 4.1-5.

Table 4.1-5. Thresholds for Daytime Low Cloud and Fog Test

Surface background	Threshold (K)
Nondesert	12.0
Desert	20.0
Sun glint regions	54.0

The test is extremely sensitive to desert surface and Sun glint, since they increase the 3.7-μm radiance relative to the 10.8-μm radiance. Potential sun glint areas are identified prior to testing for cloud contamination and a larger threshold is applied at sun glint regions.

4.1.3.8. Daytime Precipitating Cloud Test

The Precipitating Cloud Test exploits the reflective nature of thick ice clouds at 3.7 μm. Optically thick ice clouds, such as towering cumulonimbus, reflect more strongly than optically thin cirrus. Therefore, the brightness temperature from channel 3, T_{B3}, is much higher than the true physical temperature of clouds, represented by T_{B4}. The test is defined as

$$T_{B3} - T_{B4} > \text{Threshold}_{\text{precip(1)}} \tag{4.1-16}$$

where $\text{Threshold}_{\text{precip(1)}} = 20.0$ K is a cloud detection threshold.

Two additional checks should also be performed to discriminate cumulonimbus clouds from low liquid water clouds and optical thin ice clouds, such as cirrostratus.

$$T_{\text{clear sky}} - T_{B4} > \text{Threshold}_{\text{precip(2)}} \tag{4.1-17}$$

$$\rho_2 > \text{Threshold}_{\text{precip(3)}} \tag{4.1-18}$$

where $T_{\text{clear sky}}$ is the clear sky brightness temperature, ρ_2 is reflectance of channel 2, and $\text{Threshold}_{\text{precip(2)}}$ and $\text{Threshold}_{\text{precip(3)}}$ are precipitating cloud detection thresholds. $\text{Threshold}_{\text{precip(2)}} = 30.0$ K and $\text{Threshold}_{\text{precip(3)}} = 0.45$.

The $T_{\text{clear sky}} - T_{B4}$ test eliminates any low clouds that pass the $T_{B3} - T_{B4}$ test by ensuring that the true physical cloud top temperature is significantly lower than the clear scene brightness temperature. The ρ_2 test eliminates ice clouds that are not optically thick, and hence not as bright as precipitating clouds.

4.1.3.9. Daytime Thin Cirrus Cloud Test

The Daytime SERCAAThin Cirrus Cloud Test utilizes the results from the solar independent Cirrus Cloud Test and the reflectance of channel 1 and channel 2. Recall the Cirrus Cloud Test requires the following conditions to be met:

$$T_{B4} - T_{B5} > \text{Threshold}(T_{B4}, \theta) \tag{4.1-19}$$

where Threshold (T_{B4}, θ) is the cloud detection threshold obtained through interpolation from Table 4.1-2.

If the background is classified as snow or ice covered, an additional test is required:

$$T_{\text{clear sky}} - T_{B4} > \text{Threshold}_{ci} \tag{4.1-20}$$

where $T_{\text{clear sky}}$ is the clear sky brightness temperature, and Threshold_{ci} = 5.0 K is the cirrus cloud snow/ice filter threshold.

In addition to the tests listed above, the Daytime Thin Cirrus Cloud Test uses reflectance of channel 1 (ρ_1) and channel 2 (ρ_2) to discriminate thin cirrus. The criterion used is dependent on the background surface type:

$$\rho_2 < \text{Threshold}_{dci\_w} \qquad \text{(Over water)} \tag{4.1-21}$$

$$\rho_1 < \text{Threshold}_{dci\_l} \qquad \text{(Over land)} \tag{4.1-22}$$

where Threshold_{dci_w} and Threshold_{dci_l} are the cloud detection threshold values over water and land, respectively, Threshold_{dci_w} = 0.2 and Threshold_{dci_w} = 0.2.

4.1.3.10. Visible Reflectance Ratio Test

The Visible Reflectance Ratio Test is based on the fact that for clouds, the spectral signature in channel 1 and channel 2 are very close to each other so that the ratio ρ_2/ρ_1 is approximately equal to 1. For clear land surfaces, the radio is greater than 1.0 and for water surfaces, the ratio is less than 1.0. Thus, the cloud test is applies by testing the ρ_2/ρ_1 ratio against upper and lower limit cloud thresholds.

The test is only used in the absence of sun glint, desert, snow/ice background, and coast regions, all of which can produce a false cloud signal.

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, which results in a higher measured channel 1 reflectance to channel 2 for cloud-free areas and produces a false cloud signature. To account for this, the value for 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 clear sky brightness temperature against a threshold:

$$T_{\text{clear sky}} > \text{Threshold}_{\text{ratio\_humid}} \tag{4.1-23}$$

where $\text{Threshold}_{\text{ratio\_humid}}$ = 295 K is the high humidity threshold. In regions where this test evaluates as true, the Visible Brightness Ratio Test is defined as

$$\text{Threshold}_{\text{ratio\_lo\_wet}} < \rho_2 / \rho_1 < \text{Threshold}_{\text{ratio\_up\_wet}} \tag{4.1-24}$$

where $\text{Threshold}_{\text{ratio\_lo\_wet}}$ and $\text{Threshold}_{\text{ratio\_up\_wet}}$ are the lower and upper limit ratio thresholds for high humidity. In regions where the humidity test evaluates as false, the Visible Brightness Ratio Test uses a different set of thresholds:

$$\text{Threshold}_{\text{ratio\_lo\_dry}} < \rho_2 / \rho_1 < \text{Threshold}_{\text{ratio\_up\_dry}} \quad (4.1\text{-}25)$$

where $\text{Threshold}_{\text{ratio\_lo\_dry}}$ and $\text{Threshold}_{\text{ratio\_up\_dry}}$ are the lower and upper limit ratio thresholds for lower humidity given in Table 4.1-6.

Table 4.1-6. Thresholds Based on Humidity for Visible Reflectance Ratio Test

$\text{Threshold}_{\text{ratio\_lo\_wet}}$	0.7
$\text{Threshold}_{\text{ratio\_up\_wet}}$	1.0
$\text{Threshold}_{\text{ratio\_lo\_dry}}$	0.75
$\text{Threshold}_{\text{ratio\_up\_dry}}$	1.1

4.1.3.11. Reflectance Threshold Test

The test described here is used in CLAVR, SERCAA, and ISCCP, and uses a visible wavelength channel threshold to discriminate high cloud reflectance from a low background reflectance. This test works well in discriminating most cloud types with the exception of thin cirrus. The clear sky background reflectance (ρ_{cs}) is calculated from clear sky albedo (α_{cs}) and the bidirectional reflectance function (BDRF). The clear-sky albedo is obtained by spatial and temporal interpolation from ISCCP's 3-hour 2.5° map. The BDRF's for ocean and land were developed from GOES East and GOES West data (Minnis and Harrison 1984a, b, c); BDRF's for other surface types are taken from the ERBE broadband bidirectional models until other models can be developed and tested.

The clear sky reflectance is shown as follows:

$$\rho_{cs} = \alpha_{cs} / \text{BDRF}(\theta_o, \theta, \phi, M) \quad (4.1\text{-}26)$$

where θ_o, θ, and ϕ are solar zenith, viewing zenith, and relative azimuth angles, and M is scene type.

A pixel is classified as cloudy if the satellite measured reflectance exceeds the expected clear-scene background value by an amount greater than a threshold. The test is only applied for the pixels with $\theta_o < 70°$ and not applied for regions containing sun glint, desert, or snow/ice background. Separate thresholds and different channels are used for land and water backgrounds. Over land, channel 1 reflectance is used, while over water channel 2 data is used. The test is defined as

$$\rho_1 - \rho_{cs} > \text{Threshold}_{\text{land}} \quad \text{(Over land)} \quad (4.1\text{-}27)$$

$$\rho_2 > \text{Threshold}_{\text{water}} \quad \text{(Over water)} \quad (4.1\text{-}28)$$

where $\text{Threshold}_{\text{land}} = 0.25$ and $\text{Threshold}_{\text{water}} = 0.16$ are cloud detection thresholds over land and water background, respectively.

4.1.3.12. Channel 3 Reflectance Test

Likewise channel 3 reflectance values > 6% are considered to be cloudy. However, "cloudy" pixels with channel 3 reflectance values < 3% are considered to be snow/ice (Davis et al. 1993). Note that the channel 3 reflectance tests are not applied over deserts. This is because bright desert regions with highly variable emissivities tend to be misclassified as cloudy with this test. Thermal contrast needs to be examined in conjunction with channel 3 reflectivity. As we gain experience with these approaches, the actual thresholds will be adjusted to ecosystem type.

4.1.3.13. Nighttime Low Stratus Test

Both SERCAA and CLAVR describe low stratus tests (LST) based upon the brightness temperature differences between the 3.7- and 11-μm channels. The test assumes that for water droplet clouds, the emissivity at 3.7 μm (channel 3) is general lower than at 10.8 μm (channel 4). For the CLAVR test, the threshold for the LST test (Threshold_{LST}) is described as:

$$\text{Threshold}_{LST} = \exp\{A + BT_{B4}\} - C \tag{4.1-29}$$

where $A = -9.37528$, $B = 0.0341962$, and $C = 1.0$ (oceans) and $C = 3.0$ (land). The constant C increases for land from the ocean value and depends on surface type. This test is applicable for the temperature range 264 K to clear-sky T_{B4}. If the threshold is exceeded, then low stratus is said to exist. The specific values of the coefficients may vary in the CERES implementation, depending on the results of testing with global GAC data.

The SERCAA test assumes that clouds are detected if T_{B4} is greater than T_{B3} by an amount greater than a cloud detection threshold:

$$T_{B4} - T_{B3} > \text{Threshold}_{LST} \tag{4.1-30}$$

where Threshold_{LST} is a surface-dependent cloud detection threshold:

$$\text{Threshold}_{LST} = 1.0 \text{ K} \qquad \text{(Over nondesert)}$$

$$\text{Threshold}_{LST} = 2.0 \text{ K} \qquad \text{(Over desert)}$$

The final determination of thresholds to use for the CERES algorithm will be determined through global analysis of AVHRR data.

4.1.4.14. Nighttime Thin Cirrus Test

Both the SERCAA and CLAVR methods use a similar test based upon the difference in brightness temperatures between the 3.7- and 12-micron channels ($T_{B3} - T_{B5}$, or BTD_{35}). The test is based on the idea that cirrus cloud transmissivity at 3.7 μm (channel 3) is generally greater than at 12 μm (channel 5), causing some radiation from warmer backgrounds to be included in the channel 3 measurement. If the difference exceeds a given threshold, then cirrus is said to exist in the pixel.

The CLAVR Cirrus Test (CIRT) is applied at night over both land and ocean. The threshold is determined by the brightness temperature of channel 4 (11 micron). This threshold was defined by using a simulation database to plot cloud-free CIRT values against the associated channel 4 temperatures. The relatively high optical transmittance of most cirrus clouds, along with the spectrally different Planck blackbody radiance dependence on temperature, can identify cirrus clouds. The CIRT threshold is given by

$$\text{Threshold}_{\text{CIRT}} = -0.485 + 1.775\times10^{3} T_{B4} \tag{4.1-31}$$

When $T_{B4} < 273$ K, this threshold is set to zero; when $T_{B4} > 292$ K, it is set to 0.033. If the threshold is exceeded, then thin cirrus is said to exist in the pixel.

The SERCAA Nighttime Thin Cirrus Cloud Test is defined as:

$$T_{B3} - T_{B5} > \text{Threshold}_{tci} \tag{4.1-32}$$

where $\text{Threshold}_{tci} = 4.0$ K is the nighttime thin cirrus cloud detection threshold.

Empirical study has found that in regions of high humidity, the water vapor can attenuate the channel 5 signal by several degrees K. As a result, clear background surfaces will appear significantly cooler in channel 5, and if the clear sky brightness temperature does not take the humidity into account, the

result could be a false detection of cloud. The high humidity regions are identified if the clear sky brightness temperature is greater than a defined threshold:

$$T_{cs} > \text{Threshold}_{tci\_\text{humid}} \tag{4.1-33}$$

where $\text{Threshold}_{tci\_\text{humid}}$ = 290 K is the high humidity threshold. If the humidity test is true, then AVHRR channel 4, which is less sensitive to water vapor attenuation, is used instead of channel 5 in the test:

$$T_{B3} - T_{B4} > \text{Threshold}_{tci} \tag{4.1-34}$$

where Threshold_{tci} is the same threshold used with channel 5.

4.1.4. Artificial Intelligence Cloud Classification Techniques

There are regions in which simple cloud mask algorithms have been shown to perform inadequately, such as in polar, rugged terrain, and coastal regions. For these areas, or when the cloud masks indicate no clear decision on whether cloud is present, artificial intelligence (AI) classification approaches will be applied. The AI classification approaches use a number of textural and spectral features, or measures, that are derived from the satellite data. The following discussion outlines the methods that will be employed in the Release 1 algorithm. We should note that if the reader wishes to skip this discussion, section 4.1.5. begins the actual description of the cloud mask implementation.

4.1.4.1. Texture Features

Texture is often interpreted in the literature as a set of statistical measures of the spatial distribution of gray levels in an image. Here it is assumed that textural information is contained in the average spatial relationships that gray levels have with one another (Haralick et al. 1973). The gray level difference-vector (GLDV) approach is based on the absolute differences between pairs of gray levels I and J found at a distance d apart at angle ϕ with a fixed direction. The GLDV difference-vector probability density function $P(m)_{d,\phi}$ is defined for $m = I - J$, where I and J are the corresponding gray levels having a value between 0 and 255. The gray level range may vary, but we will use 2^8 gray levels in our analysis. The density function $P(m)_{d,\phi}$ (henceforth $P(m)$, where the dependence of $P(m)$ on d and ϕ is implicitly assumed) is obtained by normalizing the GLDV difference vector by the total number of difference pairs. Once $P(m)$ has been formed, textural measures are computed for each of the five AVHRR spectral channels assuming a pixel separation distance of $d = 1$ and at an angle $\phi = 0°$ and 90°. The following textural features are computed for use in the classification system, and are calculated individually for each $N \times N$ pixel subarray.

Mean:

$$u = \sum_m mP(m) \tag{4.1-35}$$

Standard deviation:

$$\sigma = \left[\sum_m (m-u)^2 P(m)\right]^{1/2} \tag{4.1-36}$$

Contrast is a natural measure of the degree of spread in the gray levels. A small contrast value indicates high concentration of occurrences on the main diagonal and represents a coarse texture. Larger values of contrast indicate that the occurrences are spread out about the main diagonal and represent a finer structure:

$$CON = \sum_m m^2 P(m) \tag{4.1-37}$$

Angular second moment is a measure of homogeneity in the subscene. The measure is smallest when the gray levels are equally probable, and large values indicate that there are dominant gray levels present.

$$ASM = \sum_m [P(m)]^2 \tag{4.1-38}$$

Entropy is a measure of disorder in the scene, and is largest for equally distributed gray levels and smallest when they are unequally distributed.

$$ENT = \sum_m P(m) \log P(m) \tag{4.1-39}$$

Local homogeneity is a measure of local similarity and has a larger value for coarse textures than for finer textures.

$$HOM = \sum_m \frac{P(m)}{(1 + m^2)} \tag{4.1-40}$$

Cluster shade is a measure of the degree to which the outliers in the histogram favor one side or another of the mean.

$$SHADE = \frac{\left[\sum_m (m-u)^3 P(m)\right]}{\sigma^3} \tag{4.1-41}$$

Cluster prominence measures the effect of the outliers on the peak of the distribution.

$$PROM = \frac{\left[\sum_m (m-u)^4 P(m)\right]}{\sigma^4} \tag{4.1-42}$$

These features are described by Chen et al. (1989) in greater detail. Plots of representative cloud texture measures as a function of pixel separation distance and angle are shown in Welch et al. (1989) and for a variety of ice and snow backgrounds in Welch et al. (1990).

4.1.4.2. Spectral Features

The spectral features are formed from the gray level representation of the bidirectional reflectances for AVHRR channels 1 and 2 and from the gray level representation of brightness temperatures for the NIR and IR channels. The reflectances are calculated using the solar zenith angle θ_o at each pixel and then scaled to gray levels 0–255, representing 0%–100%, respectively. Gray level representation means that the range of possible values is scaled between 0–255. The daytime 3.7-μm measured radiance contains contributions from both solar reflection and thermal emission. For classification purposes only, the AVHRR 3.7-μm radiometric data (channel 3) are converted to bidirectional reflectance through a relationship commonly used for optically thick clouds (e.g., Allen et al. 1990; Ebert 1987; Key and Barry 1989):

$$\rho_3 = \frac{I_3 - B_3(T_{B4})}{\pi^{-1} F_3 \cos\theta_0 - B_3(T_{B4})} \tag{4.1-43}$$

where I_3 and F_3 are the radiance and incoming solar flux for channel 3, respectively. The 3.7-μm thermal emission is estimated by using the 10.8-μm brightness temperature (T_{B4}) to calculate the Planck emission at the central wavelength of channel 3, $B_3(T_{B4})$. The reflectance calculated in this fashion is used as a spectral feature because it has been shown to be effective in distinguishing between water and ice clouds (Allen et al. 1990). Note that this is a rough approximation to the true channel 3 reflectance and does not account for differences in emissivity between channels 3 and 4 or for nonblack clouds.

Additional spectral features are formed from combinations of the five AVHRR channels. Spectral features are formed from the gray-level representation of reflectances, brightness temperatures, or combinations of quantities, such as brightness temperature differences between two channels. The spectral features are calculated for a single channel quantity X (where a quantity is either a reflectance or a brightness temperature), for two different quantities X and Y, or for three quantities X, Y, and Z, as follows.

1. Mean X: This spectral feature is the mean gray level value of either reflectance or brightness temperature calculated from the array.
2. Band Difference $[X - Y]$: This spectral feature is the difference of the gray-level means of two channels.
3. Band Ratio $[X/Y]$: This feature is formed from taking the ratio of mean gray-level values between two channels.
4. Overlay $[X, Y, Z]$: This spectral feature forms a packed integer from the mean gray-level value of three quantities X, Y, and Z. It is similar in nature to the idea of using 24-bit color graphics to form false-color imagery. With the proper channel combination, warm reflective stratus clouds would have different values than cold, thin, less-reflective cirrus. The overlay of X, Y, and Z is calculated from OVERLAY = $Z * 2^{16} + Y * 2^8 + X$. This particular feature is useful in separating the clear-sky land, low-level cloud, mid-level cloud, and high-level cloud classes.
5. Low X: This feature is the percentage of pixels in the array that have a reflectance less than 10%. It is calculated only for AVHRR channels 1, 2, or 3, and is not calculated from gray-scale values. For channel 3, the reflectance is determined by Equation 4.1-43.
6. High X: This feature is essentially the complement of LOW X. It is the percentage of pixels where the reflectance is greater than 10% (again, only for AVHRR channels 1–3). For channel 3, the reflectance is determined by Equation 4.1-43.
7. Spatial coherence: For a given array, means and standard deviations are calculated for all local (2×2 or 4×4) pixel groups within the 32×32 array. The spatial coherence feature is the mean of those local pixel groups that have a standard deviation less than 2.5.

This list of spectral features demonstrates a sampling of the nature of features currently in use and will be added to in future work.

4.1.4.3. Subregion ($N \times N$ array) Labeling

A critical aspect of the algorithm development is subarray labeling. To train and test classifiers, a large number of labeled samples for each class are required. A sample is defined here as an $N \times N$ array of AVHRR data. Accurate labeling is the key to accurate classification. Therefore, it is important to provide the analyst with as much information as possible. The actual labeling process involves more than choosing samples directly from a screen image. For each scene, the analyst uses a variety of ancillary data sets to aid in gaining more information on the scene. For instance, for sample labeling over North America, we also study NMC analyses or rawinsonde temperature and humidity profiles and National Weather Service 6-hourly surface synoptic observations to gain a better understanding of the overall scene.

Figure 4.1-1 shows an example of the Satellite Image Visualization System (SIVIS) which displays three-band color overlays. A series of pull-down menus are available to the analyst which allow a wide

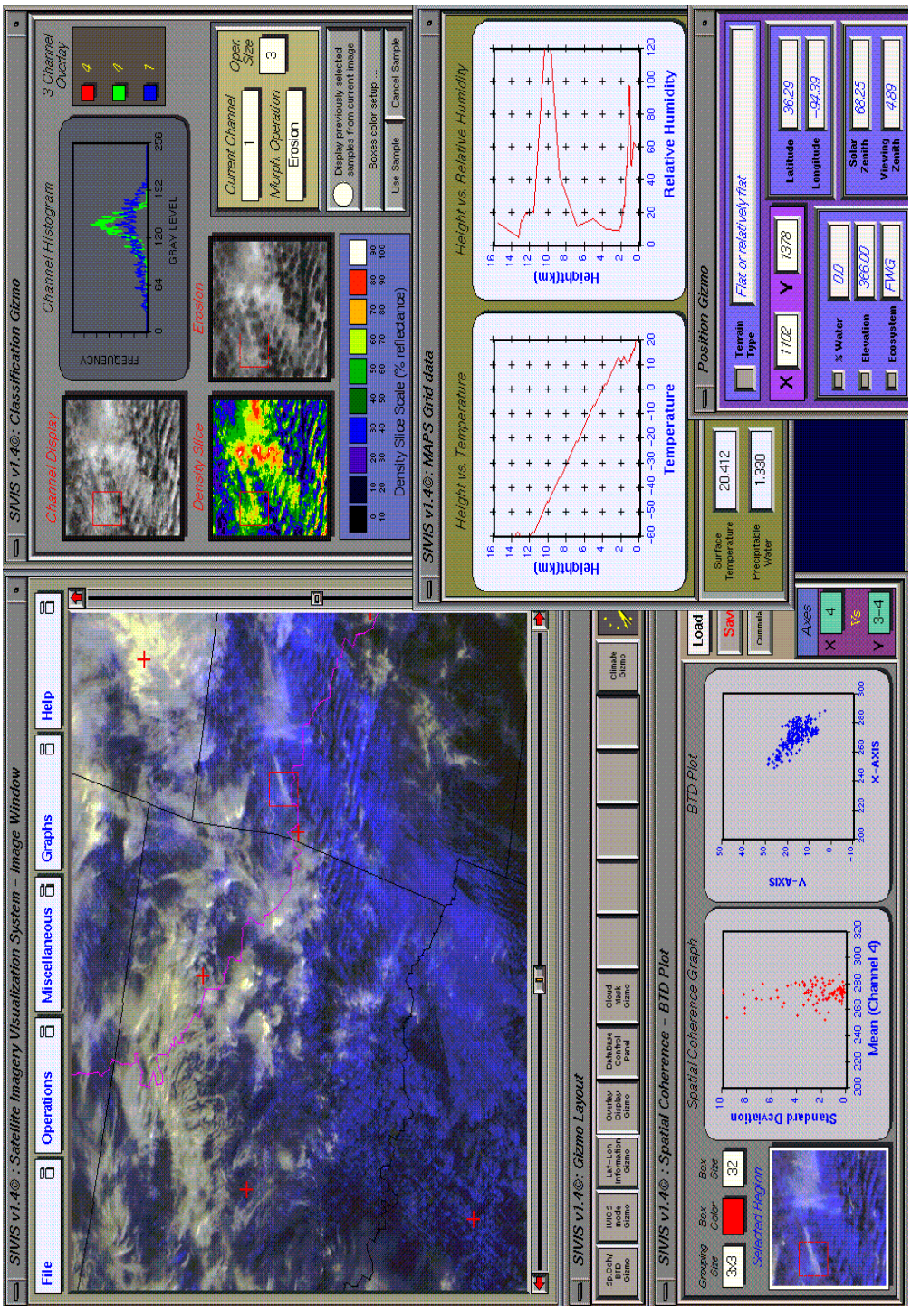


Figure 4.1-1. The Satellite Image Visualization System (SIVIS), Version 1.2.1, is a visualization analysis tool under development to perform both classification, cloud masking, image processing, and other functionality for satellite data analysis. Shown in the upper left hand corner is a segment of a 1.1-km AVHRR image over the southwestern United States. Image processing functions such as histograms, dilation, and density slicing are shown to the right of the image for the boxed region outlined in red in the image to the left. From left to right along the bottom, the boxed region in the image is shown in the bottom left corner, followed by spatial coherence analysis and brightness temperature differences between the 3.7- and 10.8-µm channels. Other functionality includes the ability to overlay coastlines, political boundaries, cloud masks, and navigation information.

range of channel displays and image processing functions. By default all bands are histogram equalized for contrast enhancement. However, any combination of band differences and band ratios can be designed and displayed on command. Additional display features such as principal components, decorrelation stretch, canonical transformations, and edge finding are being implemented.

4.1.4.4. Automated Classification Methods

4.1.4.4.1. Overview. The tests described in section 4.1.3. are to one degree or another based upon thresholds. The NOAA CLAVR algorithm is a decision tree based upon thresholds. The NASA ISCCP algorithm combines uncertainties from multiple tests in a similar fashion.

Texture is a powerful tool for defining the context of a region. Textures were not invoked in the previous analysis because they about double the cpu requirements of the algorithm. However, they are applied to regions which are uncertain. It is probably desirable to apply textures from the outset for difficult regions such as coastlines, deserts, snow covered regions, and areas of variable topography. A texture-flag is set to inform the algorithm manager that texture is to be used.

The textures, along with the results from the previous tests, are then passed to an artificial intelligence classifier. It is a common misconception that AI techniques are operationally more cpu intensive. While AI techniques often do take longer to train, they are no more cpu intensive than are traditional approaches such as Maximum Likelihood. Indeed, many of these AI classifiers are more cpu efficient in the operational mode.

Thresholds are never global. There are always exceptions. For example, the ρ_2/ρ_1 ratio test (section 4.1.3.10.) identifies cloud for values in the range $0.9 < \rho_2/\rho_1 < 1.1$. However, new analyses (McClain 1993) suggest that the lower value may need to be lowered to about 0.8, at least for some cases. The same is true for the other tests. Indeed, it seems unrealistic to label a pixel with $\rho_2/\rho_1 = 1.1$ as cloudy, and a neighboring pixel with the ratio of 1.11 as noncloudy. Rather, as one approaches the threshold limits, the certainty of the labeling becomes more and more uncertain, or “fuzzy.” In situations where the threshold results are uncertain, we will use test the use of either a neural network or fuzzy logic based classification system. In a nutshell, fuzzy logic may be thought of as following the entire decision tree, keeping a running total of the uncertainty accumulated along each path. Instead of hard-and-fast thresholds, fuzzy membership functions are used. Then, at the termination of the decision tree, a “defuzzication” function is applied to the results. Uncertainty estimates of clear and cloudy at each pixel are the outcome.

4.1.4.4.2. Don’t care neural network classifier. A perceptron network consists of an input layer, an output layer, and weights which define linear separating surfaces. Each pattern class C_i is separated by hyperplanes from all other surfaces. It has long been known that this network has very limited capabilities. Consider three tangent circles, each of which represents a class in 2-space. Neither traditional classifiers nor the perceptron network can find separating surfaces to correctly classify the points in the circles. However, the problem can be solved by a three-layer network or by training the network to find pairwise linear separating surfaces. Training a network to produce pairwise linear separating surfaces requires that for any class C_m, the linear function corresponding to the separating hyperplane C_i/C_j will have the value 1 if $m = i$, a value of 0 if $m = j$, and a “don't care” x output otherwise.

For a two-layer network, the surfaces separating the various classes are linear. Similarly, in a multilayer network, nonlinear surfaces separate the classes. Again, pairwise separating surfaces can be constructed using “don’t care” outputs. In the perceptron case, the addition of “don't care” outputs broadens the repertoire of problems the network can solve. For multilayer networks, a different benefit results. The hidden layer allows the decision surfaces to be formed into arbitrarily complex shapes. The surfaces initially are “simple,” and additional training (i.e., iterations) introduces the more complex elements into the separating surface. The network can be trained to find the simpler pairwise separator surfaces

and then construct a more complicated separating surface from pieces of these simpler curves. As a result, fewer iterations are required to train the network. Our studies show that this approach can simplify the training significantly and reduce the training time by two orders of magnitude.

The steps in the algorithm can be summarized as

Step 1: Determine the number of output nodes needed to represent the pattern classes.

Since the network will produce pairwise separating surfaces, the number of output nodes required for this technique is:

$$\binom{N}{2} = \frac{N(N-1)}{2} \tag{4.1-44}$$

where N is the number of classes. In contrast, traditional approaches only require N output nodes.

Step 2: Build the class representations.

Consider the desired node outputs for a class to be a bit string, where each position in the bit string serves as a discriminator between two classes. For each pair of classes, select a bit not previously chosen to be the discriminator and set that bit in one string to 0; set that same bit to 1 in the second string. After all pairs have been processed, fill the remaining positions with “don’t care” symbols. This simple process can be easily automated and introduces only a small overhead penalty to the training algorithm.

For example, a 4-class problem requires six output nodes. Using the above algorithm, one possible assignment of output values to classes can be found in the following table.

Table 4.1-7. Possible Assignment of Output Values

Class	Bit number—					
	1	2	3	4	5	6
1	1	1	1	x	x	x
2	0	x	x	1	1	x
3	x	0	x	0	x	1
4	x	x	0	x	0	0

Note that bit 1 discriminates between class 1 and 2, bit 2 discriminates between class 1 and 3, and so on. The symbol “x” denotes a don’t care value.

Step 3: Train the network.

During training, error is measured at the output nodes and used to adjust the network weights using back-propagation. In our experiments, the error measure

$$\text{Network} - \text{Error} = \sum_k (\text{Actual}_k - \text{Desired}_k)^2 \tag{4.1-45}$$

was used. However, unlike the standard back-propagation algorithm, the above error is not calculated at the nodes which have a don't care designation. The set of weights that will be adjusted during a particular training episode is, therefore, a function of the input pattern. Note, however, that all input to hidden weights are updated.

Step 4: Classify the pattern.

To classify the pattern, simply compare the outputs to the bit strings for each class. Note that an output pattern can match at most one class since there is a discrimination bit for each pair of classes. However, it is possible that an output pattern will not match any class. As with standard back-propagation,

the option exists to force a match by selecting the class to which the output pattern is in closest agreement.

4.1.4.4.3. Cloud mask using a fuzzy logic classifier. The "fuzzy logic" classifier methodology is described in Tovinkere et al. (1993). The classifier uses the concept of class membership to determine what classes are present within a given data array. For the cloud mask process, the initial set of classes will be cloud, land, snow, and water. Since the Tovinkere et al. (1993) study addresses only cloud classification in the Arctic, modifications to the methodology will be necessary for use in a global algorithm. This approach is moving from the conceptual stage to a test stage at the current time.

Class mixtures are often classified as a single class, thereby leading to poor information extraction. This is due to uncertainty in the membership concept of the classical set theory. This representation scheme has difficulty in dealing with elements that partially belong to two or more sets. In order to improve the information representation, the concept of fuzzy set theory has been used. Fuzzy logic is concerned with formal principles of approximate reasoning; i.e., it aims at modeling imprecise modes of reasoning to make decisions in an environment of uncertainty.

The greater expressive power of fuzzy logic derives from the fact that it contains, as special cases, not only the classical two-value and multivalued logical systems but also probability theory and probabilistic logic. The main features of fuzzy logic that differentiate it from traditional logical systems are the following:

1. In two-valued logical systems, a proposition *p* is either true or false. In multivalued logical systems, a proposition may be true or false or have an intermediate truth value.
2. The predicates in two-valued logic are constrained to be crisp in infinite truth value set *T.* In fuzzy logic, truth values are allowed to range over the fuzzy subsets of *T.* Predicates may be either crisp (e.g., "mortal," "even") or fuzzy (e.g., "tired," "tall," "cold").
3. Two-valued as well as multivalued logics allow only two quantifiers: "all" and "none." By contrast, fuzzy logic allows the use of fuzzy quantifiers exemplified by "most," "many," "several," and so on. Such quantifiers may be interpreted as fuzzy numbers that provide an imprecise characterization of the cardinality of one or more fuzzy or nonfuzzy sets. In this way, a fuzzy quantifier may be viewed as a second-order fuzzy predicate. On the basis of this view, fuzzy quantifiers may be used to represent the meaning of propositions containing fuzzy probabilities and thereby make it possible to manipulate probabilities within fuzzy logic.

4.1.4.4.4. The fuzzy expert system (ES). A fuzzy ES includes two other elements, in addition to the components of a conventional system, "fuzzifiers" which convert inputs into their fuzzy representations, and "defuzzifiers" which convert the output of the inference process into a single numerical value within the range of values of the output variable. The numerical output is used to adjust the state of the system being controlled.

A fuzzy control variable may have several states, each state being represented by a membership function. Suppose we are able to classify cloud from clear land and open water by just using the reflectances computed from channel one (CH1) and temperature from channel four (CH4). Figure 4.1-2 shows the different states for these two measures. CH1 is defined by the five albedo states: very low, low, medium, high, and very high. CH4 is defined by the three temperature states: cold, normal, and warm. The albedo measured in CH1 generally is higher for clouds than for land and water. CH4 generally is warm for land and cold for clouds. The above reasoning might lead to the following set of fuzzy rules:

Rule 1: IF CH1 is very low and CH4 is normal THEN class is water

Rule 2: IF CH1 is low and CH4 is warm THEN class is land

Rule 3: IF CH1 is medium and CH4 is cold THEN class is cloud

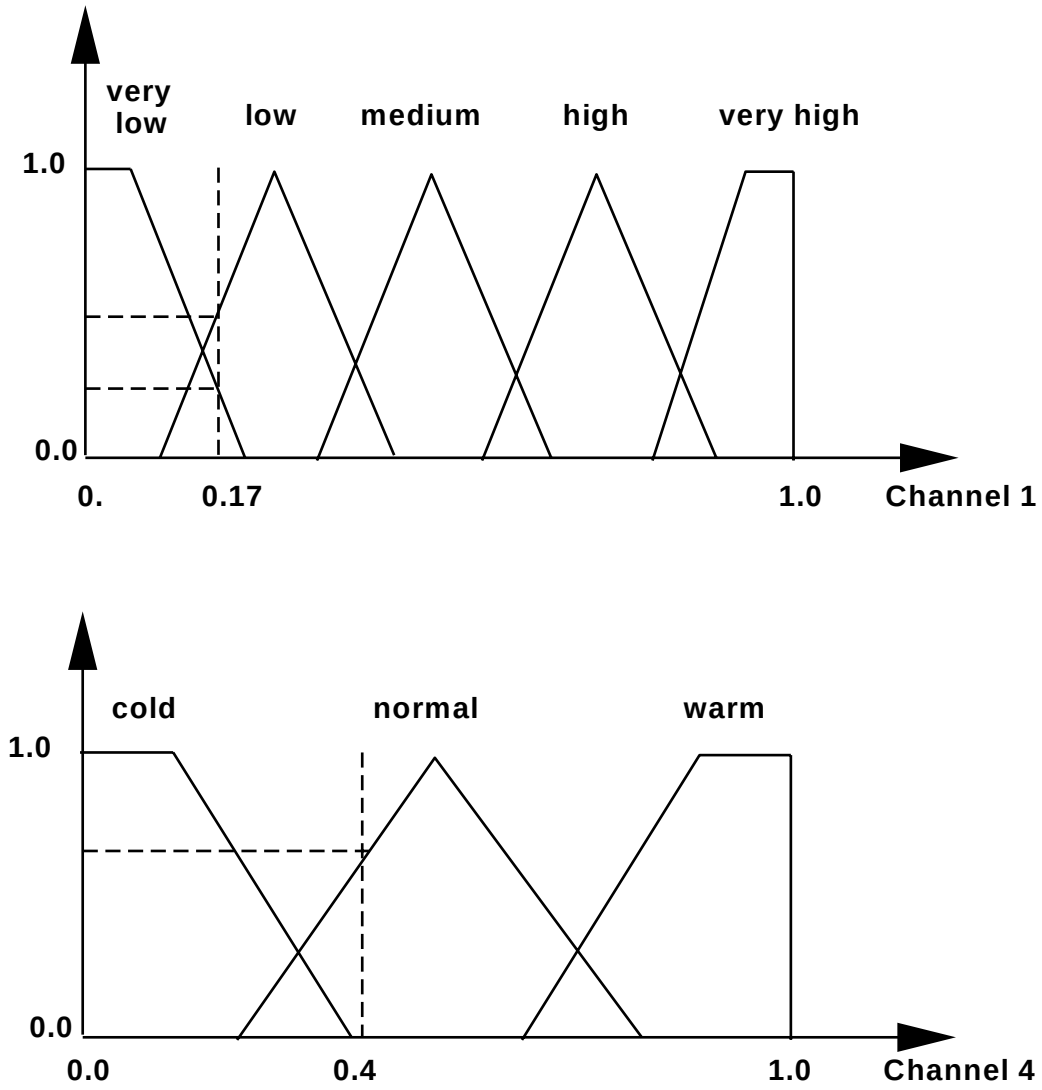


Figure 4.1-2. Schematic showing the concept of class membership in the fuzzy logic classification approach as discussed in section 4.1.4.4.3.

The CH1 reflectance and CH4 temperature values are rescaled to an integer ranging from 0 to 255. As shown in Figure 4.1-2, for a given image sample, the input value for CH1 is 0.17 and 0.4 for CH4; the fuzzifier then computes the degree of membership (DM) for one or more of these fuzzy states. In this case, the states "very low" and "low" of CH1 have membership values of 0.25 and 0.5, respectively. The other states for CH1 are zero. Similarly, the only state of CH4 with a value different from zero is "normal," with a value of 0.60. The confidence level (CL) for each rule is computed by combining the DM's associated with each condition using the following certainty theory formula (Luger and Stubblefield 1989):

$$\mathrm{CL}(C1;C2) = \min[DM(C1);DM(C2)] \tag{4.1-46}$$

where C1 and C2 are the conditions of the rule. The CL for rules 1, 2 and 3 are

Rule 1: min(0.25, 0.60) = 0.25
Rule 2: min(0.5, 0.0) = 0.0
Rule 3: min(0.0, 0.0) = 0.0

Since rule 1 has the higher confidence level, the class selected is “water,” which corresponds to the action of rule 1.

The classification process is performed with the aid of a general fuzzy expert system (GFES). GFES can handle different membership functions for describing the different states of the control variables. These functions are triangular; trapezoidal; one-, two-, and three-dimensional normal distributions; PI function; S function; and elliptical cones. The height for all these functions is equal to 1, since any membership function can have any real value between 0 and 1. The multivariate normal distribution is an extension of the one-dimensional normal distribution.

Usually, triangular, trapezoidal, PI, and S functions (Giarratano and Riley 1990) are used for the definition of fuzzy ES’s. Since our classifier uses control variables which are often assumed to belong to normal distributions, we have extended the usual set of function types to accommodate the definition of fuzzy states with one- and multi-dimensional normal distributions. Our experiments show that by increasing the number of dimensions, the classifier is able to separate better the different classes.

Three input files are required to run GFES: a control variable file, a rule file, and a facts file. The control variable file requires the following information for each control variable: the name of the variable (e.g., temperature), the type of membership function used to approximate the mean and standard deviation of the feature vector, the number of states, the state names (e.g., hot, cold), and the values that define each state’s membership function. The output consists of the class or classes present in the region or pixel with an associated value representing the percentage of the class within the region or pixel.

4.1.5. Cloud Mask Algorithm Description

4.1.5.1. Ancillary Data Set Requirements

A number of preprocessing steps will be made to the AVHRR GAC data before the cloud masking algorithm is applied. An example of navigated GAC data is shown in Figure 4.1-3. These preprocessing steps are described below:

1. The NAVY 10-min database is a 1080×2160 array covering 180° in latitude from North to South Pole and 360° in longitude (Fig. 4.1-4). This database provides surface elevation (Fig. 4.1-4), the percentage (an integer between 0 and 100) of water in the 10-min box, and character type as shown in Table 4.1-8 and Figure 4.1-5. Note that multiple characteristics are defined in this system; an example is code 14 = flat lake country or atoll.

Table 4.1-8. Navy Character Map that Provides a General Surface Classification

Code	Feature
0	Salt or lake bed
1	Flat or relatively flat
2	Desert (or for high latitudes, glaciers, or permanent ice)
3	Marsh
4	Lake country or atoll
5	Major valleys or river beds
6	Isolated mountains, ridge, or peak
7	Low mountains
8	Mountainous
9	Extremely rugged mountains
62	Ocean

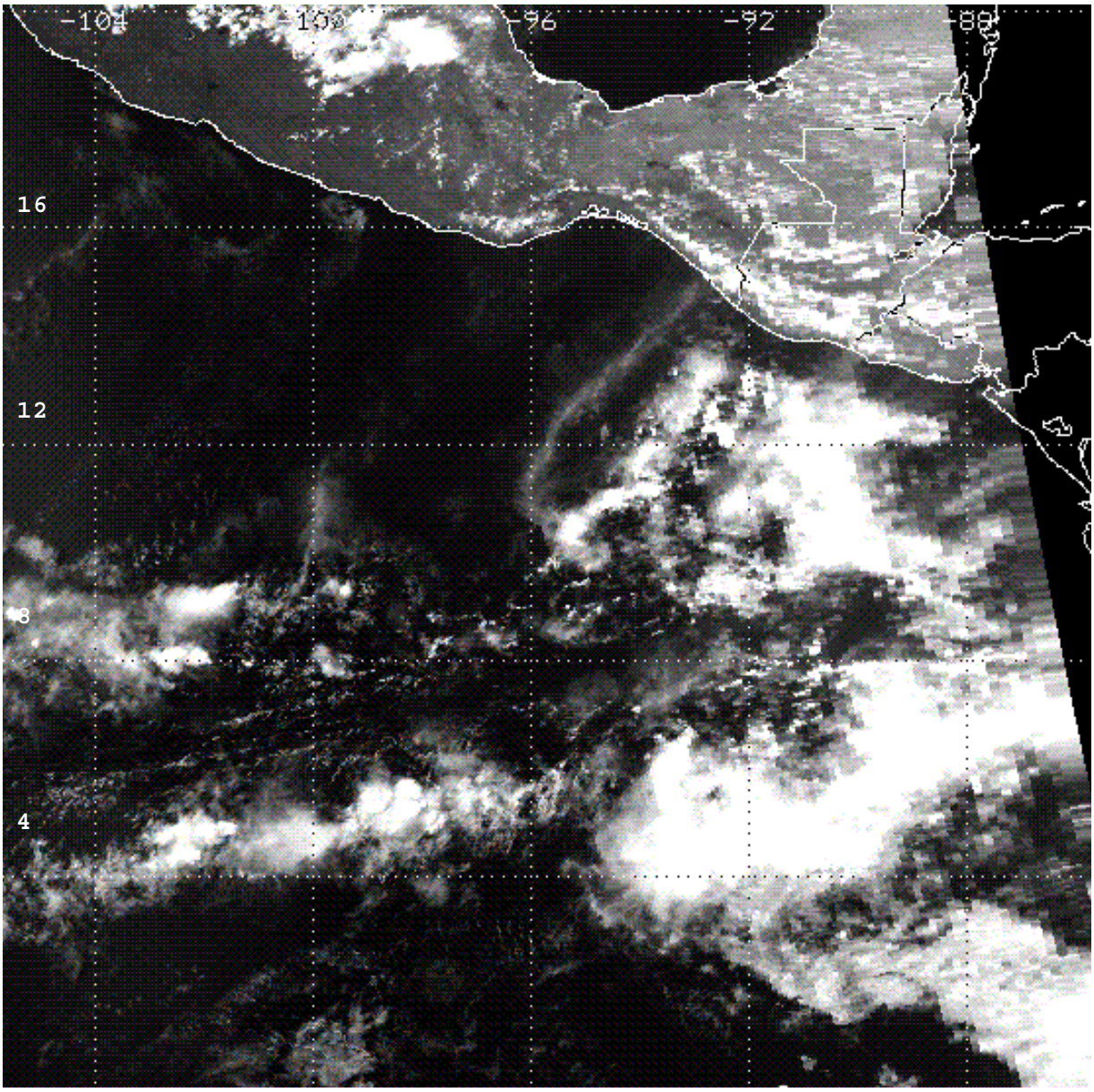


Figure 4.1-3. A navigated AVHRR GAC (~4 km resolution) image from 88°W to 104°W longitude and 2°N to 20°N latitude. Overlaid upon the satellite radiometric data are the coastline boundaries.

2. The EPA Global Ecosystem (WE1.4D) Database also is a 1080×2160 byte array which contains 59 different ecosystems classes (Fig. 4.1-6).
3. The US NAVY/NOAA Sea Ice Product provides weekly reports of fractional ice coverage at spatial resolution of about 18 km.
4. The NOAA Snow Data Product provides weekly report of snow cover at a spatial resolution of 150–200 km; snow is reported if the grid cell is more than 50% covered.
5. The NMC 3-hour surface analyses of temperature and wind speed.

Ancillary data will be subset into scenes of about 1000 lines each consisting of 409 pixels (the full swath of AVHRR GAC data). First each pixel in the scene will be tagged as being land or water, and if land, a land/water percentage. Second, each land pixel will be designated as relatively flat, valley,

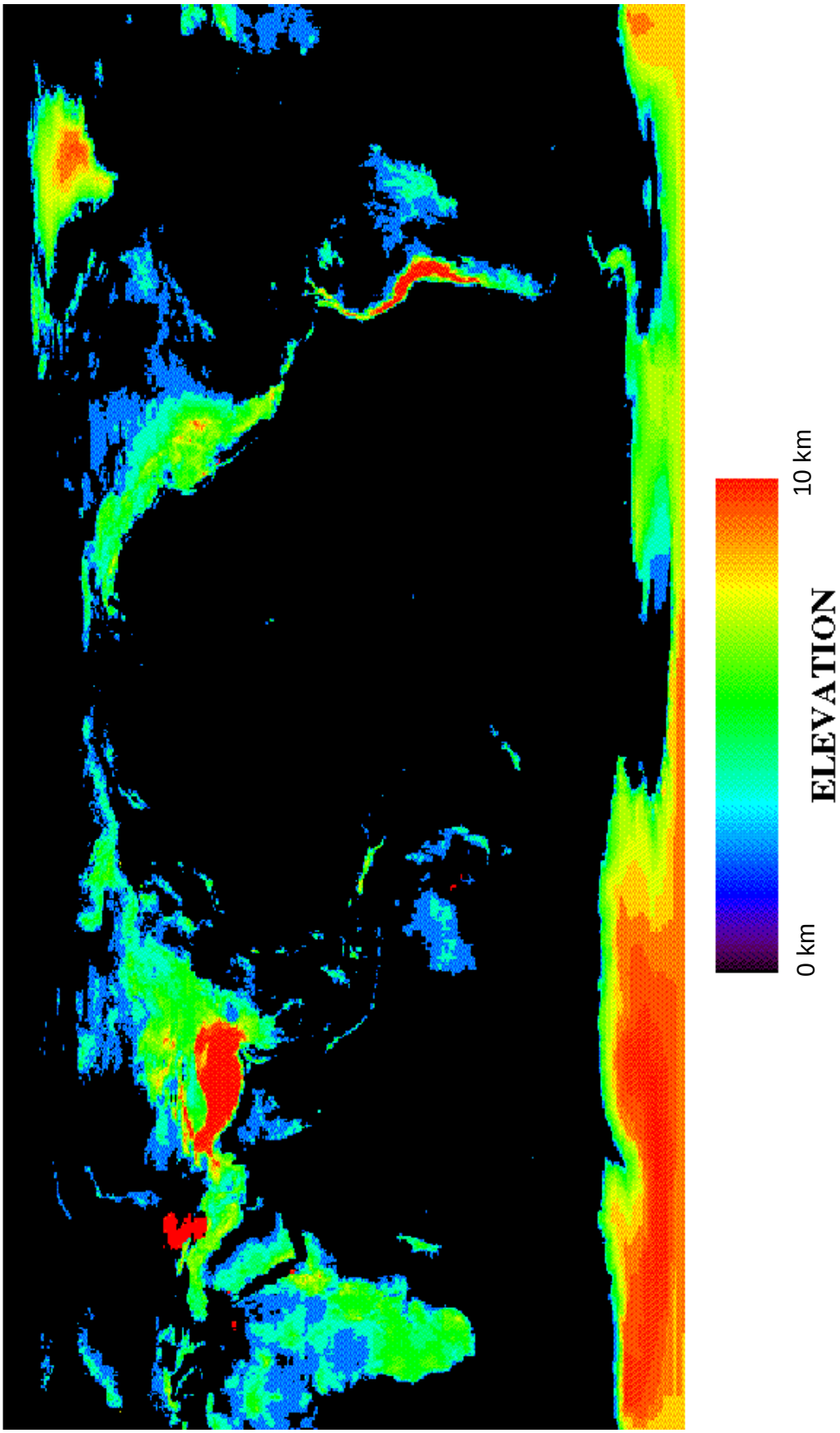


Figure 4.1-4. False color image of the global elevation map derived from the Navy 10-min database discussed in section 4.1.5.1.

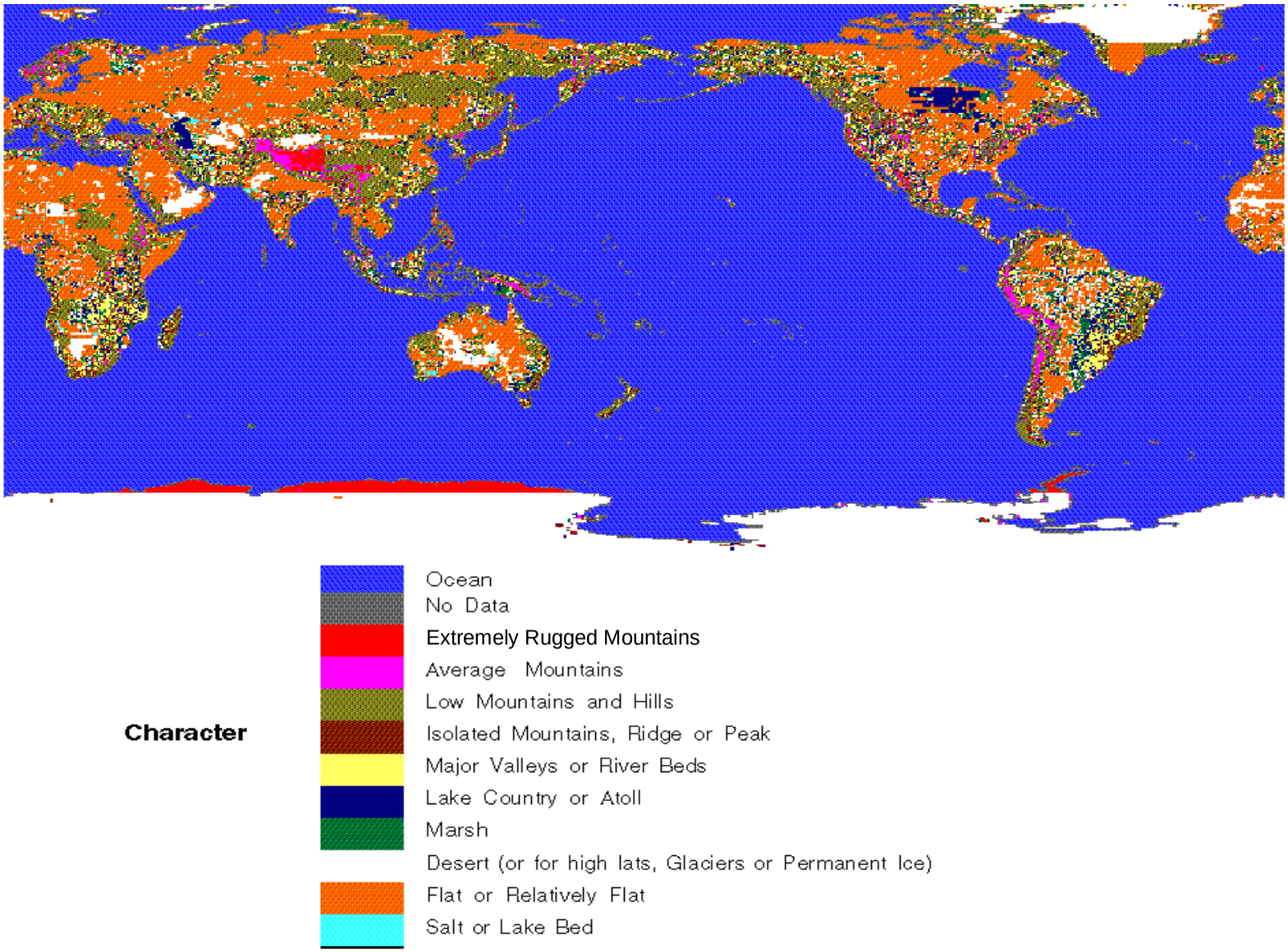


Figure 4.1-5. False-color image of the global surface character type derived from the Navy database discussed in section 4.1.5.1.

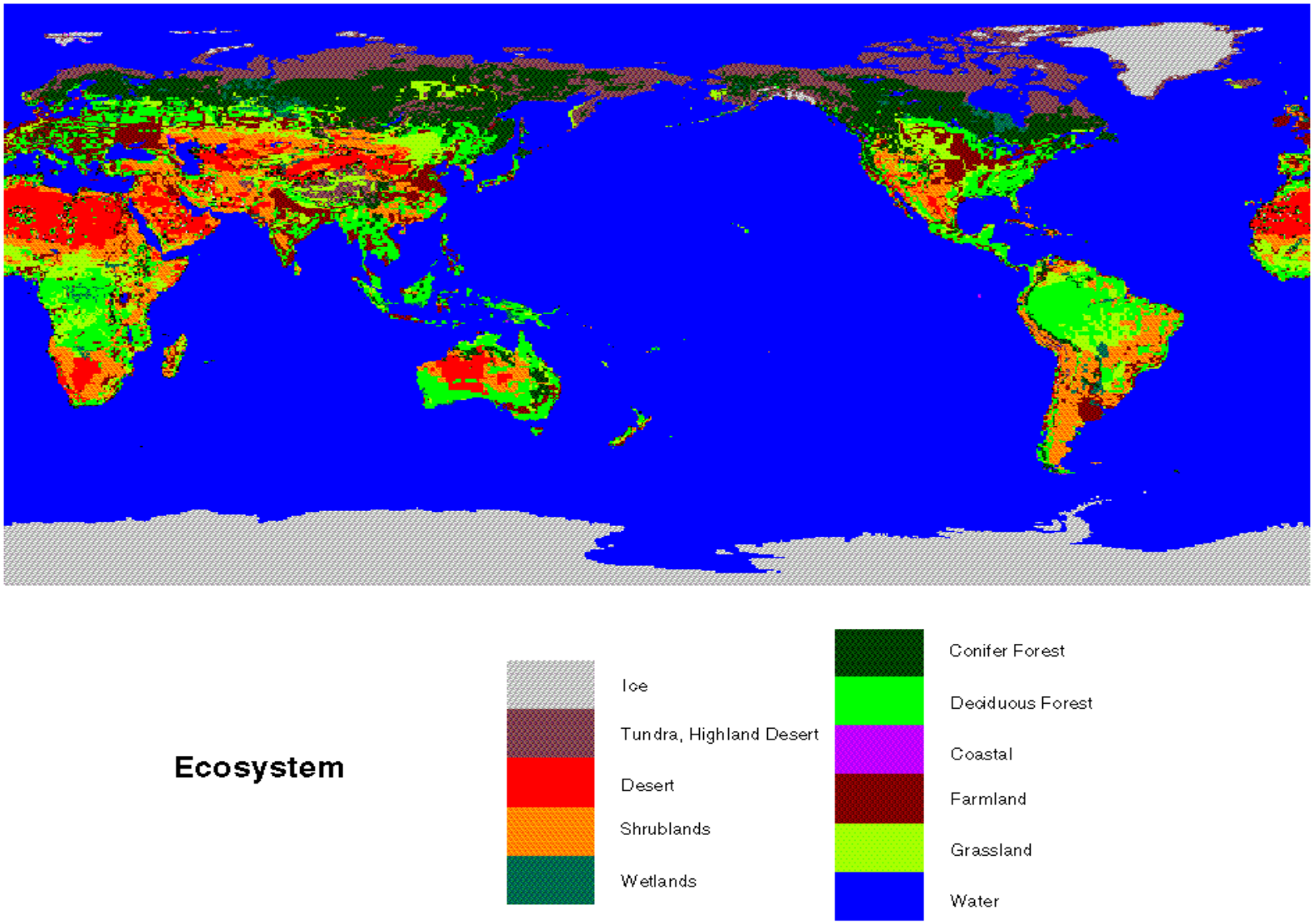


Figure 4.1-6. False-color image of the EPA global ecosystem map as discussed in section 4.1.5.1. The number of classes has been reduced to facilitate ease of interpretation in this image.

isolated mountainous region, low mountains or hills, generally mountainous, or extremely rugged mountains. From the NOAA Snow Data Product each pixel will be designated as probably/probably not snow covered. Each land pixel will be classified as to its ecosystem, along with a more general ecosystem classification of urban, forest, woodland, grassland, shrubland, tundra, arid vegetation and highland vegetation. Ocean regions will be classified as water, coastline (including islands), possibility of isolated icebergs, marginal ice zone, and nearly solid sea ice (leads may be present).

4.1.5.2. Cloud Mask for Daylight Oceanic Areas

The procedures outlined here will be applied between 60°N and 60°S for oceanic regions and seasonally for large lakes (for which ice is very unlikely). Solar zenith angles are constrained to be less than 85°. The nighttime algorithm is used for sun glint areas.

4.1.5.2.1. SERCAA Sun glint test. The Sun glint test is applied over water surfaces. Reflectance over open water is strongly influenced by illumination and viewing geometry. Sun glint is also a function of surface wind. To determine whether sun glint is present, we will implement a series of tests that are used operationally by SERCAA. Two sets of tests are run. The first set of three tests determine if the background surface type and solar/satellite geometry will support sun glint. The three tests are

1. Surface type must be water
2. $|\theta_o - \theta| < \text{Threshold}_{\text{zenith}}$
3. Azimuthal angle must fall within a certain range

If pixels passed the first set of tests, a second set of spectral tests is applied:

4. The reflectance must be high in the visible channels
5. The 3.7-μm brightness temperature must be high (near saturation)
6. The IR brightness temperature must be relatively high (not indicative of cold clouds)

Pixels that pass the sun glint test but have little illumination due to high solar zenith angle ($\theta_o > 85°$) will be passed to the nighttime ocean algorithm.

4.1.5.2.2. Cloud mask tests. The hierarchical approach we will use in the Version 1 code has six stages:

1. Filter pixels that have sun glint
2. Filter pixels that have high solar zenith angle
3. Spatial coherence (to identify clear and cloudy pixels in a 256-km by 256-km region)
4. Apply masking tests to individual pixels in the following order:
   - Imager pixel IR channel threshold test
   - Imager pixel NIR-IR brightness temperature difference threshold tests
   - Imager pixel visible channel reflectance threshold tests
   - Imager pixel visible channel ratio test
5. Apply masking tests to pixel arrays, or tiles, in the following order
   - Spatial contrast
   - Spatial/temporal uniformity
   - Artificial intelligence classification
   - IR clear-sky composite consistency test
   - VIS clear-sky composite consistency test
6. Determine final result of mask tests for each pixel

4.1.5.3. Cloud Mask for Daylight Land Areas

This portion of the algorithm is applied to land areas at latitudes from 60°N to 60°S, including islands. The hierarchy of cloud mask algorithm application is as follows.

1. Check ecosystem (vegetation) map and land/water percentages data bases. Determine land/water percentage and vegetation for every pixel.

2a. Compute sun glint angles for every pixel. Within the sun glint region the percent water database is examined (at 10-min resolution). If water is present, then a flag is set indicating possible sun glint. Eliminate use of visible channels for sun glint pixels and apply nighttime algorithm to those pixels.

2b. Over land pixels that are vegetated, compute scattering angle. If the scattering angle is close to 0°, there may be enhancements in the visible channel reflectances (hot spots). The scattering angle is computed from

$$\cos\gamma = \sin\theta \sin\theta_o \cos\phi - \cos\theta \cos\theta_o$$

3. Check surface elevation, clear-sky radiance/temperature data base, and snow-cover data base.
4. If the pixel or subregion is snow-covered, rugged terrain, coastline, or other special cases, pass imager data to automated classification algorithms.
5. Apply spatial coherence to large-scale 256-km by 256-km regions.
6. Apply masking tests to individual pixels in the following order
   - Imager pixel IR channel threshold tests
   - Imager pixel NIR-IR brightness temperature difference threshold tests
   - Imager pixel visible channel reflectance threshold tests.
7. Apply masking tests to pixel arrays, or tiles, in the following order
   - Spatial contrast
   - Spatial/temporal uniformity
   - IR clear-sky composite consistency test
   - VIS clear-sky composite consistency test
   - Artificial intelligence classification
8. Determine final result of mask tests for each pixel.

The databases are examined for elevation characteristics and ecosystem type. A separate database is examined for probability of snow. The snow probability index is set high if either the NOAA Snow Data Product is positive or if snow was indicated on the previous clear day. If the previous day was not clear, then local regions of the same ecosystem type are examined. If these regions also were cloudy on the previous day, then the pixel and its local regions are examined for two additional prior days. If these tests fail, then the snow flag is set to a low value. This test is seasonal; it is not run for tropical regions (except for regions of high elevation) or during summer months. NMC analyses also are examined. Prior days with surface temperatures > 50°C decrease the snow probability index values. On the other hand, during the late fall to early spring in the mid- to high-latitudes, cloud cover on the previous day coupled with surface temperatures in the NMC surface analysis increases the snow probability index. Note that this index has values between 0 and 100. As explained in section 4.2.3., artificial intelligence and texture analysis are applied to regions which are uncertain. The fuzzy logic algorithm directly uses the snow probability index values.

4.1.5.4. Cloud Mask for Nighttime Ocean and Land Scenes

These algorithms are applied to all surface regions between 60°N and 60°S, for solar zenith angles greater than 85°, and for daytime pixels labeled as having sun glint, vegetation hot spots, or low illumination angles. Spatial coherence is used over oceans and over land areas of the same ecosystem.

1. The initial step is to label each pixel with ecosystem, elevation, surface characteristics, snow/ice, and land/water percentage. If snow-covered, set flag.
2. Retrieve short-term clear night (within last 72 hours) and radiance values for channels 3, 4, and 5. If there are no short-term clear night data available, use longer-term values. If the subregion is

snow-covered, rugged terrain, coastline, or under the heading of other special cases, pass the sub-region to automated classification algorithms.
3. Apply spatial coherence to large-scale 256-km by 256-km regions, noting that the algorithm will be applied only to areas having the same ecosystem (e.g. water or ocean).
4. Apply masking tests in the following order.
5. Apply masking tests to individual pixels in the following order
   - Imager pixel IR channel threshold tests
   - Imager pixel NIR-IR brightness temperature difference threshold tests
6. Apply masking tests to pixel arrays, or tiles, in the following order
   - Spatial contrast
   - Spatial/temporal uniformity
   - IR clear-sky composite consistency test
   - Artificial intelligence classification

4.1.5.5. Daytime Polar Region Cloud Mask Analysis

A daytime polar region cloud mask algorithm is currently under development. The daytime polar scene classification system currently separates pixel subarrays into the following classes:

1. Water
2. Solid sea ice or snow-covered land
3. Broken sea ice
4. Snow-covered mountains (or regions of high relief)
5. Stratus-type clouds over water
6. Stratus-type cloud over ice
7. Cirrus clouds over ice
8. Cumulus clouds over water
9. Multilayer cloudiness
10. Nonsnow-covered land

These classes need to be expanded somewhat to include, for example, cirrus clouds over water. To date, high accuracies are achieved for pure classes. However, additional work is in progress to extend the classes and to include a broader range of textural measures.

The current algorithm is applied poleward of 60°N and 60°S and is based upon Ebert (1987, 1989), Key and Barry (1989), Welch et al. (1992) and Tovinkere et al. (1993). The following eight spectral and textural measures were used in a polar scene identification study by Tovinkere et al. (1993):

1. $\rho_1 - \rho_2$
2. Low 3
3. ρ_3/ρ_1
4. Mean ASM 3
5. Mean 1
6. Mean 4
7. Max Ent1
8. Max Ent4

Measure 1. $\rho_1 - \rho_2$: The reflectance difference between channels 1 and 2. This measure is positive for classes with snow and ice surfaces and negative for land. The reflectance difference $\rho_1 - \rho_2$ tends to have a positive value for the cloud classes and for water and snow and negative for land.

Measure 2. Low 3: This is the percentage of pixels in channel 3 that have a reflectance less than 10%. This is the greatest for water and snow and least for stratus and stratocumulus cloudiness.

Measure 3. ρ_3/ρ_1: The mean reflectances of channels 1 and 3 are computed, and the ratio is calculated using these values. This value is low for ice covered surfaces and cirrus cloudiness.

Measure 4. Mean ASM3: Angular second moment of channel 3. This textural feature is a measure of homogeneity in the scene. This measure is smallest when all the gray levels occur with equal probability.

Measure 5. Mean 1: This is the mean reflectance of channel 1.

Measure 6. Mean 4: This is the mean brightness temperature computed from channel 4.

Measure 7. Max Ent1: This is a measure of disorder in the scene. The entropy measure is calculated from channel 1. It has low values for water, solid sea ice, and land.

Measure 8. Max Ent4: This is the entropy measure of the region in channel 4. Max Ent4 has low values for classes which have display distinct scales of organization and relatively uniform temperatures.

These features are defined for daylight ($\theta_o < 85°$). Also, the pixel arrays are defined over a single ecosystem type. The pixel array may be modified to suit individual regions, by altering the size the array to be larger or smaller and by altering the shape of the region as necessary. If a single ecosystem type cannot be defined for a given region, then a broader categorization of forest, tundra, etc., is used.

The Navy weekly 18-km sea ice product is utilized to define the marginal ice zone. Oceanic regions within 100 km of the ice edge are labeled as probably broken sea ice. Regions further poleward are labeled with distance as more and more probable of being solid sea ice. Regions in the opposite direction are labeled with distance as less and less probable of being broken sea ice.

The NMC surface temperature analysis is used for consistency checks. For example, high surface temperatures (>273 K) indicate ice melt, probable melt ponds, and lower ice/snow reflectivities. On the other hand, very low surface temperatures generally are consistent with the lack of open water and with higher surface reflectivities. Such low surface temperatures also mean that the various thermal tests need to be applied over more restricted domains.

Artificial intelligence classifiers may be applied from the outset to provide the context of the local region. Once the surface is known (water, solid sea ice or snow-covered land, broken sea ice, land, patchy snow over land, frozen lakes and rivers) and once the basic cloud cover is known (stratus, stratocumulus, cumulus, cirrus, or multilayer), then the previously defined tests may be used at the pixel level. Somewhat different sets of the tests described previously are used for each of the various scenarios.

In the near future (F.Y. 95), we will be deriving and applying new polar cloud mask algorithms. The polar algorithm will be exercised using both AVHRR 1-km and 4-km data over both poles. The final Version 1 algorithm is expected to be modified extensively over the next year.

4.1.6. Short-Term and Long-Term Clear-Sky Composite Maps

4.1.6.1. ISCCP Clear-Sky Composite

The ISCCP developed clear-sky reflectance and temperature composites to detect clouds over a given 32-km square area by comparing the pixel radiances to the clear-sky composite values with some added thresholds (Rossow and Garder 1993). These composites are based on the observation that variations in VIS clear reflectances usually are smaller in time than in space, especially over land. Variations of surface VIS reflectances generally are smaller than variations of cloud reflectances. Therefore, it is assumed that the characteristic shape of the darker part of the VIS radiance distribution is at most weakly dependent upon surface type (Seze and Rossow 1991a, b). The minimum reflectance value for channel 1 is used to estimate clear values. Corrections to the minimum values are inferred from the shapes of the visible reflectance distribution associated with different surface types.

Rossow and Garder (1993) classify the surface into nine types depending on the time scale and magnitude of the reflectance variations (see Tables 5 and 6). The clear sky reflectance values for land and ocean regions whose surface characteristics vary the most rapidly are estimated using short-term values of ρ_{min} such that $\rho_{cs} = \rho_{min}(ST) + DEL2$. Sparsely vegetated surfaces generally exhibit more spatial variability than heavily vegetated surfaces (cf. Matthews and Rossow 1987), but are also generally less cloudy. For these, $\rho_{cs} = \rho_{min}(LT) + DEL2$. Vegetated areas show less small-scale spatial variability. They also tend to be more uniform from one geographic location to another. For vegetated regions, the clear-sky reflectance is determined by first calculating $\rho_{cs} = \rho_{min}(ST) + DEL2$. Then the individual pixel reflectance values within each latitude zone are compared to the distribution of ρ_{cs} values for the same ecosystem type; they are required to be within DEL1 of the distribution mode value, ρ_{mode}.

Similar assumptions are used for the determination of T_{cs} fields. The time scales of VIS and IR variability for different classes and the associated ISCCP thresholds are shown in Tables 4.1-9 through 4.1-12.

Table 4.1-9. Time Scales of Variability for Different Surface Types for Visible Channel

VIS classes	Short term (ST)	Long term (LT)
Ocean	-	30 days
Lakes	-	15 days
Polar ocean (open water)	-	15 days
Ice-covered water	5 days	-
Forests, woodlands, shrublands	-	30 days
Grasslands, tundra	-	30 days
Arid vegetation, deserts	-	30 days
Polar land (snow free)	-	15 days
Snow- or ice-covered land	5 days	-

Table 4.1-10. Values Used in VIS Clear-Sky Composite Logic (after Rossow and Garder 1993); VIS Threshold Values are in Percent Reflectance

IR surface types	DEL1	DEL2
Ocean, near-coastal, lakes	3.0	1.5
Forests, woodlands, shrubland	6.0	3.5
Grasslands, tundra	-	3.5
Arid vegetation, deserts	-	3.5
Ice-covered water	-	5.0
Ice- or snow-covered land	-	5.0

Table 4.1-11. Time Scales of Variability for Different Surface Types for IR Channels (after Rossow and Garder 1993)

IR classes	Short term (ST)	Long Term (LT)
Open ocean	15 days	30 days
Near-coastal ocean and lakes	5 days	15 days
Polar seas and ice-covered water	5 days	15 days
Land	5 days	15 days
High and rough topography land	5 days	15 days
Ice- or snow-covered land	5 days	15 days

Table 4.1-12. Test Values Used in IR Composite Logic (after Rossow and Garder 1993); IR Values are in Kelvins

IR surface types	DEL1	DEL2	DEL3
Ocean	2.0	2.0	2.5
Near-coastal ocean, lakes	3.0	3.0	4.0
Ice-covered water	3.0	3.0	4.0
Land	6.0	5.0	8.0
High and rough topography	9.0	7.0	11.0
Ice- or snow-covered land	9.0	7.0	11.0

One of the primary difficulties in using the ISCCP approach as currently formulated is the angular dependence of clear-sky reflectance. Although cross-track scanning Sun-synchronous satellites such as the NOAA-AVHRR repeat the angular viewing conditions on a regular cycle, the solar zenith angle slowly varies and the cloudiness conditions may prevent the determination of clear-sky reflectance at some points in cycle. The ISCCP relies on an empirical bidirectional reflectance model for clear-sky ocean reflectance (Minnis and Harrison 1984a). Thus, over ocean, the angular problems are minimized. Over land, the ISCCP assumes isotropic clear-sky reflectance, although it has been established that the anisotropy of land scenes is significant (e.g., Kriebel 1978; Tarpley 1979; Minnis and Harrison 1984c; Suttles et al. 1988). For $\theta_o < 85°$, the vegetated land clear-sky anisotropic reflectance factor $R(k, \theta_o, \theta, \phi)$, where k is a surface type that can vary from 0.6 to 1.6 (e.g., Suttles et al. 1988) for $\theta < 70°$. Thus, there is the potential for clear-sky reflectance errors as great as 300% if one assumes that the measurement taken at a particular set of viewing conditions represents the reflectance at all viewing angles for a given value of θ_o. Systematic changes of albedo with θ_o are also not considered for land surfaces. The reflectance anisotropy over snow and desert scenes is generally not as great as that over vegetated surfaces, but the absolute changes in reflectance are as great because of the higher albedos over these surfaces.

The CERES processing will begin with a set of global clear-sky radiances matched to the Navy 10-min database at a 3-hourly resolution. Thus, a relatively high-resolution clear-sky field is required. The clear-sky radiance maps currently available from the ISCCP are the C1 datasets that have a 250-km and 3-hour resolution and that lack the anisotropy corrections noted above. The following processing steps using the ISCCP data are applied to historical AVHRR data to obtain the clear-sky radiances at the higher spatial resolution and to account for reflectance anisotropy.

From the ISCCP C1 data, the clear-sky reflectance at a given day d, synoptic hour h, nominal regional latitude Θ_{C1}, and longitude Φ_{C1} is $\rho_{csC1}(\theta_o, \theta, \phi, \Theta_{C1}, \Phi_{C1}, h, d)$. The corresponding clear-sky albedo is

$$\alpha_{csC1}(k, \theta_o, \Theta_{C1}, h, d) = \frac{\rho_{csC1}(k, \theta_o, \theta, \phi, \Theta_{C1}, \Phi_{C1}, h, d)}{R(k, \theta_o, \theta, \phi)} \tag{4.1-47}$$

where the value of R is taken from Minnis and Harrison (1984a) for vegetated land and from Suttles et al. (1988) for snow and desert. Over ocean, α_{csC1} is estimated using an updated version of the clear ocean bidirectional reflectance model of Minnis and Harrison (1984a). The updated version includes calibrated data from more angles than the original model. The value of $\alpha_{csC1}(\text{ocean}, \theta_o = 0) = 0.045$. The standard deviation of α_{csC1} is $\sigma_{\alpha C1}(k, \theta_o, \theta, \Theta_{C1}, \Phi_{C1}, h, d)$. For mixed land-water regions, the reflectance for the land portion is, leaving off the dependence on the parameters $\theta_o, \theta, \phi, \Theta_{C1}, \Phi_{C1}, h,$ and d:

$$\rho_{csC1}(\text{land}) = \frac{[\rho_{csC1} - (1 - f_{\text{land}})\rho_{csC1}(\text{ocean})]}{f_{\text{land}}} \tag{4.1-48}$$

where f_{land} is the land fraction in the $C1$ region. The standard deviation of α_{csC1} is $\sigma_{\alpha C1}(k, \theta_o, \theta, \Theta_{C1}, \Phi_{C1}, h, d)$. Average values of these parameters, $\langle\alpha_{csC1}(k, \Theta_{C1}, \Phi_{C1}, h)\rangle$ and $\langle\sigma_{\alpha C1}(k, \Theta_{C1}, \Phi_{C1}, h)\rangle$ are computed for each region and month.

The corresponding ISCCP clear-sky, 11-μm temperatures and their standard deviations are $T_{csC1}(k, \Theta_{C1}, \Phi_{C1}, h, d)$ and $\sigma_{TC1}(k, \Theta_{C1}, \Phi_{C1}, h, d)$, respectively. Monthly mean values, $\langle T_{csC1}(k, \Theta_{C1}, \Phi_{C1}, h)\rangle$ and $\langle\sigma_{TC}(k, \Theta_{C1}, \Phi_{C1}, h)\rangle$, are also computed for these parameters. All 10-min regions falling within the 250-km $C1$ region are initially assigned the clear-sky radiances for the ISCCP region if the $C1$ region is all land or water. If the $C1$ region is mixed, the 10-min boxes that are entirely water are assigned the empirical model values for ocean albedo and the land boxes are given the land clear-sky albedos computed from (4.1-48) and (4.1-47). The coastal boxes retain the nominal $C1$ albedo. The $C1$ temperature is assigned to the 10-min box regardless of the geotype. These mean values constitute the starting point for the development of the high-resolution clear-sky radiance set.

To derive the high-resolution dataset, AVHRR GAC data are analyzed to determine if the pixels belonging to a particular 10-min box are clear. During a given AVHRR orbit at time t, the reflectance ρ, and 11-μm brightness temperature T_{B4} of all pixels located within a given 10-min box are compared to the monthly mean dataset. The pixels are assumed to be clear if, again leaving off the dependence on the parameters k, θ_o, θ, ϕ, Θ_{C1}, Φ_{C1}, h, and d:

$$\rho(t) < \rho_{csC1}(t) + 2\sigma_{\alpha C1} \tag{4.1-49}$$

and

$$T_{B4} > T_{csC1} - 2\sigma_{TC1} \tag{4.1-50}$$

where $h < t < h + 1$, and

$$\rho_{csC1}(k, \theta_o, \theta, \phi, \Theta_{C1}, \Phi_{C1}, t) = R(k, \theta_o, \theta, \phi)\alpha_{csC1}(k, \theta_o, \Theta_{C1}, \Phi_{C1}, t) \tag{4.1-51}$$

The last term in (4.1-51) is the albedo at time t found by linearly interpolating the $C1$ albedos in time. First, the albedos are extended to θ_o using the directional reflectance models derived from the results of Minnis and Harrison (1984a, c) based on the mean θ_o at the synoptic times. This approach is the same employed by the ERBE time-space averaging subsystem (see Brooks et al. 1986). Simple linear inter-polation is used to determine the expected standard deviation. When albedos do not exist at h or $h + 1$, the available albedo is extrapolated to t using the directional reflectance models. Over snow scenes, additional tests using the $T_{B3} - T_{B4}$ differences supplement the standard clear-sky tests to insure that the scene is cloud free. The expected clear-sky temperature, $T_{csC1}(k, \theta_o, \Theta_{C1}, \Phi_{C1}, t)$ and its standard deviation are interpolated using linear interpolation.

For some areas, such as deserts, the surface emittance at 3.7 μm will not be unity. When the surface emittance is less than unity, the task of determining the expected clear-sky 3.7-μm brightness temperature will be difficult. For this reason, we will develop a surface emittance map at 3.7 μm using nighttime data so that there is no solar contribution. The effective surface emittance ε_{3s} for channel 3 is also estimated for each 10-min box by first correcting the nighttime clear-sky values of T_{B3} and T_{B4} for water vapor attenuation. Assuming that the clear-sky downwelling radiance is zero for channel 3 and $\varepsilon_{4s} = 1$, then $\varepsilon_{3s}s = [B_3(T_{B4s}) - B_3(T_{B3s})] / B_3(T_{B4s})$, where the subscript s indicates values at the surface.

The pixel values selected as clear are then analyzed as in Minnis et al. (1987) to determine an estimate of ρ_{cs} and T_{cs} for the 10-min box and new values for their standard deviations. The procedure is reversed to estimate the clear albedo and temperature at the nearest synoptic hour. These new values plus the mean channel-3 emittances are then used to construct a new clear-sky map. The results from different days at a given h are averaged to yield the new detailed clear-sky fields that will become the initial CERES clear-sky radiance fields.

Examples of applying this procedure to a day of October 1986 NOAA-9 AVHRR data are shown in Figs. 4.1-7–4.1-10. The initial clear-sky reflectance field based solely on the ISCCP land "albedos" and the ocean reflectance model have a somewhat blocky appearance due to the low-resolution of the *C*1 dataset. The scattering of some of the data values near the orbit overlaps is due to overwriting of previous results by pixels in the following orbit. The ocean model produces a realistic pattern of reflectance including the distinct sun glint areas. Application of the clear-sky procedure yields a somewhat finer resolution of various features such as the Arabian Peninsula and the Pampas region in South America. Bright areas of sun glint appear in the middle of the predicted sun glint during some orbits. Changes did not occur in many areas because of clouds. The clear-sky temperature fields (Figs. 4.1-9 and 4.1-10) show even more dramatic changes because of more local variability, especially over land.

The procedure used to produce the results in Figures 4.1-8 and 4.1.10 will be applied to the AVHRR data for months during four different seasons. Over some particularly clear areas, the resulting means for a given hour will be examined closely to determine the sensitivity of the technique to the values of *R*. New anisotropic reflectance and thermal infrared limb-darkening models will be tested as they are developed. This methodology will be continuously refined prior to the TRMM launch.

The logic employed here will be combined with the other clear-sky detection methods and with a modified version of the ISCCP approach to provide updates of clear-sky radiances during CERES on the time scales suggested in the ISCCP method. The CERES clear-sky composite relies on high-resolution data applied to a higher-resolution grid than that employed by the ISCCP. Thus, accounting for local variability becomes very important. The ISCCP thresholds that bound the clear-sky domain for a particular surface category will be used as guidelines and as default values for the CERES clear-sky composite development. The local standard deviations in the clear-sky radiances computed using the above analysis procedure on preflight AVHRR data will be used to set the thresholds for cloud detection during CERES.

4.1.7. Version 2: Future Directions

4.1.7.1. Detection of Cloud Shadows

The detection of cloud shadows is a problem that has not been addressed adequately in the literature. The following strategy is the first method we will employ to begin determining cloud shadows. The following discussion is only meant to provide an idea of the approach. Further work in this area has been initiated on this problem.

A 3×3 median filter first is applied to reduce noise in the image. It has the following desirable properties: (1) it does not affect the presence or position of the shadow edges, (2) no new brightness values are created, and (3) performance of the Laplace of Gaussian (LOG) zero crossing edge detection algorithm is improved.

4.1.7.1.1. Oceans. Histogram equalization of the AVHRR channel 1 image is made first. The histogram equalization transform produces a histogram that is quasi-uniform on the average. It is based upon the discrete cumulative histogram with quantized brightness values. The dark values on the histogram equalized image are those due to cloud shadows.

4.1.7.1.2. Land. The algorithm over land is more complex because shadows may fall upon both land surfaces of varying reflectances as well as water surfaces such as lakes, rivers, and marshes. The application of a Laplacian filter to a Gaussian filter image is made first. This operation aids in the recognition of shadow and cloud regions. Edge locations are determined by the zero-crossings of the LOG-filtered image. Details are given in Berendes et al. (1992). Many more edges are produced than just cloud and shadow ones. These are due to background variations and to noise.

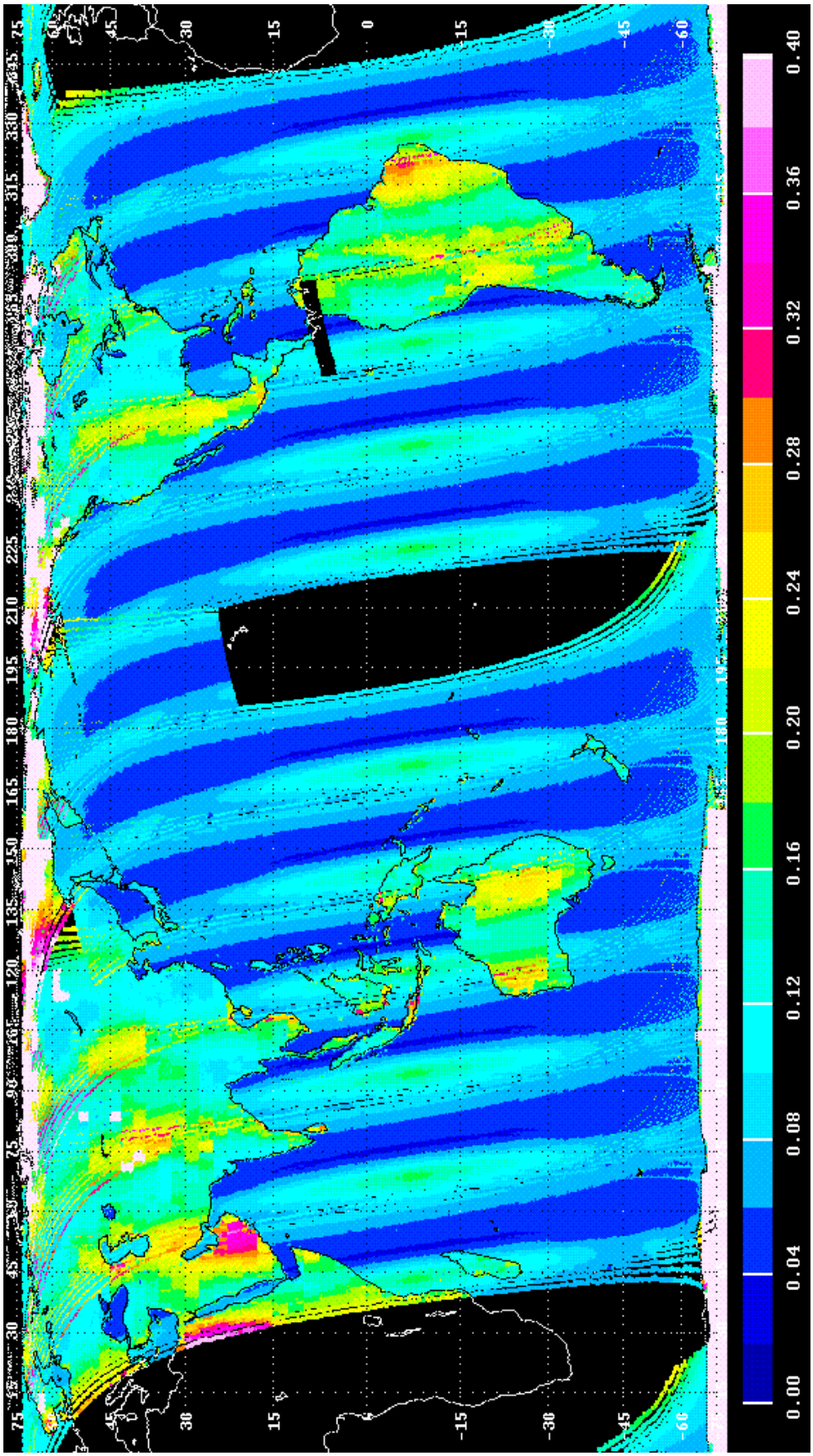


Figure 4.1-7. ISCCP clear-sky reflectance, Oct. 22, 1986. (See section 4.1.6.1.)

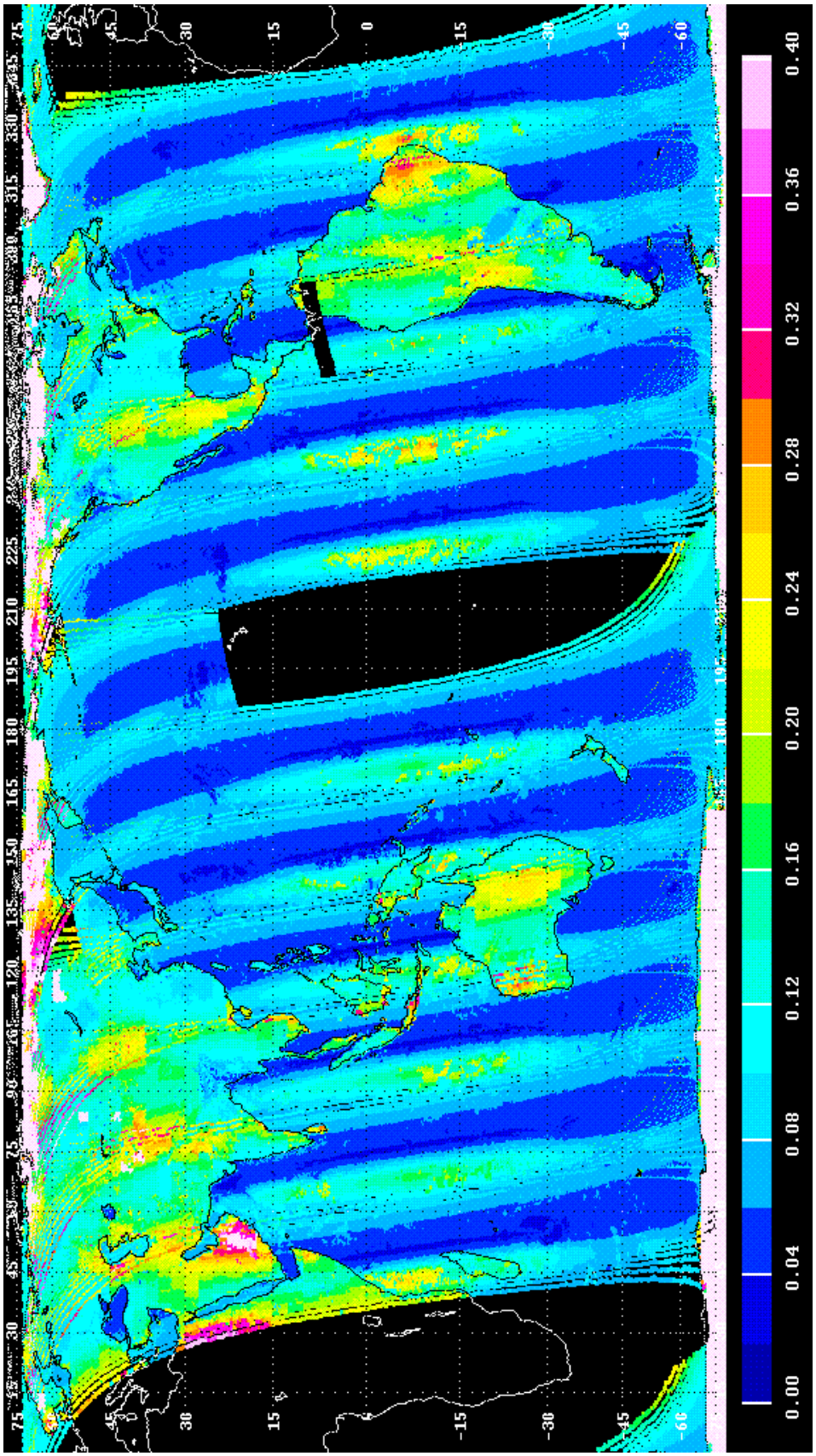


Figure 4.1-8. Revised clear-sky reflectance, Oct. 22, 1986. (See section 4.1.6.1.)

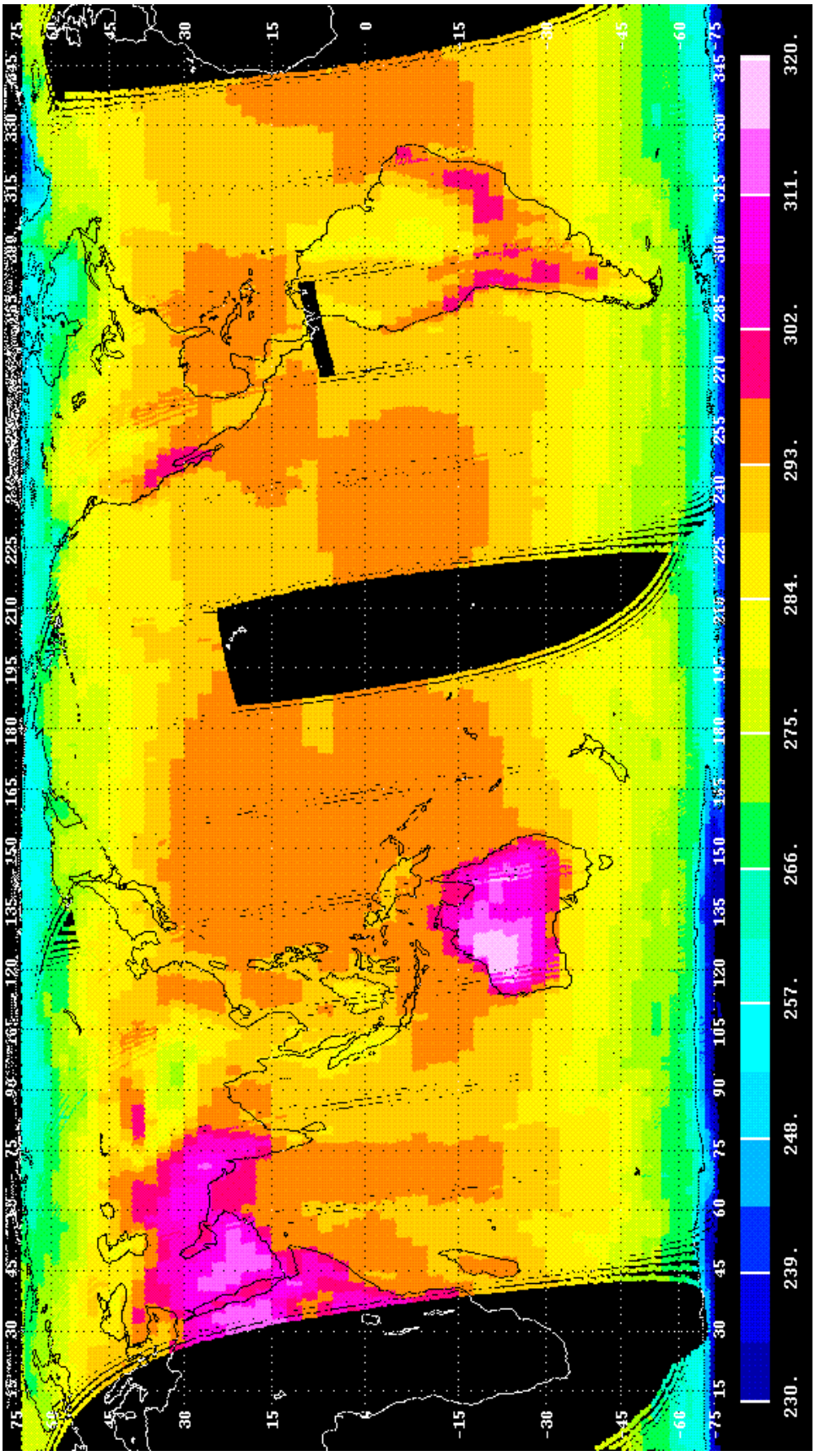


Figure 4.1-9. ISCCP clear-sky temperature (K), Oct. 22, 1986. (See section 4.1.6.1.)

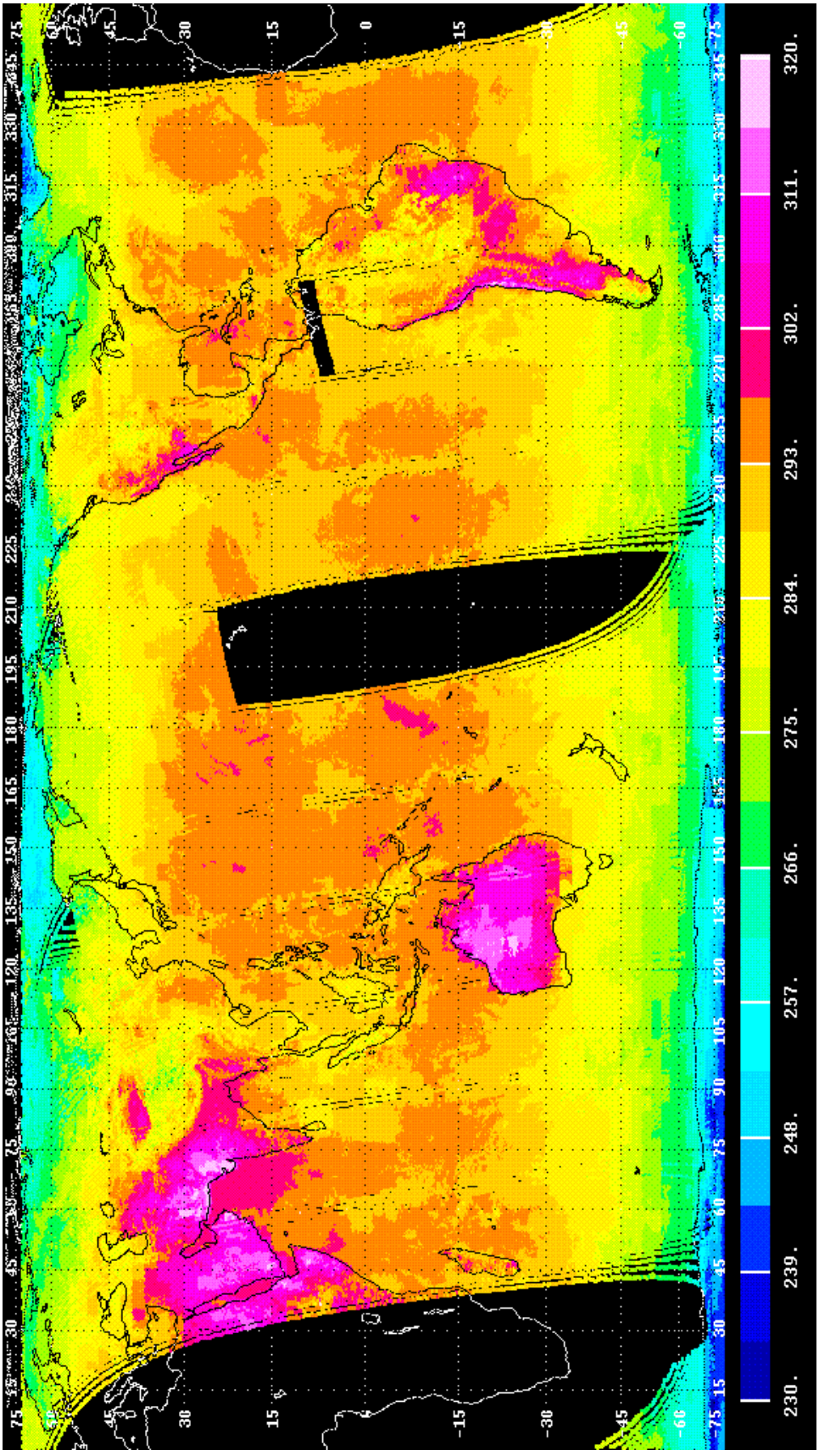


Figure 4.1-10. Revised clear-sky temperature (K), Oct. 22, 1986. (See section 4.1.6.1.)

To isolate the relevant shadow (or cloud) edge pixels, a thresholding procedure is used which is based on a restricted histogram, called the Max/Min histogram. This is constructed from the 3×3 neighborhood surrounding the potential edge pixels. The intention is to capture the modes of the transition pixels generating the edge elements. Generally, there are three distinctive peaks, due to (1) shadows, (2) background, and (3) clouds.

A weighted averaging of the peak values of the Max/Min histogram is used to determine the appropriate threshold between shadow and background (and between cloud and background). This is accomplished by taking into account the size of the distributions. The procedure is iterated to convergence. When water is present in the scene, then a four-mode Max/Min histogram is produced. The same iterative procedure is used to eliminate the background pixels, retaining both shadow and water pixels. The ancillary percent water data set is used to identify probable regions of water.

4.1.7.2. Nighttime Polar Classification

Nighttime polar cloud/surface classification is an extremely difficult problem. Yamanouchi et al. (1987) describe a nighttime polar (Antarctic) cloud/surface discrimination algorithm based upon brightness temperature differences between the AVHRR 3.7- and 10.8-μm channels and between the 10.8- and 12-μm channels. Their cloud/surface discrimination algorithm was more effective over water surfaces than over inland snow-covered surfaces. A number of problems arose over inland snow-covered surfaces. First, the temperature contrast between the cloud and snow surface became especially small, leading to a small brightness temperature difference between the two infrared channels. Second, the AVHRR channels are not well-calibrated at extremely low temperatures (<200 K). As noted in their study, the temperature resolution of channels 4 (10.8-μm) and 5 (12-μm) are approximately 0.6 K at 180 K, while the temperature resolution of channel 3 (3.7-μm) is about 3.5 K at 220 K, and only 7.5 K at 210 K. Therefore, the channel 3 data are not generally useful for cloud detection at the low temperatures expected at the Antarctic. Additionally, the AVHRR data have a digitization problem at extremely low temperatures due to mechanical noise and also because of the nonlinear temperature dependence of the Planck function. The brightness temperature differences between AVHRR channels 4 and 5 offer the most hope for discriminating clouds from a snow- or ice-covered surface. Much further work needs to be done in this area.

4.1.8. References

Allen, Robert C., Jr.; Durkee, Philip A.; and Wash, Carlyle, H. 1990: Snow/Cloud Discrimination With Multispectral Satellite Measurements. *J. Appl. Meteorol.*, vol. 29, pp. 994–1004.

Berendes, Todd; Sengupta, Sailes K.; Welch, Ron M.; Wielicki, Bruce A.; and Navar, Murgesh 1992: Cumulus Cloud Base Height Estimation From High Spatial Resolution Landsat Data—A Hough Transform Approach. *IEEE Trans. Geosci. & Remote Sens.*, vol. 30, no. 3, pp. 430–443.

Brooks, D. R.; Harrison, E. F.; Minnis, P.; Suttles, J. T.; and Kandel, R. S. 1986: Development of Algorithms for Understanding the Temporal and Spatial Variability of the Earth's Radiation Balance. *Rev. Geophys.*, vol. 24, pp. 422–438.

Chen, D. W.; Sengupta, S. K.; and Welch, R. M. 1989: Cloud Field Classification Based upon High Spatial Resolution Textural Features. Part II—Simplified Vector Approaches. *J. Geophys. Res.*, vol. 94, pp. 14749–14765.

Coakley, J. A., Jr.; and Bretherton, F. P. 1982: Cloud Cover from High-Resolution Scanner Data—Detecting and Allowing for Partially Filled Fields of View. *J. Geophys. Res.*, vol. 87, pp. 4917–4932.

Davis, P.; Stowe, L. L.; and McClain, E. P. 1993: Development of a Cloud Layer Detection Algorithm for the Clouds From AVHRR (CLAVR) Phase II Code. *Proceedings of SPIE Symposium*, SPIE.

Ebert, Elizabeth 1987: A Pattern Recognition Technique for Distinguishing Surface and Cloud Types in the Polar Regions. *J. Climat. & Appl. Meteorol.*, vol. 26, pp. 1412–1427.

Ebert, Elizabeth 1989: Analysis of Polar Clouds From Satellite Imagery Using Pattern Recognition and a Statistical Cloud Analysis Scheme. *J. Appl. Meteorol.*, vol. 28, pp. 382–399.

Garand, Louis 1988: Automated Recognition of Oceanic Cloud Patterns. Part I—Methodology and Application to Cloud Climatology. *J. Climat.*, vol. 1, pp. 20–39.

Giarratano, Joseph C.; and Riley, Gary 1989: *Expert Systems: Principles and Programming*. PWS-Kent Publ. Co.

Gutman, G.; Tarpley, D.; and Ohring, G. 1987: Cloud Screening for Determination of Land Surface Characteristics in a Reduced Resolution Satellite Data Set. *Int. J. Remote Sens.*, vol. 8, pp. 859–870.

Haralick, R. M.; Dinstein, I.; and Shanmugam, K. 1973: Textural Features for Image Classification. *IEEE Trans. Syst., Man, & Cybern.*, vol. SMC-3, pp. 610–621.

Hecht-Nielsen, Robert 1990: *Neurocomputing*. Addison-Wesley Publ. Co.

Inoue, Toshiro 1987: A Cloud Type Classification with NOAA 7 Split-Window Measurements. *J. Geophys. Res.*, vol. 92, pp. 3991–4000.

Inoue, Toshiro 1989: Features of Clouds Over the Tropical Pacific During Northern Hemispheric Winter Derived From Split Window Measurements. *J. Meteorol. Soc. Japan*, vol. 67, pp. 621–637.

Key, J.; and Barry R. G. 1989: Cloud Cover Analysis With Arctic AVHRR Data. I—Cloud Detection. *J. Geophys. Res.*, vol. 94, pp. 18521–18535.

Kidwell, K. B., 1991: NOAA Polar Orbiter User's Guide, NOAA Climate Data Center, Satellite Data Services Division, Washington, D. C.

Kriebel, K. T. 1978: Measured Spectral Bidirectional Reflection Properties of Four Vegetated Surfaces. *Appl. Opt.*, vol. 17, pp. 253–259.

Lee, Jonathan; Weger, Ronald C.; Sengupta, Sailes K.; and Welch, Ronald M. 1990: A Neural Network Approach to Cloud Classification. *IEEE Trans. Geosci. & Remote Sens.*, vol. 28, pp. 846–855.

Luger, George F.; and Stubblefield, William A. 1989: *Artificial Intelligence and the Design of Expert Systems*. Benjamin/Cummings Publ. Co.

Matthews, Elaine; and Rossow William B. 1987: Regional and Seasonal Variations of Surface Reflectance From Satellite Observations at 0.6 micron. *J. Climat. & Appl. Meteorol.*, vol. 26, pp. 170–202.

McClain, E. P. 1993: *Evaluation of CLAVR Phase-I Algorithm Performance—Final Report.* U.S. Department of Commerce/ NOAA/NESDIS, Report 40-AANE-201-424.

Minnis, P.; and Harrison, E. F. 1984a: Diurnal Variability of Regional Cloud and Clear-Sky Radiative Parameters Derived From GOES Data. Part I—Analysis Method. *J. Climat. & Appl. Meteorol.*, vol. 23, pp. 993–1011.

Minnis, P.; and Harrison, E. F. 1984b: Diurnal Variability of Regional Cloud and Clear-Sky Radiative Parameters Derived from GOES Data. Part II—November 1978 Cloud Distributions. *J. Climat. & Appl. Meteorol.*, vol. 23, pp. 1012–1031.

Minnis, P.; and Harrison, E. F. 1984c: Diurnal Variability of Regional Cloud and Clear-Sky Radiative Parameters Derived From GOES Data. Part III—November 1978 Radiative Parameters. *J. Climat. & Appl. Meteorol.*, vol. 23, pp. 1032–1051.

Minnis, Patrick; Harrison, Edwin F.; and Gibson, Gary G. 1987: Cloud Cover Over the Equatorial Eastern Pacific Derived From July 1983 International Satellite Cloud Climatology Project Data Using a Hybrid Bispectral Threshold Method. *J. Geophys. Res.*, vol. 92, pp. 4051–4073.

Rossow, William B. 1989: Measuring Cloud Properties From Space—A Review. *J. Climat.*, vol. 2, pp. 201–213.

Rossow, William B.; and Garder, Leonid C. 1993: Cloud Detection Using Satellite Measurements of Infrared and Visible Radiances for ISCCP. *J. Climat.*, vol. 6, no. 12, pp. 2341–2369.

Rumelhart, D. E.; Hinton, G.; and Williams, R. 1986: Learning Internal Representations by Error Propagation. In *Parallel Distribution Distributed Processing—Exploration in the Microstructure of Cognition*. D. Rumelhart and J. McClelland, eds., MIT Press, pp. 318–362.

Saunders, R. W.; and Kriebel K. T. 1988: An Improved Method for Detecting Clear Sky and Cloudy Radiances From AVHRR Data. *Int. J. Remote Sens.*, vol. 9, pp. 123–150.

Seze, Genevieve; and Desbois, Michel 1987: Cloud Cover Analysis From Satellite Imagery Using Spatial and Temporal Characteristics of the Data. *J. Climat. & Appl. Meteorol.*, vol. 26, pp. 287–303.

Seze, Genevieve; Rossow, William B. 1991a: Time-Cumulated Visible and Infrared Radiance Histograms Used as Descriptors of Surface and Cloud Variations. *Int. J. Remote Sens.*, vol. 12, pp. 877–920.

Seze, Genevieve; and Rossow, William B. 1991b: Effects of Satellite Data Resolution on Measuring the Space/Time Variations of Surfaces and Clouds. *Int. J. Remote Sens.*, vol. 12, pp. 921–952.

Stowe, L. L.; McClain, E. P.; Carey, R.; Pellegrino, P.; and Gutman, G. G. 1991: Global Distribution of Cloud Cover Derived From NOAA/AVHRR Operational Satellite Data. *Adv. Space Res.*, vol. 11, no. 3, pp. 51–54.

Stowe, L. L.; Vemury, S. K.; and Rao, A. V. 1994: AVHRR Clear Sky Radiation Data Sets at NOAA/NESDIS. *Adv. Space Res*., vol. 11, pp. 113–116.

Suttles, J. T.; Green, R. N.; Minnis, P.; Smith, G. L.; Staylor, W. F.; Wielicki, B. A.; Walker, I. J.; Young, D. F.; Taylor, V. R.; and Stowe, L. L. 1988: *Angular Radiation Models for Earth-Atmosphere System. Volume I: Shortwave Radiation*. NASA RP-1184.

Tarpley, J. D. 1979: Estimating Incident Solar Radiation at the Surface From Geostationary Satellite Data. *J. Appl. Meteorol.* vol. 18, pp. 1172–1181.

Tovinkere, V. R.; Penaloza, M.; Logar, A.; Lee, J.; Weger, R. C.; Berendes, T. A.; and Welch, R. M. 1993: An Intercomparison of Artificial Intelligence Approaches for Polar Scene Identification. *J. Geophys. Res.*, vol. 98, no. D3, pp. 5001–5016.

Welch, R. M.; Sengupta, S. K.; Goroch, A. K.; Rabindra, P.; Rangaraj, N.; and Navar, M. S. 1992: Polar Cloud and Surface Classification Using AVHRR Imagery—An Intercomparison of Methods. *J. Appl. Meteorol.*, vol. 31, no. 5, pp. 405–420.

Welch, Ronald M.; Kuo, Kwo-Sen; and Sengupta, Sailes K. 1990: Cloud and Surface Textural Features in Polar Regions. *IEEE Trans. Geosci. & Remote Sens.*, vol. 28, pp. 520–528.

Welch, R. M.; Navar, M. S.; and Sengupta, S. K. 1989: The Effect of Spatial Resolution Upon Texture-Based Cloud Field Classifications. *J. Geophys. Res.*, vol. 94, pp. 14767–14781.

Yamanouchi, Takashi; Kawaguchi, Sadao; and Suzuki, Kazuya 1987: Detection of Clouds in Antarctica From Infrared Multispectral Data of AVHRR. *J. Meteorol. Soc. Japan*, vol. 65, pp. 949–962.